Innovation on Job Accessibility with General Transit Feed Specification (GTFS) Data

Final Report
METRANS Project 16-08

December 2018

Principal Investigator: Dr. Gary Painter
Co-Principal Investigator: Dr. Marlon Boarnet
Madison Swayne & Matthew Miller

Sol Price School of Public Policy
University of Southern California
Los Angeles, California 90089

METRANS Transportation Center
USC | CSULB

Disclaimer: The contents of this report reflect the views of the authors, who are responsible for the accuracy of the data and information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, the California Department of Transportation and the METRANS Transportation Center in the interest of information exchange. The U.S. Government, the California Department of Transportation, and the University of Southern California assume no liability for the contents or use thereof. The contents do not necessarily reflect the official views or policies of the State of California, USC, or the Department of Transportation. This report does not constitute a standard, specification, or regulation.

TABLE OF CONTENTS

ABSTRACT

This paper investigates job access in Los Angeles County via public transportation. We develop a measure of accessibility based on travel time, via transit, to jobs. We use open source General Transit Feed System (GTFS) data to model the transit network and interface with a variety of computing tools including JavaScript, ArcPy for ArcGIS and statistical packages to measure the number of jobs accessible from these tracts within a 60-minute commute time. The result is a method that leverages GTFS data to allow us to simulate job access with different transit investments. We measure transit job access from these tracts for three scenarios. Each scenario compares transit job access with completed or proposed additions to the Los Angeles Metro network relative to the status quo. We illustrate that job access measures can go beyond the common "static" measures of access over existing networks and be used to evaluate different transit investment scenarios. Our results suggest that the opening of the Expo Phase II light rail line increased job access in tracts that contain a higher percentage of opportunity youth populations and the opening of the proposed Crenshaw Extension promises to increase transit job access even further.

LIST OF FIGURES AND TABLES

Figures

Tables

ACKNOWLEDGEMENTS

We are grateful to Parth Kundaliya for his assistance in data management and code design. His help greatly improved the efficiency of our analysis. We are also thankful to Remix for providing us with access to their online platform which was useful in the analysis.

All errors and omissions are the responsibility of the authors.

BACKGROUND

Public transportation is undergoing rapid change in Los Angeles. Recent voter-approved ballot measures are funding new transit across Los Angeles County making questions of access and opportunity particularly salient. The literature first connected job proximity to labor market outcomes with the Spatial Mismatch Hypothesis (SMH), suggested by John F. Kain (1968). Kain argued that residential discrimination constrained African American populations to live in the central city. He hypothesized that the black-white employment gap could be partly explained by the spatial isolation of African American populations from suburban jobs. Since Kain's seminal 1968 work, there have been additional studies investigating the Spatial Mismatch Hypothesis. These studies have clarified that the SMH applies to low-skilled Blacks living in the center city who are spatially isolated from entry-level jobs in the suburbs. Yet the SMH does not apply only to central cities. As an example, Painter and Liu (2012) showed that the suburban population of African Americans is growing. Similarly, the question of access to jobs and the role on labor markets has been broadened to other demographic groups (e.g. Blumenberg and Ong, 2001).

In the 1990s, urban sociologists and regional scientists broadened the conception of the SMH to frame the policy debate on solutions; we call this the geography of opportunity literature. In 1995, George C. Galster and Sean P. Killen of The Urban Institute proposed a "geography of metropolitan opportunity" conceptual framework to understand how regions create unequal "opportunity structures" across the domains of housing, labor, crime, politics, education, and social networks particularly for inner-city youth. Their metropolitan opportunity structure idea involved both a "process" dimension (i.e. markets, institutions, and service delivery systems) and a product (or outcome) dimension (i.e. future streams of income, consumption, and productivity). Thus, they define more "opportunity" for youth as being able to make more "choices about education, fertility, work and crime" within this model (Galster and Killen, 1995: 11). Their main argument and contribution to the equal opportunity debate was that geography matters both conceptually and empirically.

A critical component of understanding the geography of opportunity is a proper measure of job accessibility. There have been some previous studies of accessibility to jobs by transit. Blumenberg and Ong (2001) used a gravity-based model of transportation and census data to estimate the number of accessible jobs by transit. The authors find that in their study of Detroit, the central business district (CBD) does not have the highest job accessibility but that instead, the highest accessibility zones are located several kilometers out from the CBD. Blumenberg (2004) again measured the number of low-wage jobs accessible within a 30-minute commute using transit from seven neighborhoods in Los Angeles using the gravity model developed by Blumenberg and Ong (2001). Other studies of accessibility by transit have used Euclidean distance, which does not account for land use patterns including travel speeds, headways, irregular service frequency, or other inconsistencies to transit access. In geography, the use of mixed methods such as geographic information systems (O'Sullivan, Morrison, & Shearer, 2000), remote sensing, (Delamater, Messina, Shortridge, & Grady, 2012), network analysis (Biba, Curtin, & Manca, 2010; Delamater et al., 2012), and spatial statistics has allowed the development of more nuanced estimations that approximate the mobility of a population in relation to access to various services such as health care providers (Guagliardo, 2004;

Schuurman, Bérubé, & Crooks, 2010; Wan, Zou, & Sternberg, 2012). However, none of these methods have been used to model how realistic changes in the transit network will change job access, and then how those changes can be related to demographic sub-groups. Most of the literature to date has focused on measuring the access provided by the network as it exists. We innovate beyond that by developing a method that can allow comparisons of different changes to the existing network, allowing planners to see accurate calculations about how different infrastructure changes can affect the geography of job access.

We focus on developing access measures based on travel time while using public transit, not car, because many poor people take transit out of necessity. This is especially true for urban youth who seek entry-level jobs; they often will not have access to cars. Our work is a granular investigation into measuring accessibility to jobs using transit. We use transit system data and measures of transit travel time to measure the number of jobs that can be accessed using transit (with walking as the mode of station access-egress.) Our use of Remix and General Transit Feed Specification (GTFS) data provides us with the best and most widely available representation of the transit network and operation, and so relying on it in studies of accessibility by transit provides a sound measure for access and one that can be applied in a large, and growing, number of urban areas.

Background: Rail Transit in Los Angeles

Since 1980, Los Angeles voters have approved four half-cent sales tax measures, with the largest amount of those funds supporting transit investment and operations. Since the opening of the Blue Line in 1990, the rail transit network operated by the Los Angeles County Metropolitan Transportation Authority (L.A. Metro) has opened 93 rail transit stations. The system has six lines, which include at-grade light rail (Blue line, Expo line), light rail in the median of a freeway (Green line), heavy rail subway (Purple and Red lines), and light rail that is a combination of at-grade and freeway median (Gold line). Figure 1 shows the network and lines.

The rail construction project in Los Angeles accelerated and expanded with passage of the most recent half-cent sales tax measures, Measure R in 2008 and Measure M in 2016. Measure R included funds for a project to extend the Expo Line light rail from its terminus in Culver City to a new terminus in Santa Monica. The project extended the 8.5 mile line an additional 6.6 miles (Hymon, 2016). This extension known as Expo Phase II was the first rail transit to reach the westside of Los Angeles since 1953 (Hymon, 2016). Expo Phase II opened May 20, 2016 and now links two large job centers in Los Angeles — Santa Monica and the historic Downtown — with stops at many neighborhoods in between.

In November 2016, a majority of Los Angeles, California voters again approved a ballot initiative to fund transit investment. Voters approved a permanent sales tax increase that will raise $120 billion over the next 40 years for improvement to the public transit system (Elkind, 2017). This ballot measure, Measure M, includes transportation infrastructure projects across Los Angeles County including improvements to buses, rail service, and roads. Included in the Measure are provisions to build large-scale projects including the Metro Crenshaw/LAX Line and the Vermont Avenue Bus Rapid Transit.

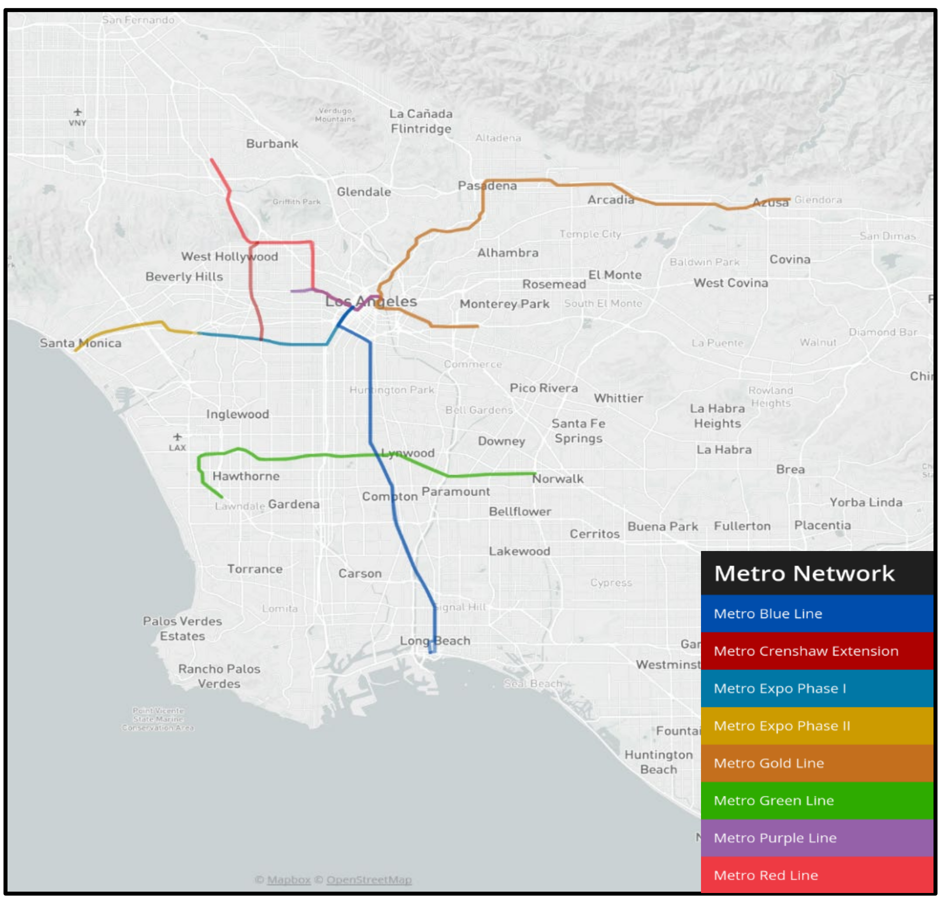
We chose those three projects – Expo Phase II, the Crenshaw line, and the Vermont Bus Rapid Transit (BRT) corridor, for analysis. Note that Expo Phase II is the most recently opened light rail line in Los Angeles, and the Crenshaw light rail and Vermont BRT are notable proposed or (for Crenshaw) under construction projects.

Since Los Angeles is undergoing rapid changes to its public transit system, it is important for research and planning to understand the impacts this new system will have to accessibility and labor markets. In this project, we use an online transit mapping platform known as Remix[1] to model changes to the transit network and their subsequent impact on job accessibility. While previous studies have measured access to jobs by transit, we expand upon these studies by modeling the future transit network. We analyze the following transit scenarios:

1. Current Transit System **without** Expo Phase II
2. Current Transit System (includes Expo Phase II)
3. Current Transit System + Crenshaw Extension

The Metro rail network is shown below (Figure 1) with Expo Phase II and the Crenshaw Extension included in the network.

Figure 4 - Metro Rail Network with Expo Phase II & Crenshaw Extension Shown



[1] Remix is a private, online platform which uses real-world data for transit planning. A full explanation of our interaction with Remix and its tools is presented in the "Methods" section of this report.

Our Contribution

This research is, to our knowledge, the first of its kind to develop a simple metric that can forecast how changes to the transit network will affect job accessibility, using transit network data that are available for a large number of cities. This work relies on actual travel times across the transit network instead of estimated distances and uses new open source data to understand the effect of changes to the network. Previous studies have had to rely on the existing network, without simulating the job-access impacts of changes to that network. Further, the automated tools developed in this project can be applied to different geographies in the future and provide us with data on accessibility across the United States and beyond. In this report, we focus on developing and illustrating the method.

DATA

Open Source Transit Data

Remix relies on General Transit Feed Specification (GTFS) data to map the existing transit network. GTFS data is public transportation schedule and geographic information that have been assembled in a common format for legibility. Public transit agencies have the ability to publish their transit data so that developers, and researchers, can interface with the GTFS database (Google Transit APIs, 2018). GTFS is a central database of past and present transit stations, stops, and schedules (Rodyansky, 2018). In Remix, we have the ability to augment GTFS by drawing new transit lines, including new stops and stations, into the existing GTFS network. This feature is built into the Remix platform, where users draw lines and designate stops along the line. We can also customize the schedule for each new line that is drawn.

Employment Data

Employment data used in this study came from the U.S. Census Bureau's Longitudinal Employment Household Dynamics (LEHD) survey. The LEHD survey reports on a number of workforce characteristics including quarterly workforce indicators and origin-destination employment statistics (LODES). To measure job accessibility, we relied on 2015 LODES Workplace Area Characteristics (WAC) data. The data provides job totals, broken down into job types based on NAICS sector codes. We used counts of all jobs per census tract, designated in the data as "C000."

Opportunity Youth Data

To focus the analysis on distinct locations with large numbers of transit dependent persons, we focused on tracts that contain large populations of Opportunity Youth. Opportunity youth – sometimes referred to as "not employed, not in education, not in training" or NEETs – are defined as people aged 16 to 24 who are not in school and not employed. Such youth are a unique population at a critical time in their lives for shaping future life trajectory (Leon, Choi & Rosen, 2018). These opportunity youth face higher risk of being isolated from social activity, dependence on public assistance, and engagement with criminal activity (Leon, Choi, & Rosen, 2018; Belfield, Levin & Rosen, 2012; Besharov & Gardinier, 1998). Using ACS 2011-2015 data we estimated the percent of opportunity youth per census tract using the methodology described by Leon, Choi & Rosen (2018) and we identified the top five tracts with the largest proportion of opportunity youth in the City of Los Angeles. The locations of the top five tracts with the highest percentage of opportunity youth, among persons in the tract aged 16 to 24, are presented below in Figure 2. Table 1 shows the percent of the 16-24 year-old population in these top five tracts who qualify as opportunity youth.

Figure 5 - Top Five Tracts by Percent Opportunity Youth for Those Aged 16-24

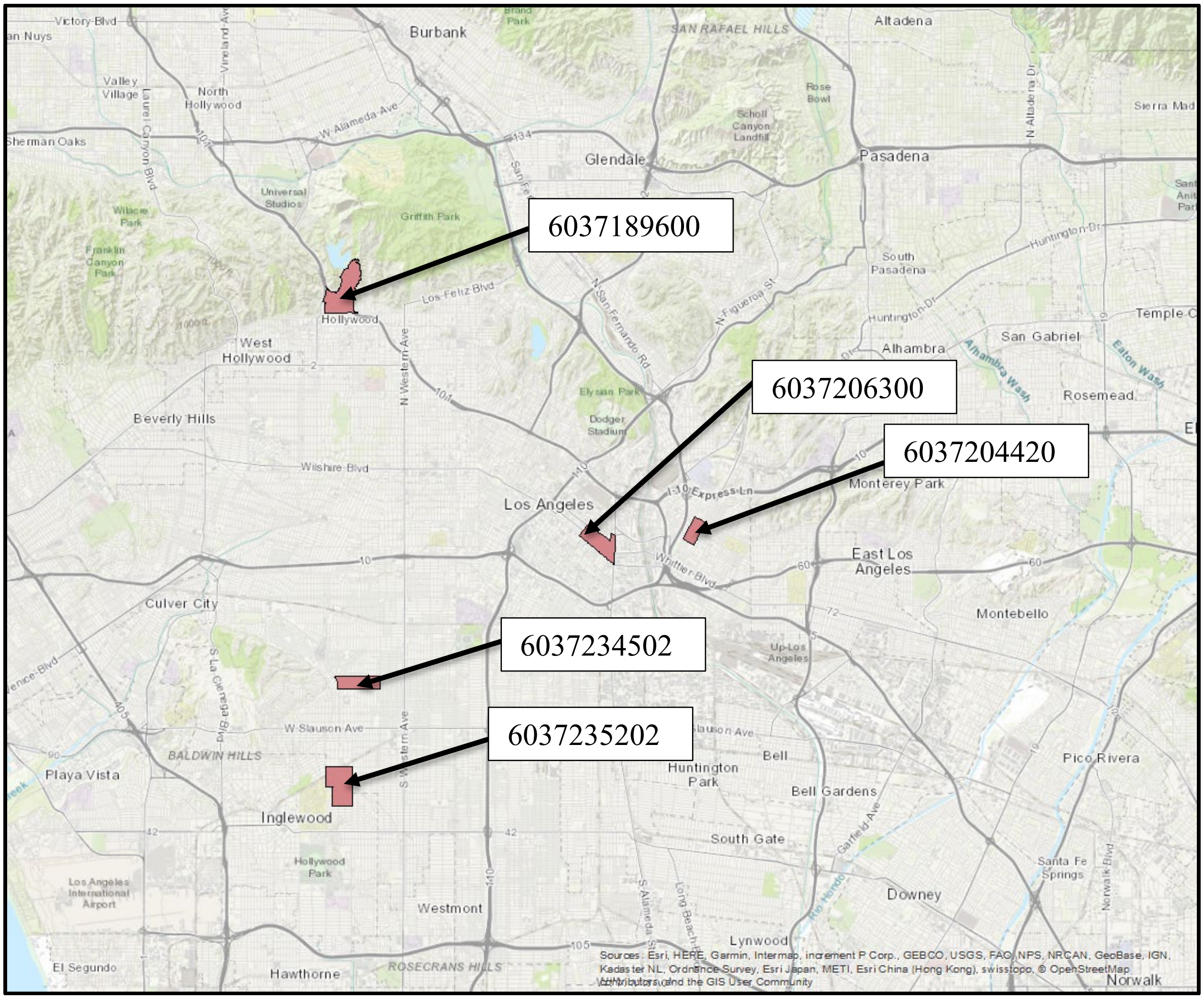


Table 3 - Top Five Tracts by Percent Opportunity Youth for Those Aged 16-24

Tract Number	Percent of Population Aged 16-24 who Qualify as Opportunity Youth
6037189600	64.1%
6037206300	58.8%
6037234502	41.9%
6037235202	39.3%
6037204420	37.1%

METHODS

Automated Tools – Summary

To assess job accessibility, we have created the 'Transit Access to Jobs' tool. This tool uses Remix, a private web-based application, to calculate travel times for public transit (bus, rail, and light rail) and walking modes in different commute sheds (15, 30, 45, and 60 minutes), using General Transit Feed Specification (GTFS) data. Remix can adapt to schedule and route changes, additions, and deletions. Remix then allows users to place "Jane" at any location within the transit network and an access isochrone is generated. The access isochrone is the distance (shown on a map as an area) that a person can travel from "Jane's" location (a point in space) within 15, 30, 45, or 60 minutes via transit, with walking access/egress to/from stations. Users can drag and drop Jane downloading isochrones one at a time. Through a customized JavaScript macro we developed (Appendix 1), we have automated the Jane tool in Remix to quickly produce access isochrones from the centroid of census tracts in Los Angeles County. Outside of the Remix tool, we then match the travel isochrones to U.S. Bureau of Labor Statistics' LODES employment data by census tract using an ArcpPy macro in ArcGIS. The ensuing results display an index of job accessibility by transit or walking modes from specific tracts in Los Angeles County. This method provides unparalleled insight into transit and walk mode travel from any tract within Los Angeles County. We have created our tool such that it can be updated to simulate future transit buildouts associated with future Los Angeles County investments in public transit infrastructure.

Remix Summary

Remix is a private, web-based transit mapping software which relies on GTFS data to map the existing transit network in an online platform. Users have the ability to edit the transit network by adding or subtracting lines and making adjustments to existing transit schedules. Remix also provides quick data overlays with sociodemographic information to allow planners to visualize their communities. We used the Remix tool to map the transit network in Los Angeles County including all bus and rail lines to model Angelenos' transit options. [2] In Remix, commuters are routed over the fastest route using any combination of bus and rail with no distinction between the two.

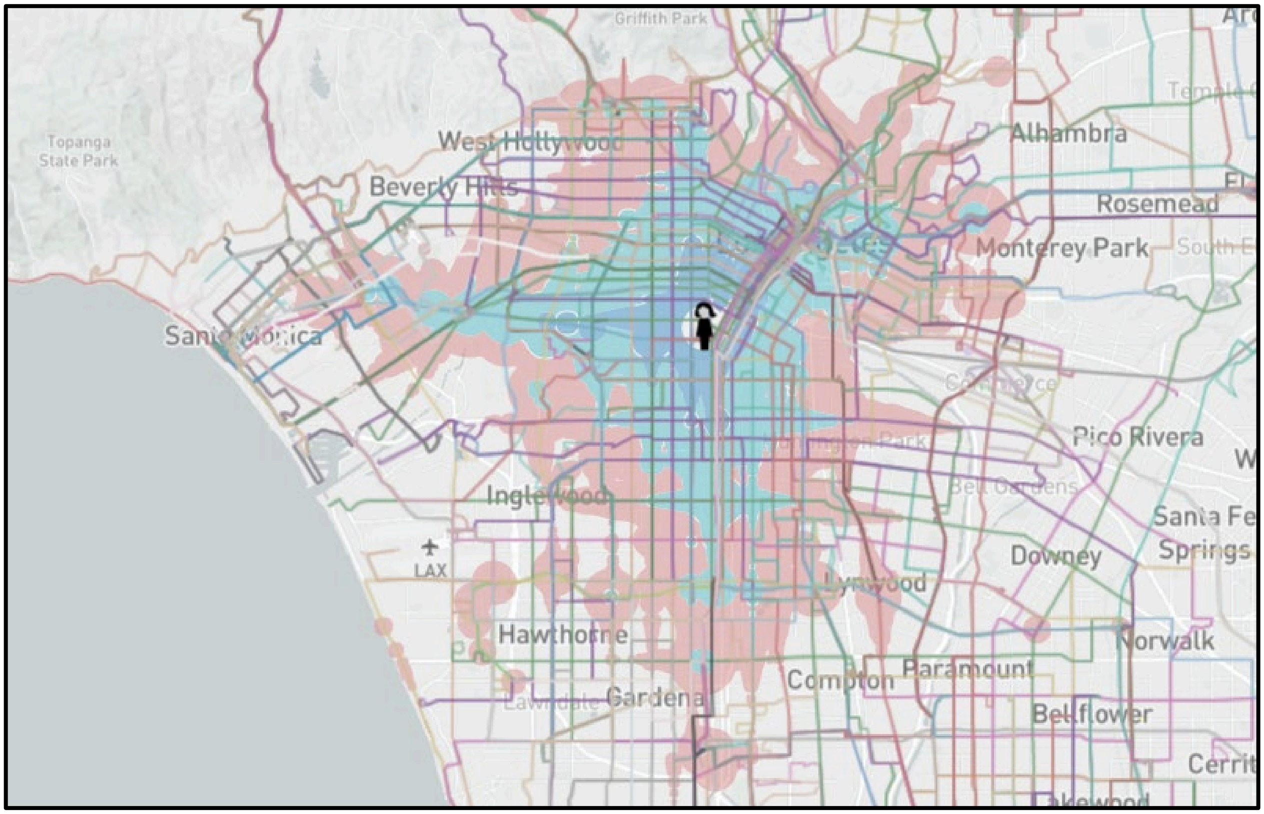
Jane in Remix

As part of the Remix transit mapping interface, users have access to a tool known as "Jane." The Jane Tool is used to show how far someone can travel using the existing transit network and walking for station access/egress. To use the tool, the user first uploads all relevant transit routes to the online map to build the network. Next, the user drags and drops Jane at any location of interest on the network map. Users can drop Jane at a street intersection, at their house, or any X-Y coordinate pair. After Jane is dropped at a location, Remix generates an isochrone[3] showing how far Jane can travel in 15, 30, 45, and 60 minutes. In this access calculation, Remix makes some assumptions. Built into the tool is the assumption that Jane will wait half of the headway time before boarding a bus or train. If a bus has a headway time of 30 minutes, Jane will wait 15 minutes before boarding. If Jane has to transfer between lines, she will again wait half of the headway time before boarding again. Between stations or after getting off the bus or train, Jane walks at a pace of 3 miles per hour. As Jane is moved across the city and throughout the network, the user can download access isochrones for each location. These isochrone shapefiles provide a good representation of what access is provided by the transit service and can be compared from one location to the next. An example set of isochrones from one location is shown in Figure 3 below.

[2] Transit authorities included in the Remix Los Angeles network, with number of lines in parentheses: LA Metro Bus (139), City of Commerce Municipal Bus Lines(11), LA Metro Rail (6), Metrolink (7), Beeline (13), Santa Monica Big Blue Bus (20), GTrans (5) Pasadena Transit(10), Torrance Transit (11), Culver City Bus (8), UCLA (8), Norwalk Transit System (6), Carson Circuit (8), LADOT (45), Foothill Transit (37), Long Beach Transit (35), Santa Clarita Transit (42), Antelope Valley Transit Authority (20), Orange County Transit Authority (82)

[3] Isochrones are defined as, "a line on a map connecting points of equal elapsed time; especially, travel time to or from a given location" (Esri GIS Dictionary, 2018).

Figure 3 - Example of Jane's Isochrone



In Figure 3, Jane has been placed in Downtown Los Angeles, an area with a high density of transit. Jane's smallest isochrone, shown in white, is the area she can reach within 15 minutes via transit. The blue isochrone is at the area that can be reached within 30 minutes via transit, green is the area that can be reached in 45 minutes by transit, and red is the area that can be reached in 60 minutes when traveling by transit.

Automating Jane

Remix allows users to drop Jane at any location on the map to generate her transit access isochrones. A Remix user has the ability to visualize and download transit travel isochrones for anywhere in the city. We have exploited this capability and through code, have generated isochrones for multiple tracts in different transit scenarios.

In order to quickly generate transit access isochrones for each of our tracts of interest in Los Angeles County, we developed a JavaScript macro to automatically fetch and download Jane's isochrone shapefile from Remix. Remix generates 15, 30, 45, and 60 minute isochrones for Jane, based on her location on the map using the available transit in her location. Moving Jane to each of the tracts under study is necessary as her location automatically becomes the center of the isochrones. The Remix platform relies on X and Y coordinates to place Jane on the map. Each time Jane is moved across the map, the Remix URL changes in a predictable manner to reflect Jane's new location [4].

Because of the data demands of Remix and multiple downloads, the JavaScript macro has been set on a timer to make server requests every 15 seconds. While this timer slows the processing down, it is important so as not to compromise internet servers and ensures that each set of four isochrones (15, 30, 45, and 60 minutes) is downloaded at an average rate of 15 seconds per tract. The macro has also been programmed to restart automatically if the server returns an error, eliminating the need for a human to watch the code run.

This JavaScript macro was used for looping through each of the census tracts in our investigation. The macro however can quickly be edited to perform the loop through other census tract centroids or to perform the query to Remix on a different transit map.

Arcpy/GIS scripting

After collecting transit isochrones for each of the tracts, the resultant shapefiles were then processed in ArcGIS by way of an ArcPy script. This ArcPy script uses the Census TIGER Line census tract shapefile as an input feature. The script then loops through each of the access isochrones using each as the clip feature. This process uses the isochrone as a "cookie cutter" on the census tracts creating a new feature class with the clipped area of each tract. For example, a census tract may have an initial area of 1000 square feet but once clipped by the access isochrone will have an area of 500 square feet. In other words, only half of the census tract may be accessible by transit. Knowing which proportion of each tract is accessible is essential for jobs calculations.

Using R for Final Analysis

Each of the .csv output files from the ArcGIS processing was then transferred into R. When Jane is dropped at the centroid of a tract, she may have access to a number of adjacent tracts using transit. The ArcPy script calculates which portion of the adjacent tracts Jane can reach within the 15-30-45-60 minute travel times. These proportional areas then have to be translated into job counts. The R-script calculates the portion of each census tract Jane can reach within unit travel time, and multiplies it by the jobs available in that census tract. It repeats the entire process for all census tracts with one census tract as the starting point in each iteration. It then merges data for all census tract in one data frame and calculates the sum (sum total of jobs a person in the census tract can reach within unit travel time). Finally, the R-script exports the data frame to a CSV file for final summary and analysis.

RESULTS

We have performed the above methodology on the five tracts in Los Angeles County selected based on the highest percentage of opportunity youth populations. The table below summarizes

[4] A complete copy of the Remix macro is provided in Appendix I. With this macro, we are able to download REMIX travel-time isochrones from each census tract. While we illustrate the results with only a small number of census tracts, we have verified that the macro can automate the process of placing Jane at the centroid of all Los Angeles County census tracts and downloading the resulting isochrones. We note that because REMIX is a private tool, we purchased a REMIX license – a necessary part of the method that we developed here.

the number of additional jobs accessible from each of the five tracts with the addition of Expo Phase II and the Crenshaw Extension to the network. All numbers presented in Table 2 below represent job access in a 60-minute commute time from the centroid of each tract. We chose 60 minutes to illustrate the tool. A 60-minute travel time is larger than average commute times by car in many metropolitan areas, and we chose the 60-minute isochrones because that illustrates the largest access changes. We caution that those access changes may not be the most reasonable representations of job access, and a more complete analysis of job access would include changes for smaller travel time isochrones. We also note that we did not model possible reductions in service from the removal of nearby bus service, but it is possible to do so.

Table 4 - Number of Jobs Accessible, Top 5 Opportunity Youth Tracts

	Number of Jobs Accessible in 60 Minutes from Five Tracts with Largest Opportunity Youth Population					
Tract Number	**Scenario 1. Current without Expo Phase II**	**Change from Scenario 1 to Scenario 2**	**Scenario 2. Current with Expo Phase II**	**Change from Scenario 2 to Scenario 3**	**Scenario 3. Current with Phase II & Crenshaw**	**Change from Scenario 1 to Scenario 3**
6037206300	1,035,934	16,773	1,052,707	412	1,053,119	17,185
6037234502	587,561	12,904	600,465	27,433	627,898	40,338
6037204420	1,088,012	2,840	1,090,852	344	1,091,196	3,184
6037235202	726,308	18,153	744,461	8,995	753,457	27,149
6037189600	734,812	22	734,834	(0)	734,834	22
Average	**834,525**	**+10,138**	**844,664**	**+7,437**	**852,101**	**+17,576**

As noted in Table 2, the extension of the Expo Line to Santa Monica increased the number of jobs accessible to residents of these tracts by 10,138 or approximately 1.2%. The addition of the Crenshaw line further increases the number of jobs accessible to these tracts in 60 minutes by an additional 7,437. The completion of these two light rail lines increases overall number of jobs accessible by over 2%.

The selection of these tracts also demonstrates that accessibility increases are not homogenous across space. For example, tract 6037234502 is proximate to both the Expo line extension and the proposed Crenshaw line and therefore residents of this tract will experience a 6.7% increase in the number of jobs compared to Scenario 1. Tract 6037235202 is proximate to the Crenshaw line and experiences a 3.6% increase in the number of jobs. Conversely, tracts like 6037189600, which are located far away from the light rail network see almost no gain in accessibility due to these extensions.

IMPLEMENTATION & LIMITATIONS

Limitations

Currently, the work focuses only on 5 tracts, which are illustrative of the kinds of accessibility gains that are available with transit line extensions. Future work will calculate accessibility gains countywide. When this work is complete, we will be able to measure the incremental job access for each tract in Los Angeles County with the addition of a new transit line to the network. This will be important for understanding which tracts see the greatest accessibility benefits from these network improvements. Further, we have relied on the 60-minute isoline for Jane's commute in each of the tracts. We understand that 60 minutes is not an ideal commute and we are not advocating it should be the norm, although it is close to the average public transit commute for many households. Mean travel time to work for persons commuting by public transportation was 47.8 minutes in the 2009 American Community Survey (Table 5, McKenzie and Rapino, 2011). Preliminary work using the 30 minute isoline demonstrated much lower gains in accessibility, which is partially due to the estimated wait times for transit. Earlier research, by Boarnet, Giuliano, Hou, and Shin (2017), demonstrated that in San Diego County wait time, transfer time, and station access-egress time can be approximately half of total transit trip time. We will continue to refine the work so that we can provide job count changes for 15, 30, 45 and 60-minute commute times.

Each of our job counts have counted total jobs and have not been parsed into different job types. We have instead assumed that improved access is inherently better. It is still possible that there is a mismatch between job seeker type and job type even if overall access is improved.

Finally, as we expand the work, we will be able to add additional transit scenarios in line with proposed changes to the Los Angeles transit network. Two such transit scenarios will be current transit system + Vermont Bus Rapid Transit (BRT); and current transit system + Crenshaw Extension + Vermont BRT.

CONCLUSIONS AND RECOMMENDATIONS

Transit access to jobs is important for insuring populations most at need have access to jobs across the city. By using GTFS data, we have modeled the number of jobs accessible from these tracts with large populations of opportunity youth. This modeling provides us with a measure of the impact of new rail infrastructure on job access for this population that is likely more relevant than using linear distance or car commuting times.

From a methodological perspective, combining Remix, JavaScript scripting, ArcPy scripting, and data management in R was useful for setting the groundwork for future analysis. These tools will be useful for continued accessibility research including countywide or citywide access to jobs and other amenities. We believe this study has demonstrated the possibilities offered by GTFS data and ArcPy scripting in ArcGIS in fine-grained investigations of accessibility. Further improvements may still need to be made before these tools can be generalized to other geographic contexts however, these tools are not far from being useful in all geographies with GTFS transit data.

The measured improvements to the transit network in Los Angeles County have had an overall positive impact on the number of jobs which are accessible to the tracts that contain the highest percentage of opportunity youth living in the County. The opening of Expo Phase II added an average additional 10,138 transit-accessible jobs per tract in the top-five opportunity youth tracts and the addition of the Crenshaw Extension is estimated to add 60-minte transit job access to an average of 7,437 additional jobs per tract.

REFERENCES

Belfield, C.R., Levin, H.M. and Rosen, R. (2012). The Economic Value of Opportunity Youth. Corporation for National and Community Service.

Besharov, D.J. and Gardiner, K.N. (1998). Preventing youthful disconnectedness. *Children and Youth Services Review,* 20(9-10), 797-818.

Blumenberg, E., & Ong, P. (2001). Cars, buses, and jobs: Welfare recipients and employment access in Los Angeles. *Journal of the Transportation Research Board,* 1756, 22-31.

Boarnet, Marlon G., Genevieve Giuliano, Yuting Hou, and Eun-Jin Shin. 2017. First/last mile transit access as an equity planning issue. Transportation Research Part A vol. 103, pp. 296-310.

Elkind, Ethan (2017). Metro is spending billions of your tax dollars to build L.A. a world class transit system. Don't let them blow it. *Los Angeles Times.*

Esri (2018). Data Dictionary. https://support.esri.com/en/other-resources/gis-dictionary/search/

Google Transit APIs (2016). GTFS Static Overview. https://developers.google.com/transit/gtfs/

GTFS.org. n.d. Getting Started. http://gtfs.org/getting-started/#gtfs-applications

Hymon, Steve (2016). 17 things to know about Expo 2 opening on May 20. *The Source.* https://thesource.metro.net/2016/02/25/17-things-to-know-about-the-expo-phase-2-opening/

Leon, Choi & Rosen (2018). How Does Agglomeration Shape Youth Opportunity Across the Urban/Rural Divide? Evidence from the Los Angeles Region. Unpublished, under review.

McKenzie, Brian and Melani Rapino. 2011. Commuting in the United States: 2009, American Community Survey Reports. Washington, D.C.: United States Census Bureau. https://www.census.gov/prod/2011pubs/acs-15.pdf.

Rodnyansky, Seva (2018). Do It Yourself: Obtaining Updated Transit Stop and Route Shapefiles in Urban and Nonurban Areas. *Cityscape: A Journal of Policy Development and Research,* 20(1), 205-214.

APPENDIX 1

```
// change these variables as you wish
const config = {
  mapID: '0cfdf83',
  inputFile: 'input.csv',
  outputFolder: 'output',
  urlBuilder: (input) => {
    return {
      pageURL:
`https://platform.remix.com/map/${config.mapID}?latlng=${input.INTPTLAT},${in
put.INTPTLON},z10&jane=34.00144,-118.28568,17,weekday,frequency,demo-line-
jobs,walking`,
      pointsURL: `https://needledrop-
service.remix.com/isochrone_points/${config.mapID}/${input.INTPTLAT}/${input.
```

```
INTPTLON}?hours_since_midnight=17&window_type=weekday&trip_based=false&walkin
g_speed_kph=4.83`,
      shapefileURL: `https://needledrop-
service.remix.com/isochrone_shapefile/${config.mapID}/${input.INTPTLAT}/${inp
ut.INTPTLON}?hours_since_midnight=17&window_type=weekday&trip_based=false&wal
king_speed_kph=4.83`
    }
  }
}

const fs = require('fs')

// 3rd party packages
const csv = require('csvtojson')
const axios = require('axios')
const unzipper = require('unzipper')

script()

async function script () {
  let inputs = await parseInput(config.inputFile).catch(error =>
console.log(error))
  if (!inputs) return 'Script failed'

  for (let input of inputs) {
    if (!alreadyExists(input.INTPTLAT, input.INTPTLON)) {
      console.log(`Fetching file ${inputs.indexOf(input) + 1} at ${new
Date()}`)
      let urls = config.urlBuilder(input)
      let outputURL =
`${config.outputFolder}/${input.INTPTLAT}${input.INTPTLON}`

      let pointsResponse = await axios.get(urls.pointsURL).catch(error =>
console.log(error))
      let points = pointsResponse.data.results

      let shapeFileLocationResponse = await axios.post(urls.shapefileURL, {
points }).catch(error => console.log(error))
      let zipFileURL = shapeFileLocationResponse.data.resultUrl

      let shapeFile = await axios.get(zipFileURL, { responseType:
'arraybuffer' }).catch(error => console.log(error))

      fs.appendFileSync(`${outputURL}.zip`, shapeFile.data)
      console.log('Zip file downloaded, now extracting')

      fs.createReadStream(`${outputURL}.zip`)
        .pipe(unzipper.Extract({ path: outputURL }))
      console.log(`Extracted to ${outputURL} at ${new Date()}`)
    } else {
      console.log(`File ${inputs.indexOf(input) + 1} already exists`)
    }
  }
}

function parseInput (inputFile) {
  return new Promise ((resolve, reject) => {
```

```
    const inputs = []
    csv()
      .fromFile(inputFile)
      .on('json', (obj) => { inputs.push(obj) })
      .on('done', (error) => {
        if (error) reject(error)
        else {
          console.log(`Finished reading input file with ${inputs.length} data
point(s).`)
          resolve(inputs)
        }
      })
  })
}

// checks if file already exists
function alreadyExists (lat, long) {
  if (fs.existsSync(`${config.outputFolder}/${lat}${long}.zip`))
    return true
  else
    return false
}
```