Gentrification Near Rail Transit Areas: A Micro-Data Analysis of Moves into Los Angeles Metro Rail Station Areas

December 2018

A Research Report from the National Center for Sustainable Transportation

Dr. Marlon G. Boarnet, University of Southern California

Dr. Raphael W. Bostic, Federal Reserve Bank of Atlanta

Evgeniy Burinskiy, University of Southern California

Dr. Seva Rodnyansky, University of California, Berkeley

Dr. Allen Prohofsky, California Franchise Tax Board

National Center for Sustainable Transportation

METRANS Transportation Center USC | CSULB

About the National Center for Sustainable Transportation

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

U.S. Department of Transportation (USDOT) Disclaimer

Federal Reserve System Disclaimer

California Franchise Tax Board Disclaimer

Acknowledgments

This study was funded by a grant from the National Center for Sustainable Transportation (NCST), supported by USDOT through the University Transportation Centers program. The authors would like to thank the NCST and USDOT for their support of university-based research in transportation, and especially for the funding provided in support of this project. We are grateful for advice and technical assistance from Chad Angaretis of the California Franchise Tax Board.

NCST

Gentrification Near Rail Transit Areas: A Micro-Data Analysis of Moves into Los Angeles Metro Rail Station Areas

A National Center for Sustainable Transportation Research Report

December 2018

Dr. Marlon G. Boarnet, Department of Urban Planning and Spatial Analysis, Sol Price School of Public Policy, University of Southern California

Dr. Raphael W. Bostic, President and CEO, Federal Reserve Bank of Atlanta

Evgeniy Burinskiy, Department of Urban Planning and Spatial Analysis, Sol Price School of Public Policy, University of Southern California

Dr. Seva Rodnyansky, Berkeley Institute for the Future of Young Americans (BIFYA), Goldman School of Public Policy, University of California-Berkeley

Dr. Allen Prohofsky, Chief Economist, California Franchise Tax Board

[page left intentionally blank]

TABLE OF CONTENTS

List of Tables

List of Figures

Gentrification Near Rail Transit Areas: A Micro-Data Analysis of Moves into Los Angeles Metro Rail Station Areas

EXECUTIVE SUMMARY

We use annual California Franchise Tax Board (FTB) micro-data on household income and location from 1993 to 2013 in conjunction with data on the Los Angeles County Metropolitan Transportation Authority (L.A. Metro) rail system to compute household mobility and income-distribution statistics within the 0.5 mile radius area of rail stations. Specifically, we longitudinally follow the residential location of tax filers who at any point during our time span reside within a 0.5-mile radius of a Metro rail station. We describe the manner in which the income distribution in station areas changed between 1994 and 2012. We also test whether the opening of LA Metro rail stations impacted the in, out, and net (in mobility rate minus out mobility rate) mobility rates of households in station areas. We focus on descriptive and regression analyses.

When reviewing descriptive statistics, six points stand out. First, we observe a large decline in the share of households in the Extremely Poor (0 to 30% of Area Median Income, or AMI) category living in rail station areas over the study period, though there was a slight rebound in the share in 2009 that persisted to 2012. Second, the share of households with higher incomes in rail station areas increased from 1994 to 2003, and then remained largely stable from that year forward. Third, the Extremely Poor generally had the highest residential l move-out <u>and</u> move-in rates in rail station areas over the study period. Fourth, move-in rates to rail station areas declined for all income categories from 1994 to 2012. Fifth, move out rates from rail station areas remained stable for all income categories for most of the study period. Sixth, the data on numbers of households shows that more movers, whether looking at moves in or moves out, were Extremely Poor or Poor (50 to 80% of AMI).

Among households for whom we can track year to year residential location in L.A. Metro rail station areas, mobility numbers suggest that more households move out than move in. This is true for all income groups. This finding is in line with American Community Survey (ACS) data on county-to-county mobility but does not capture the growth in the number of households. We suspect that because we require households to be in the FTB data for two consecutive years to establish their mobility, our in-mobility rate only tracks households who have been in the labor-force and filed California taxes the year before. This means we miss households who newly enter the labor force and households moving in from out of state.

Our regression analyses suggest that after a rail station opens the Extremely Poor (0-30% of AMI) households decrease the rate at which they move into rail stations by about 1.0 percentage point per year from an annual average of 9.4% or by about 11%. After station opening, the Very Poor (30-50% of AMI) and Poor (50-80% of AMI) households decrease the

rate at which they move out from annual averages of 9.0% and 9.5% by about 1.0 and 0.7 percentage points per year or by approximately 11% and 7%, respectively.

Overall, the findings from our analysis of mobility rates are indicative of complexity. A decreased in-mobility rate for one group can have the same effect on the group's presence as an increased out-mobility rate. For example, after rail stations open, the Extremely Poor move into rail station areas at a lower rate which, although not tested, may lead to a smaller presence in the neighborhood. This is not inconsistent with more canonical stories of neighborhood change whereby a group's out-mobility increases but something that might be unexpected if one focuses only on residential move-out rates. More broadly, there is variation in mobility patterns over time, across rail lines, and by income. Our empirical results suggest that low-income households face more than the singular decision of whether to move out of the rail transit area or not. It is also important to recognize that there is a baseline level of mobility in these neighborhoods, independent of whether a rail station opens or not. We find that, on average, every year approximately 1 in 10 households near rail transit stations move out and a roughly equal share move in. Finally, our study does not differentiate demographic change due to changes in income against demographic changes due to household residential mobility so conclusions attributing changes in the neighborhood income distribution to changes in mobility rates are likely premature.

Introduction: Measuring Moves Into and Out of Neighborhoods Near Rail Transit Stations

Rail transit and neighborhood compositional changes are becoming clearly linked in the public mind. Examples where rail transit has been associated, at least anecdotally, with neighborhood gentrification abound. In Washington, D.C., the Green and Yellow lines are associated with neighborhood transition north and east of downtown. In Los Angeles, the Gold, Expo, and Red/Purple lines have been associated with gentrification concerns (Zuk & Chapple, 2015a), and similar concerns have been raised regarding the soon-to-open Crenshaw Line. On balance, these same concerns are present in most large metropolitan areas that are building or expanding rail transit.

Gentrification is a process of neighborhood change characterized by increasing housing prices and changing demographic and socioeconomic composition of the neighborhood. These components of gentrification are often mutually reinforcing: changing composition can further increase housing prices and vice versa. Prior studies have raised the concern that rail transit expansion catalyzes or exacerbates gentrification (Zuk et al., 2017; Rayle, 2015).

This report seeks to shed light on this latter concern. It begins with a brief summary of the evidence from prior studies on both rail-related housing price increases and changing composition. It then introduces a newly available data source, which we use to examine the relationship between new rail transit station opening and neighborhood income composition. This report aims to determine whether a rail station opening in Los Angeles County is associated with the share and income composition of residents who move in and out of neighborhoods near that rail station. Specifically, we address the following questions regarding gentrification and its tie to rail transit stations:

- Who moves into rail-station neighborhoods and when?
- Are higher income households growing as a share of station area population relative to lower-income households?
- Do rail stations cause this phenomenon or is this happening regardless of the transit investment?

The Los Angeles metropolitan area presents an ideal study area for analyzing transit-oriented development (TOD) and potential displacement. Prior to 1990, Los Angeles had not had any intra-urban rail transit service for decades. Since then, 93 new rail-transit stations (see Figure 1 for map) were opened by the Los Angeles Metropolitan Transit Authority (L.A. Metro) and an additional 17 are currently under construction (Boarnet et al., 2015). This buildout amounts to about half of the U.S. spending on new rail transit (L.A. Metro, 2009). Within L.A. Metro, 21% of its budget from 2005-2040 will go toward rail transit capital and operations expenditures (L.A. Metro, 2009). Concurrently, regional and local plans envision that over half of new housing and employment to occur within a half-mile of a well-serviced transit corridor, including rail (L.A. Metro, 2009; SCAG, 2012).

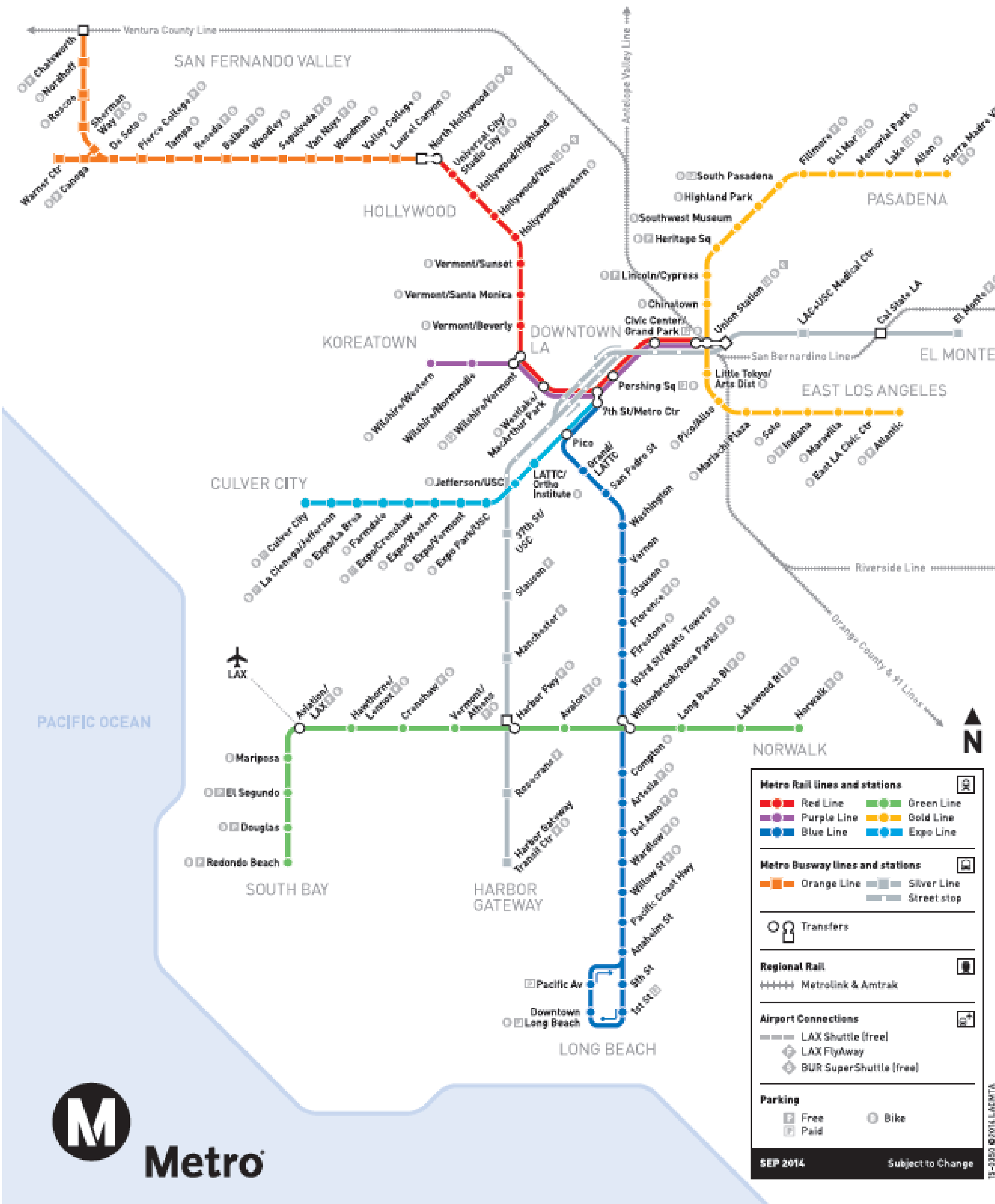

Figure 1. Map of Los Angeles Metro Rail Lines open in 2013
Source: Los Angeles County Metropolitan Transit Authority

Prior Findings on Rail Transit and Gentrification

Gentrification is a complex process of neighborhood change with a variety of definitions since the term was first coined over 50 years ago (Zuk et al., 2015). For this report, we abstract from this complexity and define gentrification as neighborhood-level changes in the income distribution of residents. Prior studies have defined gentrification more broadly, as a process of neighborhood change that can relate to changes in resident income, race, ethnicity, educational attainment, or a number of sociodemographic factors, in addition to changes in the composition of neighborhood businesses (Baker & Lee, 2017; Kahn, 2007; Grube-Cavers & Patterson, 2015; Heilmann, 2016; Chapple, 2009; Glaeser, Kahn, Rappaport 2008). Our focus, given our data, is more narrowly on income levels and household moves into and out of rail station neighborhoods. We summarize the findings of some key previous studies here.

One of the biggest concerns over expanding rail transit is increased housing prices. New rail stations can improve neighborhood accessibility for residents and businesses, represent an improvement in a neighborhood's amenities, or trigger improvements in a neighborhood's amenities through (re)investment, all of which can be capitalized into local area rents and home values (Smith, 1979; Knaap, Ding, & Hopkins, 2001). Multiple studies have confirmed that housing prices increase after rail stations open in a variety of U.S. metropolitan areas (Zuk et al., 2015): Phoenix (Atkinson-Palombo, 2010; Golub, Guhathakurta, & Sollapuram, 2012), Buffalo (Hess and Almeida, 2007), Atlanta (Immergluck, 2009), San Diego (Duncan, 2011), and in multi-city analyses (Higgins & Kanaroglou, 2016; Bartholomew & Ewing, 2011). However, results are not entirely conclusive. Results are often sensitive to station location (Dong, 2017; Lin, 2002), number of years after opening (Pilgram & West, 2017), decade of analysis (Lin, 2002), regression specification (Redfearn, 2009), and station type (park and ride versus walk and ride) (Kahn, 2007). Moreover, neighborhood house prices and average income can decrease if the primary users of rail transit are households whose incomes are relatively lower than those of the incumbent residents (Glaeser, Kahn, & Rappaport 2008).

Another concern is demographic change of the neighborhood induced or facilitated by new rail transit. New rail transit changes neighborhood composition if households who prefer transit service over other transportation options move in or if others who have not previously been as represented move in with greater numbers due to changes (real or perceived) that increase the neighborhood's appeal (Zuk et al., 2015; Pollack et al., 2010). Multiple studies have shown that in-movers to rail station neighborhoods generally alter the neighborhood composition along numerous dimensions (Zuk et al., 2017), including race or ethnicity, income, educational attainment, and age (Baker & Lee, 2017; Kahn, 2007; Grube-Cavers & Patterson, 2015; Heilmann, 2016; Chapple, 2009; Glaeser, Kahn, Rappaport 2008). When it comes to income composition, though, the literature is mixed because the direction of change depends on existing transportation options, location of jobs, present amenities, and many other factors. Income composition of a neighborhood around a station area can skew toward higher or lower incomes depending on the context.

Previous studies on rail transit and neighborhood composition have not been able to address the question of mobility and neighborhood income composition. The studies suffer from at least one of a few critiques. Some are cross-sectional and do not follow households longitudinally, some have long time gaps between observations (usually 10 years corresponding to the decennial census), some are not able to separate households by income group, or some do not look at a sufficiently small geography. Further, for all of these approaches, there are likely local changes at the neighborhood level not measured in the models, leading to omitted variable bias, which may overshadow rail station effects.

Understanding these limitations, this report focuses on better measurement of the effect of new rail station openings on changes in neighborhood composition. Our unique longitudinal micro-data on household-level mobility and income within a half-mile radius of rail stations enable us to overcome the limitations of prior studies.

This report focuses only on neighborhood composition and not on housing prices. While the California Franchise Tax Board (FTB) dataset described in the next section is ideal for studying neighborhood composition, it does not provide data to address housing prices. We leave the study of the relationship between housing price changes and rail transit for future work. Nevertheless, we believe that the application of this new dataset and method to studying neighborhood composition presents an advance in the understanding of the interconnection between rail transit and gentrification.

Data

To answer our three primary questions of interest, this report leverages annual data on household locations and incomes from 1993 to 2013. This dataset is constructed using data from income tax filings obtained from the California Franchise Tax Board (FTB). The data universe contains information on all households who filed taxes in Los Angeles County in any year between 1993 and 2013, even if they lived outside the County or California for some of the years during this period. The dataset includes information available on the standard California tax return such as household income, state taxes paid, location, and other household-specific characteristics. The data also contain the 9 digit zip code of the address at which a household filed taxes. Depending on the density of the area, a 9-digit zip code can be as small as a building or as big as a street block. All the data were "de-personalized" such that personally-identifiable information was redacted prior to our having access to the data.

Using annual household-level data on income and location, we calculate the number of households that reside, move into, and move out of areas within 0.5 miles of LA Metro rail stations between 1994 and 2012. In conjunction with data on the L.A. Metro system, we use our mobility statistics to test whether the opening of a rail station has had an impact on the rate at which households at various income levels move into and out of station areas.

Geocoding

In order to track whether people move and, if they do, where they move to, we need a data identifier of location. For our analysis, we use zip codes to document a household's location. The dataset includes either the 5- or 9-digit zip code associated with the location each household identified in its income tax filing. We choose to focus only on those households who have 9-digit zip code identifiers, as this allows us to track movements of these households with a good degree of geographic specificity. Within the hierarchy of the U.S. Postal Service delivery routes, 9-digit zip codes represent one block, one block-face, or large buildings, and hence our strategy allows us to detect household moves that span relatively small distances, even several blocks.

To detect a move, we geocode the filing location of households who have a 9-digit zip code using the geographic coordinates of their 9-digit zip code centroid using data from Geolytics, Inc., for the years 2000, 2002, 2004, 2007, 2009, 2011, 2012, 2013, and 2016. We match each 9-digit zip code to its geographic coordinates in the nearest year. If coordinates are not available for the current year, but available for a later year, the later year is used. We use this strategy because 9-digit zip code locations do not change very much from year to year. Fewer than 1% of all California 9-digit zip codes moved more than 100 meters between 2000 and 2012 (Rodnyansky et al., 2018).

To preserve a household's confidentiality, the 9-digit zip code is included by the FTB in the data only if that zip code has a population exceeding 10 households. If a household lives in a zip code that has 10 households or fewer, its location is reported as the 5-digit zip code. These Because 5 digit zip codes are too large for our analysis, we drop households geocoded to 5 digit zip codes, or 51 percent of the observations in an average year from the analysis. The ability to measure the distance between the centroid of a 1-block-long 9-digit zip code to an L.A. Metro station is of prime importance to this report and represents a refinement over prior work. Adding 5-digit zip codes to this analysis would dilute the geographic specificity necessary for this particular analysis.

Because higher residential densities are needed to report the 9-digit zip code for a tax filer, it is likely that filing households with 9-digit zip code over-represent more dense regions of Los Angeles County and under-represent mountainous, forested, and desert regions of the County. Our choice of 9 digit zip codes to represent households near L.A. Metro rail stations is appropriate because stations tend to be located in the County's most populated and dense areas.

We measure the proximity of a filer to an L.A. Metro rail station by calculating the distance between the filer's associated 9-digit zip code coordinates and the nearest station. We compute the distances for each year from 1993-2013, the time span for which we have available data. We define a rail station area neighborhood as a circle with a 0.5-mile radius around a station. This represents an approximate 10-15 minute walking distance from a station to the furthest extent of the neighborhood and is a common neighborhood distance used in the planning of land uses surrounding rail stations and transit oriented development. Depending on the street network, however, actual walking times from a station to the edge of the rail-station area may be longer than 15 minutes. The process of associating 9-digit zip codes and station area neighborhoods is detailed in the Appendix: Calculating Residential Mobility Rates by Station.

Appendix Table 15 shows the total number of 9-digit zip codes within each station neighborhood by year. There is an average of 113 (and a median of 102) 9-digit zip codes (i.e., blocks or block-faces) within each 0.5 mile catchment area. A few stations have very few 9-digit zip codes, while some have as many as 400. From 1993 to 2013, the average and the median number of 9-digit zip codes have both more than doubled in number, likely reflecting a growth in population and population density in rail-proximate neighborhoods.

A review of the data revealed several other issues that required resolution. The Robustness and Sensitivity Appendix identifies these issues and how we resolved them to create the final dataset used for the analysis presented below. To provide a sense of this process, we offer the following example. Several 9-digit zip codes had a very high number of files. Figures 4 and 5 (Robustness and Sensitivity Appendix) map 9-digit zip codes within station areas whose number of tax filers was in the .9999th percentile of number of filers in a particular year. As we observe in Figure 4, these zip codes map to large commercial areas such as malls, museums, and stadiums. We suspect that households who filed at these 9-digit zip code locations do not reside in these locations, but rather filed their taxes under their employment address. Since these zip codes are likely not representative of residential patterns but rather of employment, we exclude these zip codes from our sample. This category includes 3 zip codes total, accounting for 1718 filers across all years or an average of 192 households per year and station with a maximum of 317 and minimum of 131 filers. To ensure these outlier 9-digit zip codes do not affect our analyses, we run a robustness check re-incorporating these outlier zip codes to determine whether results are sensitive to the inclusion or exclusion of these high-county 9 digit zip codes.

Sample Construction

We construct a sample of tax filers between 1994 and 2012 as follows. First, as noted above, we find all households whose tax filing has a 9-digit zip code attributed to them. Next, we limit the sample to those filers for whom we have a 9-digit zip code in consecutive years. This is required in order to determine whether a filer has moved from one year to the next. Third, we use our geocoding to establish which households reside within 0.5 miles of an L.A. Metro rail station in a given year. This last step permits one to identify if a household's zip code changed from one year to the next; such a change is defined as a move. It also facilitates determining whether a move was into or out of a metro rail station area.

Table 1 shows statistics describing how the sample construction process affected the size of the final sample upon which the analysis was run. On average, 49% of the filer population in a given year has a 9-digit zip code for their location (column a). Over the period, an average of 74 percent of these filers were present in the subsequent year (column b). Finally, we were generally able to geocode 91 percent of household locations with 9 digit zip codes (column c). Our final sample of geocodable consecutive filers is on average 34% of the total filer population or about 100,000 filers per year (Table 1, columns D and E).[1]

Table 1 also shows that the proportion of filing population possessing the desired properties in the sample improved over a number of dimensions, which contributes to the bulk of the increase in our sample over the study period. This includes increases in the proportion of households with 9-digit zip codes in the data, with consecutive 9-digit zip codes, and with matchable geographic coordinates. As a consequence, the share of filing households we can use in our study increased from 22% in 1994 to a high of 36-40% in the 2000s. The number of in-sample filers (Table 1, column E) grew from about 50,000 to over 120,000, an increase of 137%.

[1] We discuss the representativeness of our sample later in this section in the External Data Validation sub-section.

Table 1. Sample Construction Details

Column name:	(a)*	(b)*	(c)*	(d)	(e)
Data Source:	*FTB*	*FTB*	*Geolytics*	*Author calculation*	
Year t	**% filers with 9 digit zip codes in year t**	**% of filers with 9 digit zip codes in years t and t+1**	**% of 9-digit zip codes with latitude / longitude**	**% in Sample (a) x (b) x (c)**	**All tax filers in study area x (e)**
1994	39%	63%	88%	22%	51,402
1995	44%	69%	88%	27%	60,895
1996	47%	71%	89%	30%	67,310
1997	48%	72%	89%	31%	74,083
1998	49%	75%	90%	33%	81,052
1999	50%	76%	91%	35%	87,958
2000	51%	77%	93%	36%	93,474
2001	51%	78%	91%	36%	98,045
2002	51%	77%	91%	36%	95,445
2003	51%	77%	92%	36%	97,799
2004	51%	77%	92%	36%	105,125
2005	51%	78%	92%	37%	111,691
2006	52%	79%	93%	38%	119,912
2007	52%	79%	97%	40%	124,665
2008	52%	79%	95%	39%	123,044
2009	52%	79%	96%	40%	124,822
2010	53%	79%	91%	38%	126,250
2011	54%	78%	88%	37%	123,892
2012	54%	78%	88%	37%	121,947
Average	*49%*	*74%*	*91%*	*34%*	*99,411*

*Columns (a), (b), (c) represent successive cuts to our data due to data limitations.

Income Categorization

We categorize filers into income groups so we can test whether rail station opening year impacted mobility rates differently across income brackets. We derive 7 categories using the U.S. Department of Housing and Urban Development's (HUD) Area Median Income (AMI) for Los Angeles County for that year. AMI is a flexible and relative measure of income for a metropolitan area. Three category cutoffs correspond to HUD's designation of Extremely Poor (0-30% AMI), Very Poor (30-50% AMI), and Poor (50-80% AMI) households. Filers who report incomes between -100% of AMI and 0, who we refer to as Negative Earners, had more expenses than income in a particular year; such households are most likely self-employed. The remaining three cutoffs 80-100% AMI (Middle Income), 100-200% AMI (Upper Middle Income), and 200-300% AMI (Upper Income) represent non-poor households.[2] Table 2 shows the income cutoffs for the 7 categories in each sample year.

Note that AMI is a relative measure of income that changes as the median income of the metropolitan area changes. A particular household's income may increase or decrease over time without changing their relative status within the income distribution. Conversely, if a household's income does not change over time then they may step up or down along the income distribution over time, if households moving into the County have a different income distribution or if there is a general change in economic conditions countywide that causes the AMI to shift. Nevertheless, because we are interested in the change in the income distribution of a rail station, we are not too concerned about these shifts.

Table 2. Area Median Income (AMI) by Year and Income Bracket Cutoffs

Year	Negative Earners (-100% AMI)*	Extremely Poor (30% AMI)	Very Poor (50% AMI)	Poor (80% AMI)	Area Median Income (100% AMI)	Upper Middle Income (200% AMI)	Upper Income (300% AMI)
1993	$(42,300)	$12,690	$21,150	$33,840	$42,300	$84,600	$126,900
1994	$(45,200)	$13,560	$22,600	$36,160	$45,200	$90,400	$135,600
1995	$(45,200)	$13,560	$22,600	$36,160	$45,200	$90,400	$135,600
1996	$(46,900)	$14,070	$23,450	$37,520	$46,900	$93,800	$140,700
1997	$(47,800)	$14,340	$23,900	$38,240	$47,800	$95,600	$143,400
1998	$(49,800)	$14,940	$24,900	$39,840	$49,800	$99,600	$149,400
1999	$(51,300)	$15,390	$25,650	$41,040	$51,300	$102,600	$153,900
2000	$(52,100)	$15,630	$26,050	$41,680	$52,100	$104,200	$156,300
2001	$(54,500)	$16,350	$27,250	$43,600	$54,500	$109,000	$163,500
2002	$(55,100)	$16,530	$27,550	$44,080	$55,100	$110,200	$165,300
2003	$(50,300)	$15,090	$25,150	$40,240	$50,300	$100,600	$150,900
2004	$(54,200)	$16,260	$27,100	$43,360	$54,200	$108,400	$162,600
2005	$(54,450)	$16,335	$27,225	$43,560	$54,450	$108,900	$163,350
2006	$(56,200)	$16,860	$28,100	$44,960	$56,200	$112,400	$168,600
2007	$(56,500)	$16,950	$28,250	$45,200	$56,500	$113,000	$169,500
2008	$(59,800)	$17,940	$29,900	$47,840	$59,800	$119,600	$179,400
2009	$(62,100)	$18,630	$31,050	$49,680	$62,100	$124,200	$186,300
2010	$(63,000)	$18,900	$31,500	$50,400	$63,000	$126,000	$189,000
2011	$(64,000)	$19,200	$32,000	$51,200	$64,000	$128,000	$192,000
2012	$(64,800)	$19,440	$32,400	$51,840	$64,800	$129,600	$194,400
2013	$(61,900)	$18,570	$30,950	$49,520	$61,900	$123,800	$185,700

*cut-off for -100% to 0% of AMI earners

[2] HUD adjusts incomes depending on family size while we do not, so our income bins likely underestimate measures of poverty. However, since we are primarily interested in neighborhood compositional effects we do not see this as a major issue.

In Table 3, we present the income distribution for each year in the sample. The largest number of households within 0.5 miles of Los Angeles rail stations are categorized as Extremely Poor, Very Poor, Poor, and Middle Income. The number of households in each category grew sizably from the beginning to the end of the study period. Increases were largest for the Upper Income and Upper Middle Income groups. Table 3 also reports households with incomes outside of the range covered by our seven categories. Households with incomes below -100% of AMI likely had very extreme shocks to income that would necessitate them making location decisions not related to the presence or absence of a rail station. Similarly, households with incomes above 300% of AMI might be more likely to make location decisions independent of infrastructure improvements such as rail stations. Because of these possibilities, we do not include these households, which represent less than 2% of the total number of households in any given year, in the empirical analysis.

Table 4 shows the proportion of in-sample filers by income category. The plurality of the sample, or about one third, falls into the Extremely Poor category, though this average masks several trends over the 20 years. We observe a large decline in the share of Extremely Poor households from 1994 to 2003. This share fluctuates in the same general range for the next 5 years, and then jumps to 33 percent in 2009 and remains there through 2012. The key takeaway here is that the pattern for Extremely Poor households has not been monotonic.

The next-largest income categories are Very Poor and Poor, which average 23% and 19% of sample filers respectively. In contrast to the Extremely Poor category, the proportion of Poor and Very Poor households has remained relatively stable through the study period. Like the Poor and Very Poor shares, the share of households in the Middle Income category remained stable over the period. Finally, the proportion of Upper Middle and Upper Income households is smaller than the lower-income groups, on average 12% and 2% respectively. The proportion of filers in these categories increased between 1994 and 2003, with the most notable increase in the Upper Middle Income category, but has remained largely stable since 2003.

On balance, Table 4 shows there has been a slight shift toward a higher income tax filing population near L.A. Metro stations and away from the lowest-income category. This pattern appears more striking when one looks at numbers of households (Table 3), as the growth in the number of higher income households is easy to detect. However, it is important to recognize that the number of such households was quite small in the early years of the sample. This highlights the importance of also noting population shares.

Table 3. Number of In-Sample Filers within 0.5 miles of all L.A. Metro Stations by Income Category and Year

Year	Negative Earners (<0% AMI)	Extremely Poor (0-30% AMI)	Very Poor (30-50% AMI)	Poor (50-80% AMI)	Middle Income (80-100% AMI)	Upper Middle Income (100-200% AMI)	Upper Income (200-300% AMI)	Incomes outside range	All Incomes
1994	628	21,819	12,122	8,191	2,921	4,359	713	649	51,402
1995	681	23,689	14,565	10,529	3,664	5,993	927	847	60,895
1996	849	26,290	15,842	11,582	4,043	6,629	1,114	961	67,310
1997	896	27,955	17,562	13,165	4,464	7,561	1,303	1,177	74,083
1998	768	29,814	19,749	14,567	5,041	8,333	1,471	1,309	81,052
1999	741	31,889	21,291	16,191	5,557	9,170	1,687	1,432	87,958
2000	744	32,621	22,429	17,650	6,188	10,247	1,916	1,679	93,474
2001	805	34,843	23,892	18,371	6,335	10,375	1,880	1,544	98,045
2002	905	32,955	23,104	18,154	6,391	10,532	1,878	1,526	95,445
2003	937	28,765	23,259	19,991	7,349	12,895	2,488	2,115	97,799
2004	1,029	33,330	25,101	20,705	7,546	12,788	2,416	2,210	105,125
2005	1,114	34,194	25,836	22,253	8,227	14,475	2,946	2,646	111,691
2006	1,349	35,937	27,329	23,941	9,041	15,987	3,322	3,006	119,912
2007	1,454	35,205	28,195	25,093	9,484	17,846	3,858	3,530	124,665
2008	1,507	35,651	28,394	24,277	9,251	17,182	3,686	3,096	123,044
2009	1,708	40,757	28,455	22,837	8,648	16,083	3,526	2,808	124,822
2010	1,651	41,571	28,824	22,911	8,689	16,169	3,476	2,959	126,250
2011	1,605	40,781	28,078	22,339	8,308	16,280	3,533	2,968	123,892
2012	1,414	39,744	27,340	22,077	8,294	16,261	3,667	3,150	121,947
All Years	20,785	627,810	441,367	354,824	129,441	229,165	45,807	39,612	1,888,811
Standard Deviation of Annual Population Change	108	2,214	1,004	1,065	429	921	233	259	3,946
18-year change in filing households*	125%	82%	126%	170%	184%	273%	414%	385%	137%

Source: FTB

Table 4. Income Distribution in Neighborhoods near L.A. Metro Stations (Households in 9-digit zip codes within 0.5 miles of L.A. Metro Stations)

Year	Negative Earners (<0% AMI)	Extremely Poor (0-30% AMI)	Very Poor (30-50% AMI)	Poor (50-80% AMI)	Middle Income (80-100% AMI)	Upper Middle Income (100-200% AMI)	Upper Income (200-300% AMI)	Incomes Outside Range
1994	1%	42%	24%	16%	6%	8%	1%	1%
1995	1%	39%	24%	17%	6%	10%	2%	1%
1996	1%	39%	24%	17%	6%	10%	2%	1%
1997	1%	38%	24%	18%	6%	10%	2%	2%
1998	1%	37%	24%	18%	6%	10%	2%	2%
1999	1%	36%	24%	18%	6%	10%	2%	2%
2000	1%	35%	24%	19%	7%	11%	2%	2%
2001	1%	36%	24%	19%	6%	11%	2%	2%
2002	1%	35%	24%	19%	7%	11%	2%	2%
2003	1%	29%	24%	20%	8%	13%	3%	2%
2004	1%	32%	24%	20%	7%	12%	2%	2%
2005	1%	31%	23%	20%	7%	13%	3%	2%
2006	1%	30%	23%	20%	8%	13%	3%	3%
2007	1%	28%	23%	20%	8%	14%	3%	3%
2008	1%	29%	23%	20%	8%	14%	3%	3%
2009	1%	33%	23%	18%	7%	13%	3%	2%
2010	1%	33%	23%	18%	7%	13%	3%	2%
2011	1%	33%	23%	18%	7%	13%	3%	2%
2012	1%	33%	22%	18%	7%	13%	3%	3%
All Years	1%	33%	23%	19%	7%	12%	2%	2%

Source: FTB

External Data Validation

To ensure the external validity of the FTB filer and income proportion data, we compare our sample sizes and proportions to U.S. Census data and previous studies using FTB data. While the sample restriction and geocoding approaches are unique to this particular study, we believe that the comparisons are relevant and useful even if U.S. Census data is not available in every year and if different levels of spatial aggregation are used.

Using Census data on the number of households, we compute the number of households we should expect to file an income tax return given limitations inherent in the FTB data and compare this to the number of tax filers in our sample. To do so, we spatially aggregate the number of households in each U.S. Census block and block group in Los Angeles County to half-mile buffers around the L.A. Metro station areas using ArcGIS. We use data from National Historical Geographical Information System (NHGIS) from the available years, which include U.S. Decennial Census blocks in 2010 and block groups in 1990, 2000, and 2010, and the American Community Survey 5-year sample block groups from 2005-2009, 2006-2010, 2007-2011, 2008-2012, 2009-2013, 2010-2014, 2011-2015, and 2012-2016 (Manson et al., 2018). We match these as closely as possible to each study year (see Table 5, columns A and B).[3] We next apply the proportion of tax filers reporting consecutive 9-digit zip codes that are geocodable (Table 5, columns C, D, and E) and the proportion of households estimated to file taxes in California in every year. The estimated tax filing proportion in California is estimated to be 80-90% according to the FTB reports of 2006 and 2017 (FTB, 2006; FTB, 2017). We use 80% as a conservative lower bound (Table 5, column G). We then subtract the census-derived expected number of filers from the actual FTB sample number in each year.

The expected Census block group estimate (Table 5, column H) overestimates the number of tax filers compared to the FTB in all but one year by an average 28,000 filers or about 20% of filers. The expected Census block estimate (Table 5, column I) underestimates the number of households relative to the FTB in years prior to 2005 and overestimates them for 2005 and subsequent years. The Census block pattern is monotonically increasing from 1996 to 2011. One reason for this is that our Census block estimate of number of households does not change between 1994 and 2012 whereas the true underlying number of households was likely steadily increasing from early 1990s to the 2010s.

We believe the block and block group estimates provide a lower and upper bound, respectively, for the expected number of consecutive filers geocodable to the 9-digit zip code. Since Census block groups tend to be too large to be wholly encapsulated within the half-mile radius around a Metro station area, the Census Block Group estimate likely includes households who reside beyond the half-mile area, thereby over-counting the number of households near station areas. The Census block is geographically smaller than the block group so if it is near a station then it is more likely to be wholly contained by the half-mile radius. Census blocks likely provide a more accurate count for the number of households near rail stations but because they are

[3] Appendix Table 16 reports the number of households by Los Angeles Metro Line by year.

rectangular and are fitted inside a circle, it is likely blocks do not cover the entire area encapsulated by a half-mile radius. For these reasons, we believe block groups are an upper estimate and a block is lower estimate for the number of households in station areas. The relatively similar number of filers using Census and FTB data provides some assurance of the external validity of our sample, giving reassurance that our sample is representative of the households residing in L.A. Metro station areas.

As an additional validation, we compare the income distribution in the FTB data against the Census block group data. Using income categories reported in the Census, Table 6 shows the percentage point difference between the FTB tax filing households and Census households for the L.A. Metro station study area (Census income proportion minus FTB). We see that on average income proportions are reasonably similar between the two data sources. For households with annual incomes between $30,000 and $50,000, or near the U.S. median, there is near complete similarity in the income distribution. For incomes below $30,000, the FTB data appears to overestimate by 2-5 percentage points on average off an average annual base of 34 percent in the FTB data, and by as much as 7 or 8 percentage points in the mid-1990s off a base of 44 percent. In contrast, for households with incomes between $50,000 and $100,000, the FTB data underestimates the census by 1-4 percentage points. The FTB overestimate is of lower magnitude (1-2 percentage points) for incomes above $100,000. The overall similarity of the income distribution between the data sets provides more assurance that the FTB data are representative of households residing in L.A. Metro station areas. We acknowledge the divergences in the lower and upper ends of the distribution. Perhaps these arise from geocoding issues or non-random selection of which households have 9-digit versus 5-digit zip codes. We test this next.

To test the effect of the 9-digit restriction on the income distribution of filers, we compare the current sample to two other studies that used the FTB data for this study area, but included filers with 5-digit zip codes, in addition to those with 9-digit zip codes (Rodnyansky et al., 2018; Boarnet et al., 2018). Both of those studies utilized the same study area (within 0.5 miles of L.A. Metro rail stations) and required that households be in the data consecutively and that the zip codes had geographic coordinates for both years. Table 7 compares the income distribution of the current FTB sample to that of the prior studies using the income categories from the prior studies. Table 7 shows very few differences in income distribution between this study using the 9-digit zip codes and the previous studies using 9-digit and 5-digit zip codes. When there are differences, magnitudes are very low with mostly 1 percentage-point differences and the pattern of differences appears random. The consistency of income distributions across our sample and that of the previous studies provides additional assurance of our sample restriction strategy.

We observe a growth in the number of filing households near LA Metro rail stations. The number of filing households has grown in our data from 51,402 to 121,947 between 1994 and 2012 or by almost 140%. On the other hand, the growth according to the Census Block Groups is about 13% between 1990 and 2012. Given that the Census Block Groups likely represent a larger area and do represent a longer time span, we suspect the 13% growth rate is an upper bound for the growth rate in the number of households near LA Metro Rail stations. The large difference in growth rates between the Census and FTB data imply that the majority of growth in number of filing households is most likely a result a higher share of filing households meeting the criteria required to be in our data sample.

Table 5. Census Comparison Table

Column name:	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	(l)
Data Source:	*Census Bureau*		*FTB*	*FTB*	*Geolytics*	*Author calc.*	*FTB Reports*	*Author calculation*				
Year t	**Households (Block Group)**	**Households (Block) in 2010**	**% filers with 9 digit zip codes in year t**	**% of filers with 9 digit zip codes in years t and t+1**	**% of 9-digit zip codes with latitude / longitude**	**% in Sample (c) x (d) x (e)**	**Estimated % of CA households filing taxes**	**Expected Filers via Census Block Group (a) x (f) x (g)**	**Expected Filers via Census Block (b) x (f) x (g)**	**# of Filers in FTB sample**	**FTB minus Block Group (j) – (h)**	**FTB minus Block (j) – (i)**
1994	446,277*	364,521	39%	63%	88%	22%	80%	77,266	63,112	51,402	-25,864	-11,710
1995	446,277*	364,521	44%	69%	88%	27%	80%	95,811	78,259	60,895	-34,916	-17,364
1996	458,198*	364,521	47%	71%	89%	30%	80%	108,363	86,209	67,310	-41,053	-18,899
1997	458,198*	364,521	48%	72%	89%	31%	80%	113,152	90,019	74,083	-39,069	-15,936
1998	458,198*	364,521	49%	75%	90%	33%	80%	121,191	96,414	81,052	-40,139	-15,362
1999	458,198*	364,521	50%	76%	91%	35%	80%	126,820	100,892	87,958	-38,862	-12,934
2000	458,198	364,521	51%	77%	93%	36%	80%	133,568	106,261	93,474	-40,094	-12,787
2001	458,198*	364,521	51%	78%	91%	36%	80%	133,384	106,114	98,045	-35,339	-8,069
2002	458,198*	364,521	51%	77%	91%	36%	80%	131,313	104,466	95,445	-35,868	-9,021
2003	458,198*	364,521	51%	77%	92%	36%	80%	131,737	104,804	97,799	-33,938	-7,005
2004	458,198*	364,521	51%	77%	92%	36%	80%	132,903	105,732	105,125	-27,778	-607
2005	458,198*	364,521	51%	78%	92%	37%	80%	134,136	106,712	111,691	-22,445	4,979
2006	458,198*	364,521	52%	79%	93%	38%	80%	140,398	111,694	119,912	-20,486	8,218
2007	321,844	364,521	52%	79%	97%	40%	80%	102,706	116,325	124,665	21,959	8,340
2008	476,231	364,521	52%	79%	95%	39%	80%	148,871	113,950	123,044	-25,827	9,094
2009	481,520	364,521	52%	79%	96%	40%	80%	152,614	115,532	124,822	-27,792	9,290
2010	486,508	364,521	53%	79%	91%	38%	80%	148,408	111,196	126,250	-22,158	15,054
2011	498,287	364,521	54%	78%	88%	37%	80%	147,534	107,928	123,892	-23,642	15,964
2012	504,123	364,521	54%	78%	88%	37%	80%	148,747	107,556	121,947	-26,800	14,391
Average	*457,376*	*364,521*	*49%*	*74%*	*91%*	*34%*	*80%*	*123,306*	*98,178*	*99,411*	*-23,895*	*1,233*

Source: FTB, Census, ACS; *The 1994-1995 Block Group data are from the 1990 Census; the 1996-2006 Block Group data are from the 2000 Census.

Table 6. Percentage Point Difference Census versus FTB Income Distribution (Census minus FTB) for L.A. Metro station areas BG = Block Group

FTB Year	Census Comparison Year	<$10K	$10-15K	$15-20K	$20-25K	$25-30K	$30-35K	$35-40K	$40-45K	$45-50K	$50-60K	$60-75K	$75-100K	$100-125K	$125-150K	$150-200K	>$200K
1993	1990 BG	-12%	-8%	-2%	1%	2%	3%	3%	3%	2%	4%	3%	2%	1%	0%	1%	-1%
1994	1990 BG	-7%	-8%	-3%	0%	2%	3%	2%	2%	2%	3%	3%	2%	1%	0%	1%	-1%
1995	1990 BG	-5%	-7%	-3%	0%	1%	2%	2%	2%	1%	3%	2%	2%	1%	0%	1%	-1%
1996	2000 BG	-7%	-8%	-5%	-1%	0%	1%	2%	2%	2%	3%	4%	4%	2%	1%	1%	1%
1997	2000 BG	-5%	-7%	-5%	-2%	0%	1%	1%	2%	1%	3%	4%	4%	2%	1%	1%	1%
1998	2000 BG	-3%	-7%	-5%	-2%	0%	1%	1%	2%	1%	3%	4%	3%	2%	1%	1%	1%
1999	2000 BG	-2%	-6%	-5%	-2%	0%	1%	1%	1%	1%	2%	3%	3%	2%	1%	1%	1%
2000	2000 BG	-1%	-5%	-5%	-2%	-1%	1%	1%	1%	1%	2%	3%	3%	2%	1%	1%	1%
2001	2000 BG	0%	-5%	-5%	-2%	-1%	0%	1%	1%	1%	2%	3%	3%	1%	1%	1%	1%
2002	2000 BG	1%	-4%	-5%	-2%	-1%	0%	0%	1%	1%	2%	3%	3%	1%	1%	1%	1%
2003	2000 BG	1%	-4%	-5%	-2%	-1%	0%	0%	1%	1%	2%	3%	3%	1%	1%	1%	1%
2004	2000 BG	2%	-3%	-4%	-2%	-1%	0%	0%	1%	1%	1%	2%	2%	1%	0%	0%	0%
2005	2005-2009 BG	-6%	-3%	-5%	-4%	-2%	-1%	0%	0%	1%	2%	4%	5%	3%	2%	2%	1%
2006	2006-2010 BG	-5%	-2%	-4%	-4%	-2%	-1%	0%	0%	0%	2%	4%	4%	3%	2%	2%	1%
2007	2007-2011 BG	-4%	-1%	-4%	-4%	-2%	-1%	-1%	0%	0%	2%	3%	4%	3%	2%	2%	1%
2008	2008-2012 BG	-3%	-1%	-4%	-4%	-2%	-2%	-1%	0%	0%	2%	3%	4%	3%	2%	2%	2%
2009	2009-2013 BG	-4%	-2%	-5%	-3%	-2%	-1%	0%	0%	0%	2%	3%	4%	3%	2%	2%	2%
2010	2009-2013 BG	-3%	-2%	-4%	-4%	-2%	-1%	0%	0%	0%	2%	3%	4%	3%	2%	2%	2%
2011	2011-2015 BG	-3%	-2%	-4%	-3%	-3%	-1%	0%	0%	0%	2%	3%	4%	3%	2%	2%	2%
2012	2012-2016 BG	-3%	-2%	-4%	-3%	-3%	-1%	0%	0%	0%	1%	3%	4%	3%	2%	2%	2%

Source: FTB and U.S. Census and ACS

Table 7. Income Distribution Comparison between Current Report vs. CCF / Rodnyansky et al. 2018

	Current Report: Zip 9 only				Previous Reports: Zip 9 + Zip 5 hybrid geocode				Current minus Previous			
Year	<30% AMI	30-50% AMI	50-80% AMI	>80% AMI	<30% AMI	30-50% AMI	50-80% AMI	>80% AMI	<30% AMI	30-50% AMI	50-80% AMI	>80% AMI
1994	44%	24%	16%	17%	41%	22%	17%	20%	3%	2%	-1%	-3%
1995	40%	24%	17%	19%	39%	22%	18%	21%	1%	2%	-1%	-2%
1996	40%	24%	17%	19%	41%	22%	17%	20%	-1%	2%	0%	-1%
1997	39%	24%	18%	20%	41%	23%	18%	20%	-2%	1%	0%	0%
1998	38%	24%	18%	20%	38%	23%	18%	20%	0%	1%	0%	0%
1999	37%	24%	18%	20%	37%	24%	18%	20%	0%	0%	0%	0%
2000	36%	24%	19%	21%	37%	24%	19%	21%	-1%	0%	0%	0%
2001	36%	24%	19%	21%	37%	24%	19%	20%	-1%	0%	0%	1%
2002	35%	24%	19%	21%	37%	24%	19%	21%	-2%	0%	0%	0%
2003	30%	24%	20%	25%	32%	23%	20%	25%	-2%	1%	0%	0%
2004	33%	24%	20%	24%	34%	23%	20%	23%	-1%	1%	0%	1%
2005	32%	23%	20%	25%	32%	23%	20%	25%	0%	0%	0%	0%
2006	31%	23%	20%	26%	32%	22%	20%	26%	-1%	1%	0%	0%
2007	29%	23%	20%	28%	30%	22%	20%	28%	-1%	1%	0%	0%
2008	30%	23%	20%	27%	31%	22%	20%	27%	-1%	1%	0%	0%
2009	34%	23%	18%	25%	35%	22%	19%	24%	-1%	1%	-1%	1%
2010	34%	23%	18%	25%	35%	22%	18%	24%	-1%	1%	0%	1%
2011	34%	23%	18%	25%	35%	22%	18%	25%	-1%	1%	0%	0%
2012	34%	22%	18%	26%	35%	22%	18%	25%	-1%	0%	0%	1%

Descriptive Statistics

We use the FTB dataset to describe our population sample of filers living near rail stations by year, by income category, and by rail line. We also show the number of households moving IN to and OUT of neighborhoods near L.A. Metro rail stations and compute the NET change in number of households near the stations by year, income, and rail line. The descriptive statistics, accompanying tables and charts build on the sample description above and provide context for understanding the baseline for how many households move into and out of these neighborhoods annually. We discuss the descriptive data in two steps. First, we focus on the income distribution of households near stations by line and then we describe the pattern of household moves into, out of, and net (inflow minus outflow) within the ½ mile station areas. A longitudinal comparison of income distributions informs us on whether the income distribution near metro areas have indeed shifted in a discernable manner while analysis of move-in and move-out rates informs whether changes in the income distribution were due to changed mobility patterns between income groups.

Income Distribution by Rail Line

Table 7 above provides a comprehensive view of the year-to-year changes in income proportions for in-sample filers systemwide. Figure 2 uses the same data to provide a snapshot of these income distributions in 1994, the first year of observation, and triennial snapshots from 2000-2012.

Figures 2.A through 2.E shows the income distribution for each rail line. Because of the large overlap between the Red and Purple lines, we combine the stations along both lines and title the result as the Red line. This reduces the L.A. Metro rail system down to 5 train lines: Blue, Expo, Gold, Green, and Red. For reference, we show the same graphs for the aggregate income distribution across all lines (Figure 2.F).

We first note that the Extremely Poor income category is the largest income category for each rail line in all years. The share of households in the Extremely Poor category is the highest for the station areas along the Blue Line, followed by the station areas along the Red line. Station areas on the Gold and Green lines have similar shares of Extremely Poor households, while the Expo line has the smallest proportion of Extremely Poor households. When looking at how this share evolves over the sample time period, we see that the pattern mirrors the aggregate pattern – decline in share from 1994 to 2003, followed by an increase in share through 2012 – for all lines except the Red/Purple line. For Red, the share does not increase post-2003 but rather remains steady.

Turning to the other income categories, as is the case in the aggregate, the shares of Very Poor, Poor, Middle Income, and Upper Income households for the station areas of all 5 rail lines remain stable for the entire study period. Patterns for the Upper Middle income category differed most significantly across the five lines, particularly in the years after 2000. Before 2000, the share of households in the Middle Income category increased across the board. However, we note again that there were small numbers of such households in many station areas during this time. Since 2000, the Upper Middle income share for was stable for the Expo line and generally stable, with some fluctuation, on the Green line. The Upper Middle income share continued to grow from 2000 to 2003 on the Blue and Gold lines, and then stabilized from 2003 to 2012. Unlike the case for the other lines, the Upper Middle income share in stations areas continued to grow on the Red/Purple line throughout the period. It increased by 30% between 2003 and 2012.

An interesting question is how the station area income distribution patterns compare with income distribution patterns citywide. Figure 3 reports the income distribution patterns for all filers in L.A. County for whom we have 9 digit zip code data. Citywide, we observe a pattern similar in some respects to what occurred in rail station areas. The share of households in L.A. County with lower incomes declined continuously over the period. This dynamic was pronounced at the lowest end of the income distribution; the share of households who earned less than $15,000 fell from 34 percent in 1994 to 19 percent in 2012. By contrast, the share of households with incomes greater than $65,000 nearly doubled between 1994 and 2012, growing from 14 percent to 27 percent. However, since the L.A. County graph is in static dollars whereas our other graphs present income groups in relation to AMI, any decrease across L.A. County could be explained by inflation.

Taken together, these figures paint a picture suggesting that shifts in the distribution of income over time and space are complex, and that the role of rail stations is not straightforward. The L.A. County pattern over 1994 to 2012 points to a broad-based dynamic in which lower-income households became less prevalent and higher-income households became more prevalent in the population. In that light, one should expect to see similar shifts in the income distribution in rail station areas. However, the data suggests there may be more to the story. The patterns for the rail station areas do not match exactly with those for the County. Moreover, we find that patterns vary across rail lines. These variations point to a possibility that more can be learned by examining individual behavior directly rather than aggregate statistics. We take to this exercise in the next section.

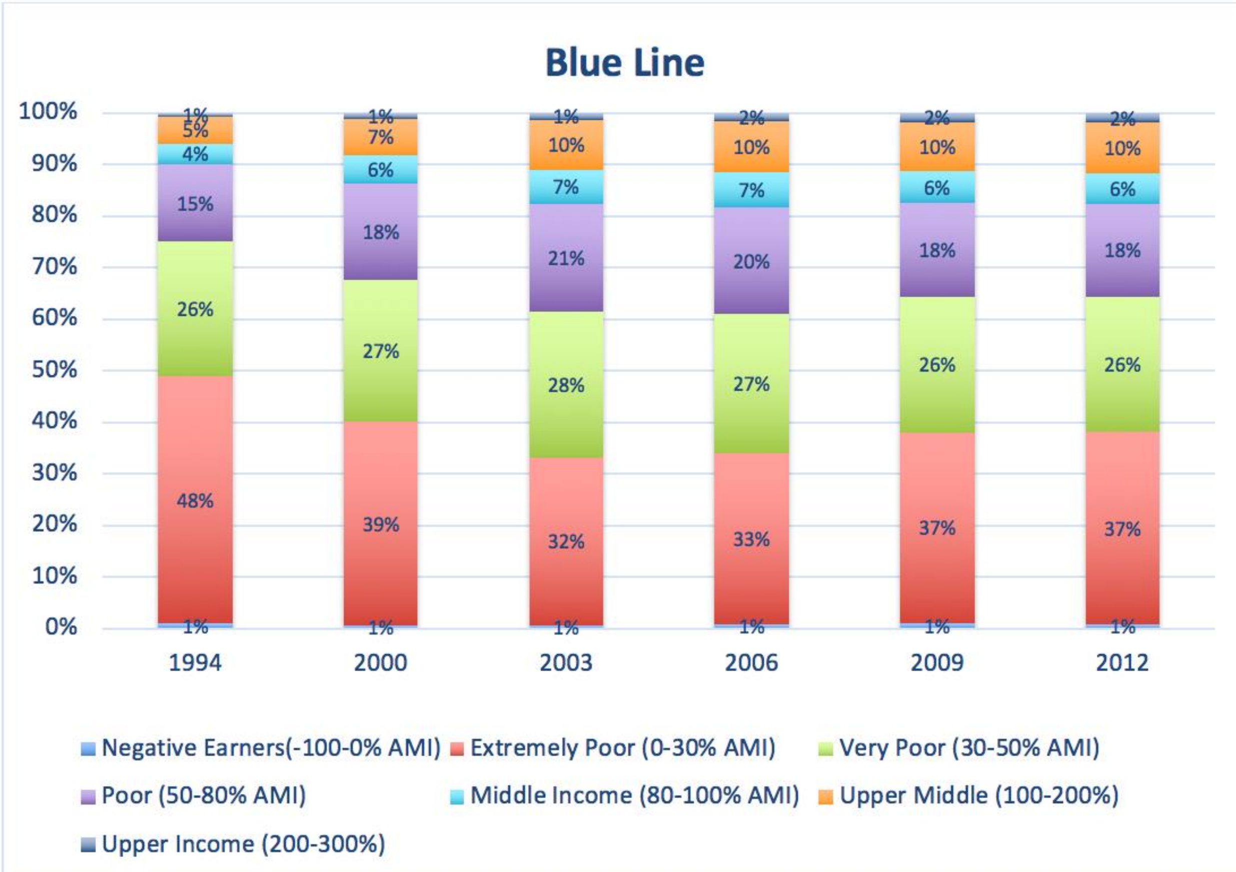

Figure 2.A. Income Distribution over Time for Households Living in 9-digit Zip Codes within 0.5 miles of L.A. Metro Station, Blue Line
Source: FTB

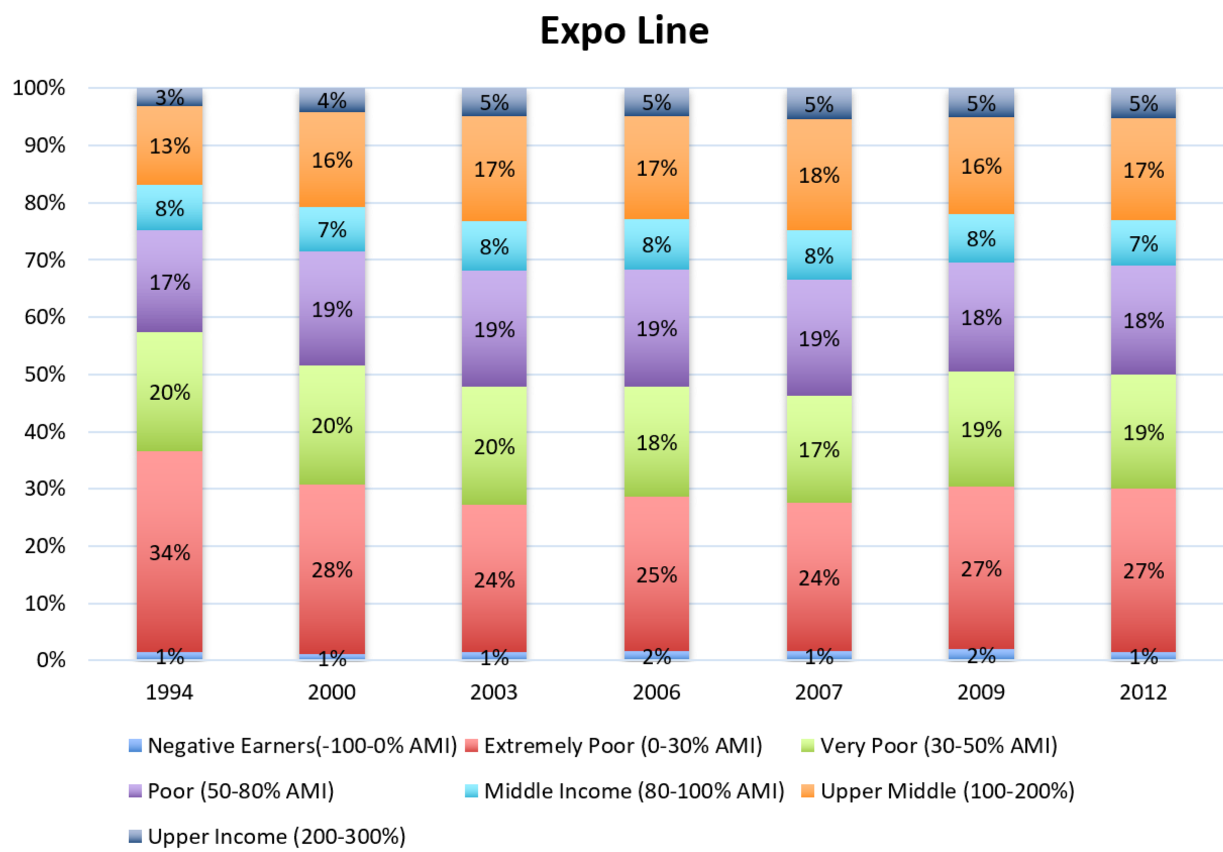

Figure 2.A. Income Distribution over Time for Households Living in 9-digit Zip Codes within 0.5 miles of L.A. Metro Station, Expo Line

Source: FTB

a

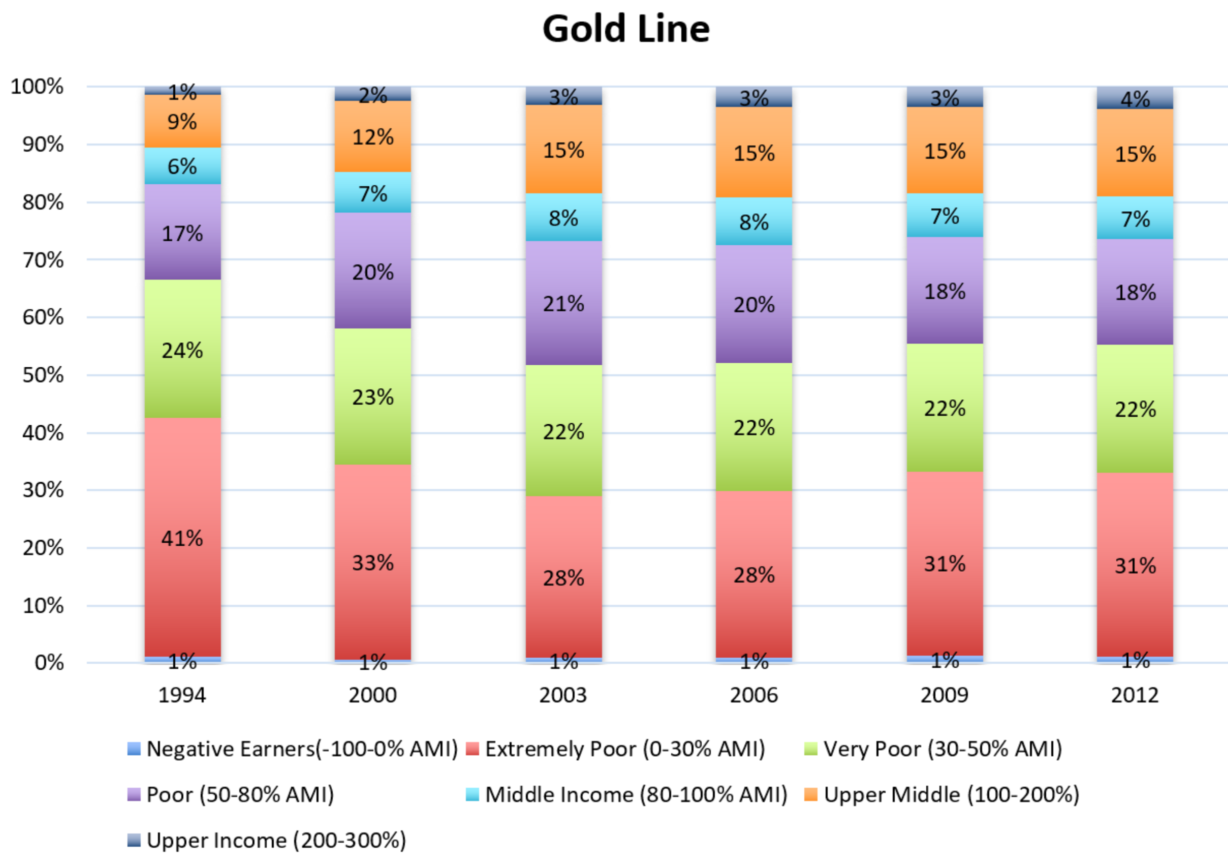

Figure 2.B. Income Distribution over Time for Households Living in 9-digit Zip Codes within 0.5 miles of L.A. Metro Station, Gold Line
Source: FTB

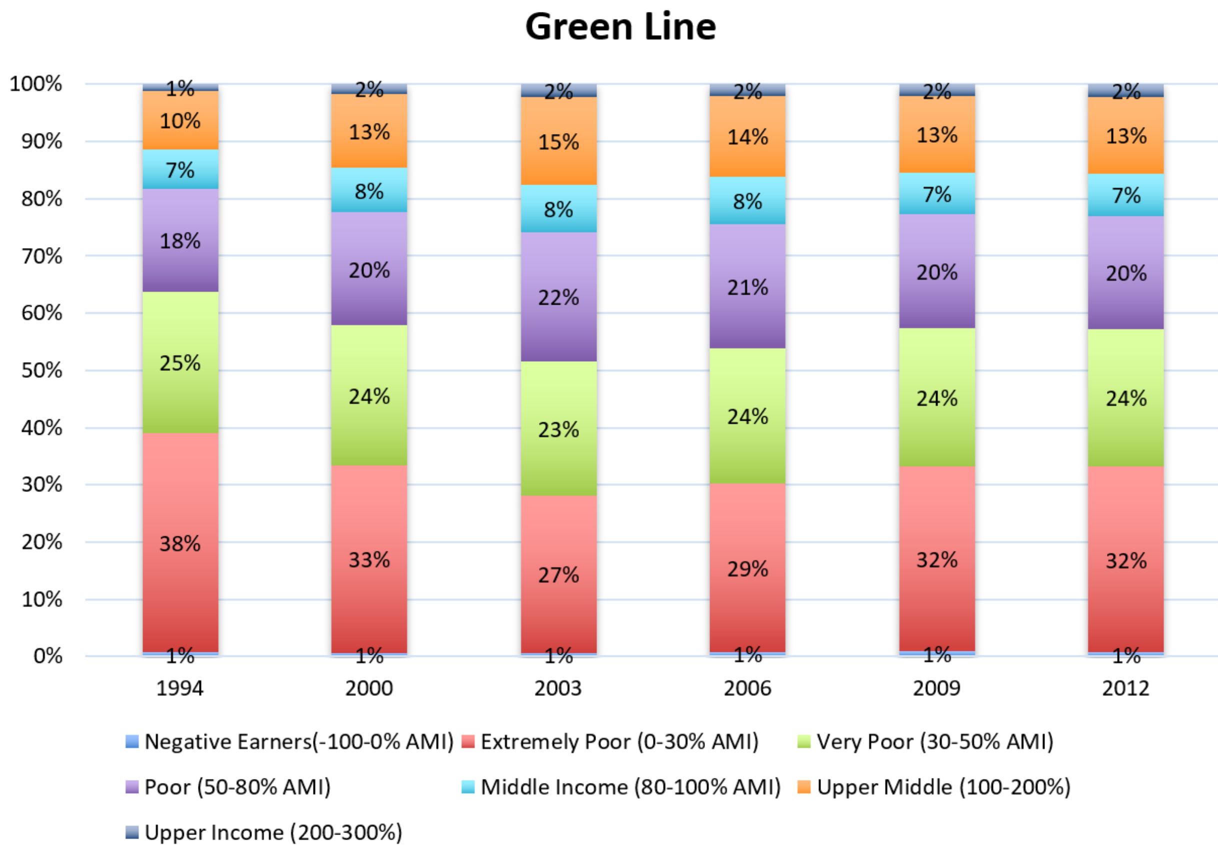

Figure 2.C. Income Distribution over Time for Households Living in 9-digit Zip Codes within 0.5 miles of L.A. Metro Station, Green Line

Source: FTB

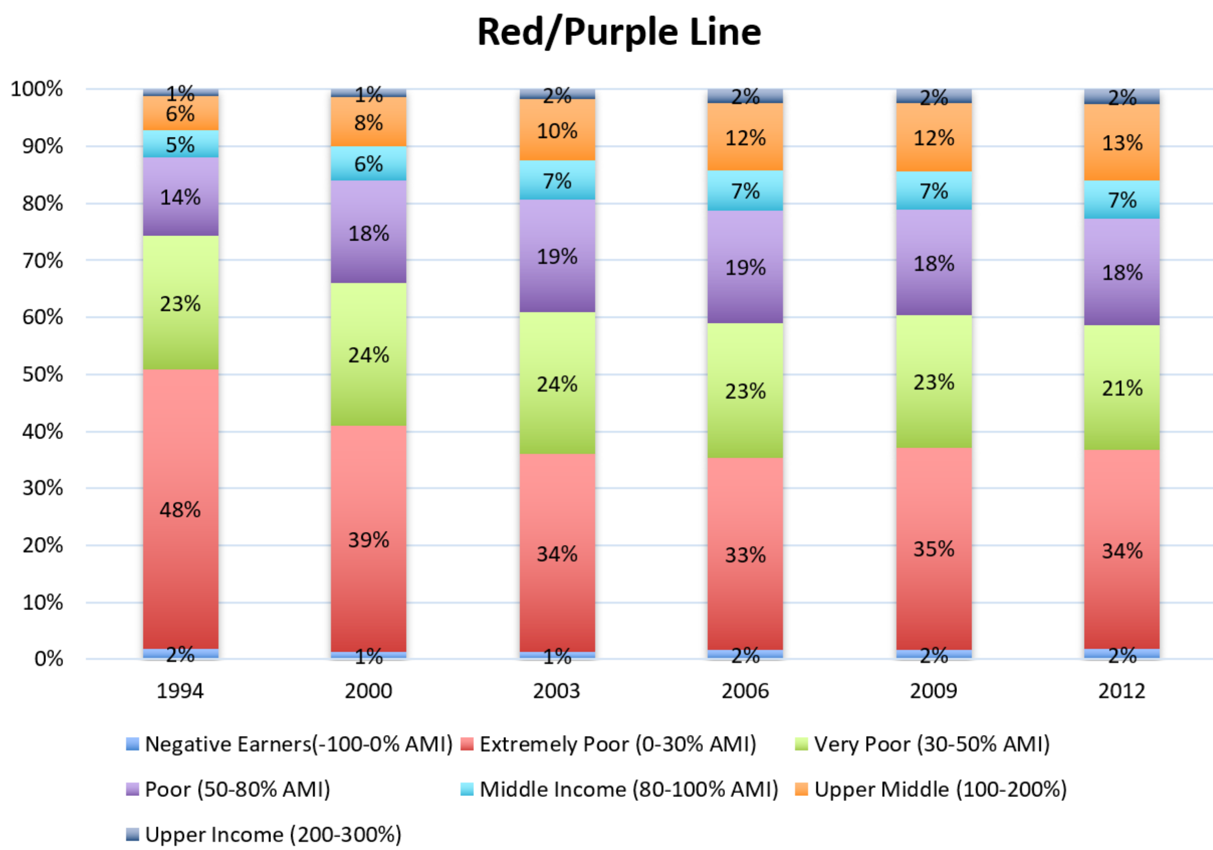

Figure 2.D. Income Distribution over Time for Households Living in 9-digit Zip Codes within 0.5 miles of L.A. Metro Station, Red/Purple Line
Source: FTB

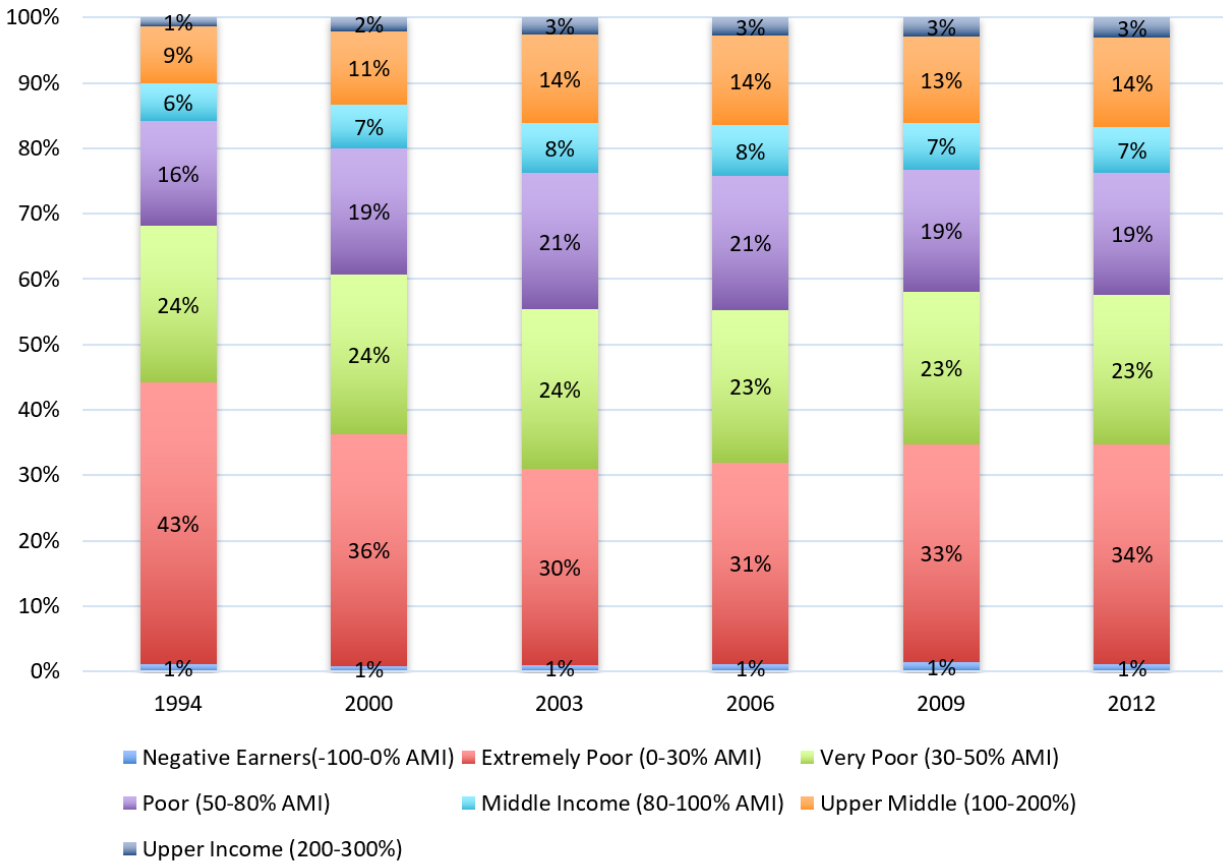

Figure 2.E. Systemwide Income Distribution over Time for Households Living in 9-digit Zip Codes within 0.5 miles of L.A. Metro Station

Source: FTB

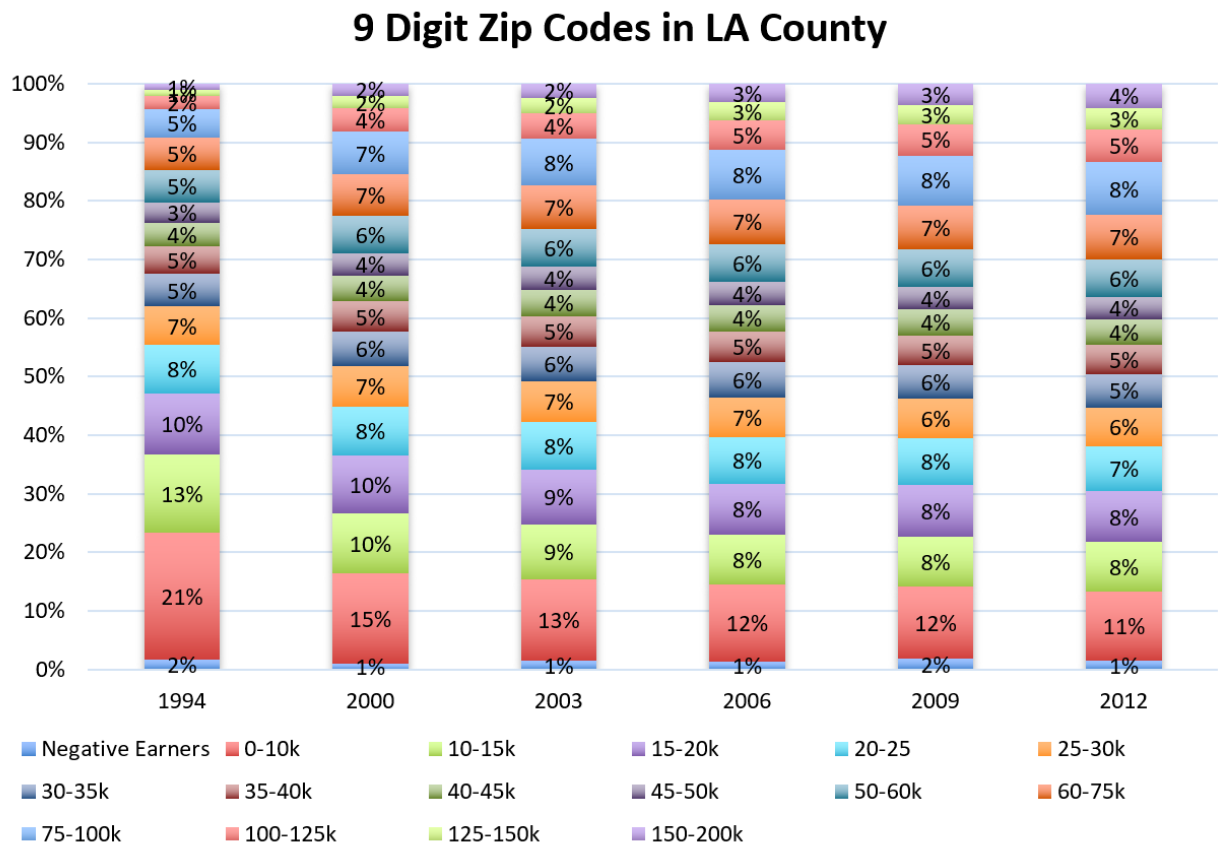

Figure 3. Income Distribution over time, by all households in the sample with 9-digit zip code locations
Source: FTB

Move Patterns Near L.A. Metro Rail Stations

We also track the number of households moving IN and OUT of L.A. Metro rail station neighborhoods, as well as the NET change in the number of households. As a reminder, we measure these transitions by tracking those households living in a station area in a given year that meet the selection criteria for creating the sample and appear in the data in an adjacent year. Thus, this analysis does not speak to overall changes in the demographics and income distribution of the area, which is also affected by households moving into the area from outside of California, households moving out of the area to places other than California, and households for whom data were not available in consecutive years for whatever reason. In addition, the discussion will not focus on trends for Negative Earners because, as discussed earlier, there are reasons to believe that these households are less relevant for the issues being considered.

With that caveat established, we begin by documenting rates of in-migration to rail station areas by income category. Table 8 shows that move-in rates were higher for lower-income households, especially Extremely Poor households. The Extremely Poor had the highest in-migration rate in 9 years and the Very Poor and Poor had the highest in-migration rate in 8 other years. Compared to in-migration rates in the earliest years of the sample, in-migration rates declined over time for all groups, except for increases seen across the board in 2009.

In assessing these rates, it is important to remember that the upper-income categories began from a smaller base than the other categories. Given this, it is also instructive to consider the numbers of households moving into a rail station area in a year (right panel of Table 8). Here we see that more Extremely Poor and Very Poor households moved into the rail station area than did households from other income categories. However, when considering changes in the volumes of household in-flows, we see that changes were most dramatic for the Upper Middle and Upper Income categories. But in terms of numbers, the volume of in-movers from these income categories still did not approach the volumes seen for less affluent households.

An interesting question is whether transition patterns differ in the more recent years compared to earlier in the study period. The last two rows of Table 8 – which show the average out-mover rates for the entire period and for 2009-2012 – offer some insight here. We see that rates of in-moving were lower in recent years than over the entire period for all income categories. The declines were between 11 and 14 percent for all categories except the Middle Income and Upper Middle Income categories, where the changes were small. Again, however, we see that the numbers of households moving in was higher in more recent years than overall. This variation, particularly in rates over time, suggests simple stories of in-migration that seek to identify a single factor as driving changes may miss important nuance.

Next, we turn to rates of out-migration (Table 9). As was the case for in-migration, the out-migration rates were higher for lower-income households, and particularly for the Extremely Poor, who had the highest rates of out-migration in more than half of the years in the sample. Also similar to the in-migration patterns, out-migration rates were higher more frequently for Very Poor and Poor households than for households with higher incomes. However, unlike the pattern for in-migration, household rates of out-migration did not fall appreciably over the study period. We find this to be especially true during the years with higher match rates (1998 to 2012).

The right panel of Table 9 shows the number of households in each income category that moved out in each year. As was the case for in-movers, more households in the poor categories moved out of the neighborhoods than did households in middle and upper income categories. Particularly for Extremely Poor and Very Poor households, these differences were sizable (ranging from, 1,000 to more than 2,500 households in the case of Upper Middle Income households) and remained so throughout the study period.

The final two rows of Table 9 again allow for a comparison of transitions in the more recent years with transitions over the entire period. The data show that rates of out-moving did not change in recent years, which is a notable difference from the pattern observed for in-moving. The largest change in the rate of out-moving was among Extremely Poor households, but this change was less than 5 percent. As was the case for in-moving, however, we observe increases in the *number* of households moving out across all income categories.

Finally, we report on the implications of these two analyses for net movements of households by income categories. To conduct this exercise, we take the difference between the number of in-movers and the number of out-movers in an income category, and then calculate the net rate based on the number of households in the initial year. For this, it is important to keep in mind that our sample includes only those households that met our inclusion criteria in consecutive years. As such, we are capturing a partial picture of net movements.

Table 10 shows that net out-movement is the prevailing trend among households when considered in terms of the income categories. In only 8 cases – a case is an income category-year combination – do we observe net in-migration during our sample period. Comparing experiences across income categories, we see that until 2008, the Extremely Poor had the highest rate of out-migration, with the lone exception being in 2003 when net outmigration among the Middle Income and Upper Middle Income households was the highest (a tie). From 2008 to 2012, we observe a different pattern. In 2008 the Middle Income category had the highest rate of out-migration. Interestingly, and perhaps counterintuitively, since 2009 we see that the highest rate of out-migration is in the Upper Income category.

As before, because the base number of households differs across income categories, it is useful to look at the aggregate number of net movers as well. We see that the Extremely Poor category consistently has the largest number of net outward movement, and the number of such households moving out on net in a year remained relatively stable over the study period. Though the Upper Income category had the highest net out-migration rate for 2009 to 2012, this appears to be a function of the small base number of such households. During these years, the net number of out-movers trailed the numbers of net out-movers for all poor income categories, sometimes by sizable amounts. Finally, when there was net in-migration, the net change in the number of households was quite small. In only three cases did the number of net in-movers exceed 25 households.

Finally, as before, the final two rows in Table 10 show how net movement occurred for the entire study period and for the most recent 4 years. In term of rates of net movement, the data show relatively few differences among families in the middle of the income distribution. However, for families in the extreme categories, we do observe significant variation. Among the Extremely Poor, net rates of out migration were more muted in the most recent years as compared with the rest of the sample period. We observe a net out migration rate that is about 25 percent less than the overall rate. By contrast, the net out migration rate for Upper Income households increases dramatically. The rate of net out migration between 2009 and 2012 is nearly twice the rate when the entire study period is considered.[4]

[4] Appendix Table 17 reports on the number of households moving in to and out of rail station areas for 1994, 2003 and 2012 by rail line. Appendix Table 18 reports on the net movement of households per year by rail line.

One should not infer from the tables that there was a NET loss of households in the system. We know the number of households in the system has grown steadily up until at least 2006, and we believe the negative NET change in the year-to-year number of households is due to our requirement for households to file in L.A. County and be geocodable in consecutive years. Because our method of determining household mobility requires filers to be in the data in consecutive years, we suspect our count of households who moved in (and the corresponding rate), does not account for households who newly enter the workforce (i.e., college students) and does not account for out-of-state in-movers (including immigrants). That said, our net mobility rates are comparable to county-to-county or domestic movers as presented by the St. Louis Federal Reserve and California's Legislative Analyst Office.

Table 8. Households moving in to of Neighborhoods near L.A. Metro Station Areas by Year and Income, rates and numbers

	By Rate of In-Moving (percent)							By Number of Households						
Year	Negative Earners	Extremely Poor	Very Poor	Poor	Middle Income	Upper Middle Income	Upper Income	Negative Earners	Extremely Poor	Very Poor	Poor	Middle Income	Upper Middle Income	Upper Income
1994	7.6	14.5	13.6	13.5	10.5	9.0	10.6	26	1,472	882	465	142	171	31
1995	8.1	12.8	12.5	10.2	10.2	7.2	7.9	40	2,574	1,606	863	248	298	47
1996	8.7	11.9	11.1	10.1	7.6	7.4	6.5	61	2,555	1,599	977	280	370	56
1997	6.8	11.6	11.0	9.6	8.7	6.3	6.3	49	2,405	1,749	1,071	336	484	87
1998	5.9	9.8	10.3	9.4	8.8	6.7	7.0	48	2,249	1,738	1,203	386	535	105
1999	7.9	9.0	9.0	8.5	8.0	7.1	6.1	47	2,271	1,677	1,178	383	624	118
2000	7.5	8.5	8.7	9.1	7.7	6.5	4.6	59	2,204	1,708	1,219	377	593	112
2001	4.5	8.7	7.2	7.1	7.1	6.3	5.4	41	2,243	1,717	1,226	446	673	111
2002	8.2	8.1	7.8	8.2	6.5	6.3	5.4	63	2,425	1,805	1,430	425	659	129
2003	8.1	7.9	8.2	7.6	7.8	8.0	8.2	70	2,225	1,724	1,281	459	730	154
2004	7.3	7.9	8.8	8.0	8.1	7.6	10.6	71	1,936	1,793	1,448	530	816	189
2005	6.6	7.9	8.1	7.7	6.8	8.0	8.9	60	2,113	1,827	1,445	518	862	205
2006	6.3	8.3	9.0	8.5	7.1	7.2	7.7	85	2,220	1,992	1,769	620	1,046	235
2007	6.2	8.6	8.4	8.7	8.6	7.7	7.5	94	2,369	2,030	1,747	665	1,142	241
2008	7.4	9.3	8.4	8.1	7.9	6.9	6.4	111	2,549	2,134	1,878	677	1,239	272
2009	7.4	9.9	9.2	9.2	9.1	8.3	7.7	127	3,367	2,387	1,897	760	1,383	262
2010	6.2	8.1	8.3	8.3	7.9	6.4	6.5	126	3,056	2,267	1,709	648	1,116	199
2011	5.6	7.7	7.9	7.6	7.1	6.1	5.9	100	3,045	2,171	1,619	581	1,064	224
2012	5.8	8.2	7.2	7.3	6.7	6.6	5.1	112	2,923	2,021	1,466	547	1,094	239
Average														
Overall	6.9	9.4	9.2	8.8	8.0	7.1	7.1	73	2,432	1,833	1,363	475	784	159
2009-12	6.3	8.5	8.2	8.1	7.7	6.9	6.3	117	3,098	2,212	1,673	634	1,164	231

Source: FTB

Table 9. Households moving out of Neighborhoods near L.A. Metro Station Areas by Year and Income, rates and numbers

	By Rate of Out-Moving (percent)							By Number of Households						
Year	Negative Earners	Extremely Poor	Very Poor	Poor	Middle Income	Upper Middle Income	Upper Income	Negative Earners	Extremely Poor	Very Poor	Poor	Middle Income	Upper Middle Income	Upper Income
1994	6.0	13.7	13.5	11.4	10.4	9.0	4.5	40	3,246	1,623	866	263	334	44
1995	10.0	12.0	10.9	9.9	8.6	8.6	11.0	44	3,044	1,778	1,082	317	425	64
1996	9.0	11.0	11.5	9.7	8.4	7.8	10.6	82	3,091	1,788	1,061	319	463	82
1997	7.0	11.9	9.9	10.9	9.1	8.1	7.2	197	3,051	1,723	1,215	351	566	92
1998	6.4	9.4	9.4	9.8	9.4	7.6	9.5	63	2,795	1,837	1,212	410	606	122
1999	7.2	8.8	8.9	7.8	8.1	8.3	8.5	62	2,819	1,841	1,329	409	640	127
2000	6.7	10.3	8.5	8.1	8.4	7.7	9.7	57	2,847	1,836	1,362	498	760	139
2001	7.3	8.3	8.1	8.2	7.9	8.2	13.2	63	2,954	2,002	1,460	491	794	160
2002	6.3	8.9	8.3	7.9	7.6	8.5	9.7	73	2,909	1,881	1,388	458	815	157
2003	6.9	8.3	8.4	7.9	8.7	7.6	8.2	53	2,466	1,968	1,612	592	960	183
2004	8.3	8.7	8.7	8.6	8.2	7.6	8.0	83	2,887	2,182	1,672	596	893	168
2005	6.9	8.8	8.5	8.1	8.7	7.1	7.3	73	3,060	2,316	1,799	628	960	194
2006	6.5	8.2	9.4	7.9	7.0	6.4	7.0	83	2,865	2,163	1,757	614	1,012	221
2007	7.4	9.5	9.1	9.4	9.0	7.7	6.8	99	3,198	2,512	2,104	811	1,330	277
2008	8.0	10.0	9.7	9.3	10.3	9.4	8.5	129	3,259	2,647	2,216	905	1,512	311
2009	7.1	9.7	10.3	9.6	9.5	9.2	9.7	126	3,962	2,734	2,093	795	1,436	328
2010	7.1	9.0	8.5	9.0	8.4	7.9	8.2	131	3,596	2,393	1,924	694	1,284	297
2011	8.0	9.2	9.2	8.4	7.7	8.2	7.7	118	3,575	2,384	1,769	619	1,334	324
2012	5.6	9.6	9.1	9.0	9.0	9.2	9.4	99	3,608	2,276	1,772	676	1,467	354
Average														
Overall	7.2	9.8	9.5	9.0	8.7	8.1	8.7	88	3,117	2,099	1,563	550	926	192
2009-12	7.0	9.4	9.3	9.0	8.7	8.6	8.8	119	3,685	2,447	1,890	696	1,380	326

Source: FTB

NCST

Table 10. Net Change in Number of Households in Neighborhoods near L.A. Metro Station Areas by Year and Income (Moving IN minus Moving OUT), rates and numbers

	By Rate of Net-Moving (percent)							By Number of Households						
Year	Negative Earners	Extremely Poor	Very Poor	Poor	Middle Income	Upper Middle Income	Upper Income	Negative Earners	Extremely Poor	Very Poor	Poor	Middle Income	Upper Middle Income	Upper Income
1994	-2.2	-8.1	-6.1	-4.9	-4.1	-3.7	-1.8	-14	-1,774	-741	-401	-121	-163	-13
1995	-0.6	-2.0	-1.2	-2.1	-1.9	-2.1	-1.8	-4	-470	-172	-219	-69	-127	-17
1996	-2.5	-2.0	-1.2	-0.7	-1.0	-1.4	-2.3	-21	-536	-189	-84	-39	-93	-26
1997	-16.5	-2.3	0.1	-1.1	-0.3	-1.1	-0.4	-148	-646	26	-144	-15	-82	-5
1998	-2.0	-1.8	-0.5	-0.1	-0.5	-0.9	-1.2	-15	-546	-99	-9	-24	-71	-17
1999	-2.0	-1.7	-0.8	-0.9	-0.5	-0.2	-0.5	-15	-548	-164	-151	-26	-16	-9
2000	0.3	-2.0	-0.6	-0.8	-2.0	-1.6	-1.4	2	-643	-128	-143	-121	-167	-27
2001	-2.7	-2.0	-1.2	-1.3	-0.7	-1.2	-2.6	-22	-711	-285	-234	-45	-121	-49
2002	-1.1	-1.5	-0.3	0.2	-0.5	-1.5	-1.5	-10	-484	-76	42	-33	-156	-28
2003	1.8	-0.8	-1.0	-1.7	-1.8	-1.8	-1.2	17	-241	-244	-331	-133	-230	-29
2004	-1.2	-2.9	-1.5	-1.1	-0.9	-0.6	0.9	-12	-951	-389	-224	-66	-77	21
2005	-1.2	-2.8	-1.9	-1.6	-1.3	-0.7	0.4	-13	-947	-489	-354	-110	-98	11
2006	0.1	-1.8	-0.6	0.1	0.1	0.2	0.4	2	-645	-171	12	6	34	14
2007	-0.3	-2.4	-1.7	-1.4	-1.5	-1.1	-0.9	-5	-829	-482	-357	-146	-188	-36
2008	-1.2	-2.0	-1.8	-1.4	-2.5	-1.6	-1.1	-18	-710	-513	-338	-228	-273	-39
2009	0.1	-1.5	-1.2	-0.9	-0.4	-0.3	-1.9	1	-595	-347	-196	-35	-53	-66
2010	-0.3	-1.3	-0.4	-0.9	-0.5	-1.0	-2.8	-5	-540	-126	-215	-46	-168	-98
2011	-1.1	-1.3	-0.8	-0.7	-0.5	-1.7	-2.8	-18	-530	-213	-150	-38	-270	-100
2012	0.9	-1.7	-0.9	-1.4	-1.6	-2.3	-3.1	13	-685	-255	-306	-129	-373	-115
Average														
Overall	-1.4	-2.1	-1.1	-1.1	-1.1	-1.2	-1.4	-15	-686	-266	-200	-75	-142	-33
2009-12	-0.1	-1.5	-0.8	-1.0	-0.8	-1.3	-2.7	-2.3	-588	-235	-217	-62	-216	-95

NCST

Summarizing the findings thus far, six points stand out. First, we observe a large decline in the share of households in the Extremely Poor category over the study period, though there was a slight rebound in the share in 2009 that persisted to 2012. Second, the share of households with higher incomes increased from 1994 to 2003 and then remained largely stable from that year forward. Third, the Extremely Poor generally had the highest move-out and move-in rates over the study period. Fourth, move in rates declined for all income categories from 1994 to 2012. Fifth, move out rates remained stable for all income categories for most of the study period. Sixth, the data on numbers of households shows that move movers, whether looking at moves in or moves out, where Extremely Poor or Poor.

Regression Setup and Results

The descriptive statistics point to a number of interesting observations, but they do not inform us whether mobility overall and mobility across incomes was impacted by the opening of rail stations. To answer this question, we set up regression models to test the effect of rail station openings on the mobility of households in the surrounding neighborhoods while controlling for neighborhood-specific and year-specific idiosyncrasies.

Regression Setup

In our statistical tests, we test the effect of rail station openings by comparing a station's mobility rate before and after opening, and by comparing those trends for stations that newly opened at some point during the period with trends for stations that were open continuously through the period. For these analyses, we include year and station fixed effects. We report additional regressions that restrict or relax sample selection criteria to test the sensitivity and robustness of our models.

Our specification measures the effect a rail station opening on the in-move, out-move, and net-move rates after differencing out station and year fixed effects and average mobility in unopened stations. Equation 1 describes this specification. Let *s* index stations, *t* index years, and *p* index income bins. $r_{p,t,s}$ is a mobility rate (in, out, or net), α is a constant term, $\lambda_{p,s}$ is a station-level fixed effect, $\gamma_{p,t}$ is a year fixed effect, and $D_{p,s,t}$ is a dummy for whether a station *s* has opened in year *t*. The coefficient of interest on the dummy variable, β, measures the impact of station opening on in-, out-, and net-mobility rates. Specifically, β is the average change in a mobility rate across all years after a station has opened. For example, if the average annual mobility rate before a station opens is 10% then a β=.01 implies that the average annual mobility rate after a station opened is 11%. We fit Equation 1 to each income category *p* separately to assess how each income group was affected by station opening.

$$r_{\rho,t,s} = \alpha_p + \lambda_{p,s} + \gamma_{p,t} + \beta_p D_{p,s,t} + \varepsilon_{pst} \qquad (1)$$

We cannot test the rail station opening effect on the entire system because not all stations opened during the span of our data. All Blue line stations opened before 1994 while some stations along Expo and Gold lines opened after 2013 (Table 11). Consequently, we can only test the effect on stations that opened between 1994 and 2012 along the Gold, Red, Green, and Expo lines. In all years that a station opened, the stations that are not yet in operation serve as controls. In example, the Gold and Expo stations that opened after 2013 serve as the controls for Expo line stations opened in 2012. In total, 54 rail stations were opened on the Expo, Gold, Green, and Red Metro lines.

Table 11. L.A. Metro Stations by Year Opened

L.A. Metro Rail Line	Years Stations Opened	Number of Stations	Number of Stations Opened within Data Span (after 1993 & before 2013)
Blue	1990	22	
Expo	2012, 2016	17	10
Gold	2003, 2009, 2016	26	20
Green	1995	13	13
Red	1993, 1996, 1999, 2000	15	11

To ensure precise estimation, we use all station observations with at least two 9-digit zip codes near them consistently throughout our data time span. The nature of our geocoding strategy requires 9-digit zip codes to be near L.A. Metro rail stations. The higher the number of 9-digit codes in a station area, the more precisely we are able to estimate the number of households and mobility rates. Conversely, too few 9-digit codes can impact the precision of the estimates and bias our coefficients. More importantly, too few zip codes suggest that the stations are not located in residential areas. Stations with few surrounding 9-digit zip codes are Downtown Long Beach (Blue line), Union Station (Red line), Monrovia (Gold line), and Aviation / LAX and Mariposa (Green line) (see Appendix Table 15).

Given our complex geocoding and sample restriction strategies, we also measure how sensitive our model is to several robustness checks, as variations of the main model described in Equation 1 above.

The regression described in Equation 1 measures the effect of station openings on mobility, controlling for year and station fixed effects. In Robustness Model 1, we remove station fixed effects to understand how coefficients change when we do not account for factors common to station areas that do not change over time. In Robustness Model 2, we account for the fact that the Green and Expo lines opened toward the beginning and end of our available data period. In this robustness check, we exclude the observations from these two lines to see if this materially changes the analysis results. In Robustness Model 3, we restrict our sample to only stations that opened within the span of our dataset by dropping certain Expo and Gold line stations that opened after 2012. This checks whether our regression results are sensitive to the selection of untreated stations. In Robustness Model 4, we include stations with less than two 9-digit zip codes to see whether that exclusion materially affected regression results. In Robustness Model 5, we include the high-count outlier 9-digit zip codes in non-residential areas to check whether their exclusion materially affected regression results. Table 12 summarizes these 5 robustness models.

Table 12. List of Robustness Models and Definitions

Model Name	Model Definition relative to Main Model (Equation 1)
Main Model	Includes available lines, fixed effects, and excludes certain zip codes
Robustness Model 1	Excludes station fixed effects from estimation equation
Robustness Model 2	Excludes Expo and Green Lines
Robustness Model 3	Excludes stations opened after 2012
Robustness Model 4	Includes stations with fewer than 2 9-digit zip codes
Robustness Model 5	Includes outlier high-count 9-digit zip codes

Our models include fixed effects for the station areas and time effects, and so can control for unique (but time invariant) station area effects and time trends that affect all stations the same way (e.g., region-wide housing market or macroeconomic levels in each year). Nevertheless, for the estimated coefficient on station opening to be interpreted causally, station opening would need to be a plausibly exogenous shock to the station area. In our estimates, the year that a station opened is unlikely to be plausibly exogenous. Rail station openings are typically preceded by lengthy approval processes and construction and can be followed by a gradual build-up of amenities and neighborhood change.

For some station areas, changes in net-migration of households might have occurred before the station opened, implying that any measured effect understates a station effect. For the later-opening stations (particularly the Expo Line, but also the Gold Line), we only observe the early years of station openings, and changes in net-migration that might take years to unfold may have not fully played out. The mix of anticipatory signals and slow evolution may bias our rail station opening coefficient toward 0, which understates the impact of rail station opening on mobility rate.

Another modeling concern is the endogenous nature of a station's location and the amenities that surround it. We cannot assume stations are randomly placed throughout Los Angeles County. L.A. Metro stations can either be expected to be placed into underinvested neighborhoods to spur investment or to be placed into already desirable locations. Our regression can address this only indirectly. Our descriptive statistics (Figure 2) suggest that the proportion of higher income households has increased in station areas under study during our data time span. This could be because either stations are placed in gentrifying areas or it could be because rail stations induce gentrification. If the coefficients on rail opening are not significantly different from 0 or do not suggest that there is an outflow of lower income households and an inflow of higher income households after a rail station opened in our models, then we have evidence consistent with the view that rail stations are located in gentrifying areas. Of course, a rail station could expedite the gentrification of an already gentrifying area but our test cannot separate the two effects.

Regression Results

We estimate the model in Equation 1 separately for in moving, out moving, and net moving. In addition, we run estimates using all households pooled together as well as for each income category separately. Table 13 present the coefficient estimates using the pooled sample. The rail station opening coefficient is estimated to be negative and significant for in and out mobility, but not for net mobility. These estimates imply that once a station opens, the average rate at which households move into the station area decreases by 12% or by 0.98% percentage points from an all-year and all-station average of 8.5%. The average rate at which households move out of the station area decreases by 4% (0.36 percentage points from a base of 8.5%). Taken together, these two results imply a lower turnover rate of households in LA Metro rail station areas after a station opens. The non-significant coefficient on net mobility, however, indicates that a rail station opening does not impact population growth or decline in a station area. In effect, the significant effects on in and out mobility cancel each other out.

Table 14 shows the estimates obtained when the regressions are run on samples including households from a single income category. Starting with the in-move results, we find that both Extremely Poor and Upper Middle Income households reduce the rate at which they move *in* to station areas. Given their average annual bases of 9.4 percent and 7.1 percent respectively, these coefficients imply a 10% (0.95%/9.4%) reduction in the rate of moving in for Extremely Poor households and a 15% (1.2%/7.1%) reduction in the move in rate for Upper Middle Income households. Regarding out moves, we see significant coefficients for the Very Poor and Poor categories. The Very Poor reduce the rate at which they move *out* of station areas by about 11.6% (1.1%/9.5%) and the Poor reduce their out-move rate by about 7% (0.7%/9%) due to a station opening.

Turning to the analysis of net moves, we find significant coefficients for three income categories. The regressions for both the Extremely Poor and Upper Middle Income categories yielded negative coefficients, while the analysis of the Very Poor produced a positive coefficient. In all three cases, the magnitudes of these coefficients are large relative to the category base rates (Table 10). The net mobility rate for households in the Extremely Poor and Upper Middle Income categories increased by 0.41% (0.87%/2.1%) and by 120% (1.44%/1.2%), respectively. For the Very Poor, the net mobility rate an increased by 83% (0.91%/1.1%).

The results from the move-in and move-out regressions help to explain the net mobility patterns. For the Extremely Poor, the negative coefficient for net moving appears to be driven primarily by the significant reduction in the move-in rate after a rail station opens. The same dynamic holds for the Upper Middle Income category. A significant negative move-in effect, coupled with an insignificant move-out effect, resulted in a significant negative net-move effect. By contrast, in the case of the Very Poor, we have the opposite pattern – a significant reduction in the move-out rate with no change in the move-in rate leading to a reduction in the net-mover rate.

Table 13. Regression Results for All Income Groups pooled

Mobility Rate	Number of Observations	R-Squared	Rail Station Opening Effect (β)
In-move	1216	0.39	-0.0098***
Out-move	1216	0.65	-0.0036***
Net-mobility	1216	0.13	-0.0063

NOTE: Statistical significance: *** p>0.01, ** p>.05

Table 14. Regression Results by Income Group

Mobility Rate	Income Group	Number of Observations	R-Squared	Rail Station Opening Effect (β)
In-move	Negative Earners	1180	0.07	0.0022
	Extremely Poor	1215	0.40	-0.0095**
	Very Poor	1215	0.29	-0.0015
	Poor	1215	0.39	-0.003
	Middle Income	1212	0.17	-0.0058
	Upper Middle Income	1214	0.29	-0.0109***
	Upper Income	1173	0.15	-0.0143
Out-move	Negative Earners	1188	0.04	0.006
	Extremely Poor	1216	0.43	-0.0009
	Very Poor	1216	0.37	-0.011***
	Poor	1216	0.31	-0.007***
	Middle Income	1214	0.31	0.0024
	Upper Middle Income	1214	0.41	0.0035
	Upper Income	1173	0.18	-0.0073
Net mobility	Negative Earners	1173	-0.01	-0.0006
	Extremely Poor	1215	0.07	-0.0087***
	Very Poor	1215	0.10	0.0091***
	Poor	1215	0.09	0.0049
	Middle Income	1211	0.04	-0.007
	Upper Middle Income	1214	0.07	-0.0144***
	Upper Income	1160	0.03	-0.001

NOTE: Statistical significance: *** p>0.01, ** p>.05

Next, we compare Main model results to those of the Robustness models to gain confidence in our main estimates. Tables 19 and 20 in the Robustness and Sensitivity Appendix report the regression results for our key variables in Equation 1 for the Main model and the five Robustness Model variants as described in Table 12.

Appendix Table 19 reports on these results for estimates derived using the pooled sample. Robustness Model 1 excludes station fixed effects which, like our Main model, produces significant effects on in- and out-mobility rates, but the signs are opposite of our Main model, suggesting station fixed effects are indeed controlling for important unobserved station-specific characteristics. Robustness Model 2 excludes the Expo and Green lines, effectively only using Red and Gold lines to test station effects. In that, we observe that estimated signs and magnitudes are similar to our main Model but only the effect on out-mobility is significant. This suggests that a large part of the Main results may be driven by either the Expo or Green lines, especially on effects for in-moves. The estimated effect on in-mobility of Robustness Model 3 is largely the same as in our Main Model but the effect on out-mobility is much smaller and insignificant, implying that the out-mobility results may be sensitive to the choice of whether the unopened stations from Expo Line Phase II are in the sample. Robustness Model 4 includes stations with poorly-estimated mobility rates and suggests that both the effect on in- and out-mobility are sensitive to the inclusion of these stations when it comes to magnitude and significance of the estimates relative to our Main Model. Our Robustness Model 5 includes high-count 9-digit zip codes and the estimates that include these zips are virtually identical to our Main Model. The results are not sensitive to the inclusion or exclusion of these commercial districts. Overall, among all Robustness Models that vary the sample, the effect signs are consistent with our Main model which provides confidence in the Main Model results.

Appendix Table 20 reports the coefficient estimates for the rail station opening effect across the Robustness Models when the samples include households from a single income category. As in the results for overall mobility rates, the consistent significant and positive coefficient estimates in Robustness Model 1 suggest that station-fixed effects capture crucial unobserved station characteristics without which coefficient estimates are not consistent. In Robustness Model 2 that excludes the Green and Expo lines, the coefficients on the Very Poor and Poor are both negative and significant as in the Main Model, though the magnitude of the estimate for the Poor is larger and the estimated coefficients on the Extremely Poor and Upper Middle Income are negative but not significant. This may suggest that effects are stronger along some lines than others. Robustness Model 3, which excludes stations opened after 2012, yields results that are very similar to the Main Model coefficient estimates in both magnitude and significance except in the case of Poor households for whom coefficients are no longer significant. This suggests that results for the Poor may be sensitive to the choice of control stations but the overall similarity to the Main Model is reassuring. Robustness Model 4, which includes stations with few zip codes, is similar to Robustness Model 3 in that all signs and significance on coefficients are similar for all income groups except for the Poor. Robustness Model 5 results are virtually identical to the Main Model results implying that the high count commercial zip-codes do not impact the overall results.

Overall, it seems our results are robust to sample specification. The lone exception is the Poor category, where we observe that both significance and coefficient magnitude are sensitive to the sample cut. On balance, this is reassuring, since the overall story suggested by the Main Model is supported by most of our Robustness Models.

Discussion

The analysis offers some answers to the initial questions raised in this report. Results from both the descriptive statistics and the regression analysis suggest that the Extremely Poor appear to be less present in rail station areas compared to other income categories after a station opens. The regression analysis suggests that Upper Middle Income households are also less represented compared to other income categories (except the Extremely Poor) in rail station areas post-opening. The Very Poor, according to the regression analysis, seem to be relatively more common post-opening, though this finding conflicts somewhat with the general trends seen in the descriptive statistics. The results also point to mechanisms driving these trends. For those in the Extremely Poor and Upper Middle Income categories, the result of less presence appears to be due to reductions in the propensity of households from these categories to move into rail station areas. For the Very Poor, the result of greater presence seems to be due to reductions in their propensity to move out.

These findings point to a number of different issues that would appear to be worthy of study. Foremost among these is research that can generate a better understanding of the reasons behind the observed relationships. While we have identified that reductions in move -n rates among the Extremely Poor are an important factor in their reduced presence in rail station areas after a station opens, we do not know why this rate declines. There are several possibilities – a reduction in the availability of affordable housing is certainly a plausible one – and these merit some exploration. Similarly, we do not identify the reason(s) why the move in rate among households in the Upper Middle Income category declines when a rail station opens. Could it be that new housing product is introduced at price points that are too expensive for these households? Future research could shed light on this. Finally, why does the move out rate for Very Poor households fall? Is it because they become more sensitized to the competitive housing market and so are more likely to accept the protection of units subject to rent control? Or is it driven by some other reason? Might it be that the Very Poor households value the rail transit connectivity and have enough means to find ways to stay in rail transit neighborhoods? Future research that sheds light on these, and the other, questions can advance knowledge and help identify possible policy solutions when sought in the face of neighborhood transitions.

There are a number of caveats regarding this analysis that we would like to offer. The first pertains to the "purity" of the signal of regression coefficients and descriptive statistics. As we noted above, the opening of a rail station typically takes several years to be fully completed. This admits the possibility that households could have started responding to the opening prior to its completion. If true, migration effects could be spread over several years. For stations that opened early in our sample, some of the effect may have played out prior to our collecting data; for stations that opened late in our sample (particularly relevant for the Expo line), changes in migration may not have fully played out. This could bias effects down. Similarly, other changes in the station area that change the amenity mix could also have lengthy processes, some of which may have begun prior to the station's opening. These could affect migration patterns as well, and could bias observed station effects up.

We also note that this exercise did not study a true experiment. Station areas were not randomly located across Los Angeles County. Rather, siting decisions by MTA were driven by a number of specific factors, not the least of which was the existence of right-of-ways that facilitated the development of some of the rail lines (Gold and Expo are two examples). We attempt to address this lack of exogeneity by including fixed effects in the regressions, but it is possible that these are not sufficient.

A regularity we find in our data is that, among tax filers for whom we can track year to year mobility in L.A. Metro rail station areas, mobility numbers suggest that more households move out than move in. This is true for all income groups. This finding is in line with ACS data on county-to-county mobility but does not capture the growth in the number of households. We suspect that because we require households to be in the FTB data for two consecutive years to establish their mobility, our in-mobility rate only tracks households who have been in the labor-force and filed California taxes the year before. This means we miss households who newly enter the labor force and households moving in from and moving out to other states. A complete picture would also characterize the behavior of these important groups.

Overall, the findings from this analysis offer some evidence consistent with the stylized picture of displacement around rail stations, but also strongly suggests that this is only a part of the total story. There is variation in mobility patterns over time and across rail lines. Further, the variation in the drivers of mobility – moving in versus moving out – suggest that households are facing more complex decisions than that story might imply. Finally, it is important to recognize that there is a very high baseline level of mobility in these neighborhoods, independent of whether a rail station opens or not, which may suggest that other issues might be more important if one wishes to make measurable progress in reducing household residential mobility that is not discretionary.

Appendix

Calculating Residential Mobility Rates by Station

Here, we describe our approach to associating and aggregating tax filers up to L.A. Metro station area neighborhoods. In dense residential areas, we know the location of each tax filer up to a 9-digit zip code. In order to associate the filer to a station area we measure the distance $d(s, z)$ between each 9-digit zip code centroid z to each train station s. If the $d(s, z)$ is less than $\bar{d} = 800m$ (approximately 0.5 miles), then we associate zip code z to train station s.

However, some 9-digit zip codes are within 800 meters of multiple stations, so in order to minimize zip code overlap we make greater restrictions. Let S be the set of all stations and Z_s be the set of zip codes that will be associated with train station s. Associating a zip code to a train station should be the same as associating a station to a zip code. For each zip code, let $\overrightarrow{d_z}$ be the vector of distance from the zip code z to each of the stations. The station that will be associated with zip code z is simply the station that has the smallest distance to the centroid of the zip code (i.e., $s_z = \arg\min\{d(s, z) \leq \bar{d}$ for $s \in S\}$). Consequently, the set of zip codes that are associated to a station is the collection of zip codes to which that station was closest station within 800 meters of the zip code.

In order to go from household-level data to station-area mobility rates, we take the following steps. We process data in year pairs from 1993-1994 to 2012-2013. Let p index our income bins, {-100-0%, 0-30%, 30-50%, 50-80%, 80-100%, 100-200%, and 300% of the Area Median Income}. If household i earning in income bin p changed its zip code between two consecutive years t and $t + 1$, we code her move status $m_{i,p,t} = 1$ and $m_{i,p,t+1} = 1$ to indicate that the household moved and $m_{i,p,t} = 0$ if the household did not move. We refine move status by marking households who moved within a station area (i.e., household's zip code changed, but household remained in the station area) as ones that did not move. For each 9-digit zip code z in Los Angeles County, we count the total number of filers in each income bin $n_{p,z,t}$ and $n_{p,z,t+1}$ which yields the population size of zip code z in years t and $t + 1$.

Once we have the number of movers $m_{p,z,t}$ and population $n_{p,z,t}$ in year t and zip code z, we compute the mobility rate r for station s in year t, ignoring the in(out) designations, as the sum of the number of movers in each zip code associated with station s divided by the station population. We follow the same process when calculating the mobility rates for households occupying particular income bins (Equations 2 and 3). The process for move and population numbers that are agnostic to income is analogous to the one described here.

Our in-mobility estimate may be downward bias. Since we do observe some growth in the number of households near rail transit areas, in any year $n_{p,z,t} < n_{p,z,t+1}$. In turn, this implies that the in-mobility rate will *always* be based off of a higher base than the out mobility rate which suggests that the in-mobility rate may be downward bias. We re-ran the regression (not shown) using $n_{p,z,t}$ as the base for the in-mobility rate and results were not categorically different from the present ones. The provides some evidence that the choice of base is not what is driving our results or the negative net mobility in light of growth in the number of households.

$$r_{p,s,t}^{in} = \frac{\sum_{z \in Z_s} m_{p,z,t}}{\sum_{z \in Z_s} n_{p,z,t}} \qquad (2)$$

$$r_{p,s,t+1}^{out} = \frac{\sum_{z \in Z_s} m_{p,z,t+1}}{\sum_{z \in Z_s} n_{p,z,t+1}} \qquad (3)$$

9-Digit Zip Codes by Station Area Neighborhood

Table 15. Number of 9-digit Zip Codes per L.A. Metro Station

Station Name	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Downtown Long Beach	1	3	4	5	4	6	4	5	9	12	24	29	28	30	40	47	51	43	39	40
1st Street	25	27	35	43	60	57	68	70	68	75	74	78	89	95	112	122	132	123	115	102
Pacific Avenue	46	62	60	67	72	91	93	104	105	103	117	125	140	162	177	165	173	155	138	127
5th Street	26	37	36	40	37	45	50	55	55	57	58	64	71	90	102	95	98	94	81	74
Anaheim Street	57	83	86	89	96	102	115	134	129	117	118	126	146	174	189	183	191	161	151	144
Pacific Coast Highway	81	96	104	107	114	130	150	170	174	169	179	193	234	262	286	270	271	247	224	213
Willow	62	78	84	81	82	87	94	101	107	106	112	112	111	118	124	124	128	119	115	116
Wardlow	75	119	119	132	144	149	153	164	151	152	153	155	158	170	176	174	169	163	165	162
Artesia	9	11	13	14	13	13	12	13	17	16	19	19	19	21	21	20	21	19	19	20
Compton	85	112	110	110	115	116	117	124	132	125	128	134	153	158	163	169	179	175	166	156
Willowbrook / Rosa Parks	70	83	85	84	92	93	100	111	115	116	119	123	127	122	128	132	131	129	128	125
103rd Street / Watts Towers	67	98	98	111	119	131	127	138	149	150	162	172	163	169	172	179	185	179	177	181
Firestone	139	168	173	178	178	183	197	195	202	207	205	219	220	222	217	213	224	225	222	214
Florence	168	202	212	211	215	212	214	222	218	220	216	225	219	224	229	225	225	219	218	221
Slauson	106	128	134	136	147	146	149	152	151	150	154	161	143	135	143	150	156	152	148	142
Vernon	102	123	126	117	125	123	134	141	142	141	139	141	125	129	129	134	134	134	130	128
Washington	25	33	34	31	32	32	34	33	33	35	38	37	36	32	33	37	38	35	35	32
San Pedro Street	98	107	109	111	109	120	124	126	122	126	128	133	138	135	141	139	141	135	133	129
Grand / LATTC	8	10	11	9	10	12	14	14	13	17	18	17	18	20	20	17	20	18	16	14
Union Station		1	1	1	1	1	1	1	1	1	1	1	3	10	10	12	14	13	12	14
Civic Center / Grand Park	7	7	12	17	36	45	47	53	52	59	66	74	78	86	91	98	106	100	98	91
Pershing Square	13	20	21	23	32	35	40	53	52	53	60	72	79	101	133	147	166	154	146	139
7th Street / Metro Center	16	20	21	26	30	32	32	41	47	48	56	70	82	112	142	153	163	152	148	144
Westlake / MacArthur Park	86	120	114	104	119	134	160	177	178	173	185	209	231	248	311	292	293	247	214	198

Station Name	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Wilshire / Vermont	125	176	208	232	266	282	298	319	314	303	311	311	331	373	421	409	437	402	388	347
Wilshire / Normandie	115	167	191	216	237	251	253	264	274	279	298	316	351	396	416	428	436	374	353	342
Wilshire / Western	119	174	186	213	233	248	257	265	266	270	290	298	314	351	394	394	421	374	354	323
Vermont / Beverly	122	188	203	207	213	235	231	242	252	258	254	270	304	323	366	367	371	345	323	284
Vermont / Santa Monica	139	167	177	183	175	195	193	206	213	206	189	200	223	262	281	268	275	254	239	226
Vermont / Sunset	74	101	116	131	137	137	142	160	152	160	166	171	174	189	194	186	186	179	166	146
Hollywood / Western	93	133	141	148	163	192	218	221	222	221	237	268	285	316	342	314	327	302	265	240
Hollywood / Vine	36	47	60	63	63	70	72	76	82	83	90	95	110	129	143	121	137	119	108	97
Hollywood / Highland	70	96	107	116	127	136	141	156	155	156	160	181	185	205	216	208	209	181	171	155
Universal City / Studio City	36	56	60	66	66	74	74	71	75	76	76	79	94	96	105	105	110	98	93	88
North Hollywood	67	80	93	94	93	101	106	113	120	117	121	124	136	165	197	218	229	196	178	172
Norwalk	36	81	92	102	110	119	123	129	132	132	129	140	149	155	154	153	153	146	149	149
Lakewood Boulevard	89	131	143	144	162	168	168	175	177	173	178	179	182	192	199	190	203	194	184	179
Long Beach Boulevard	100	123	130	129	140	140	148	163	157	154	160	159	165	174	181	179	178	173	173	168
Avalon	65	94	99	101	108	107	102	111	112	114	118	119	129	126	136	133	126	131	129	120
Harbor Freeway	82	106	116	114	114	122	127	134	137	135	135	144	148	149	161	151	149	145	138	136
Vermont / Athens	35	44	54	60	61	61	66	74	74	70	69	74	75	76	78	75	71	70	66	65
Crenshaw	143	169	177	179	178	184	188	199	201	198	194	185	195	210	214	208	207	202	205	192
Hawthorne / Lennox	12	27	30	32	30	33	34	35	34	34	36	36	30	31	50	57	58	59	58	57
Aviation / LAX					1	1					1	1	1	1	2	3	4	3	2	2
Mariposa											1	1	1	1	2	2			1	
El Segundo	23	34	36	38	40	41	41	41	44	44	44	44	29	30	27	27	28	26	25	26
Douglas	2	3	3	4	5	5	7	7	6	7	8	8	7	11	12	8	7	7	8	8
Redondo Beach	32	50	59	61	65	84	93	95	96	92	101	107	104	113	127	114	109	104	95	90
Chinatown	65	75	81	83	90	98	106	110	116	119	119	126	137	142	142	136	129	128	129	119
Lincoln Heights / Cypress Park	49	71	75	80	87	89	91	102	107	108	111	108	123	132	129	129	133	124	122	115

Station Name	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Heritage Square	81	118	137	148	157	166	188	218	212	217	226	233	234	245	260	246	233	232	229	223
Southwest Museum	44	76	95	97	102	110	118	128	120	116	124	133	130	135	142	144	149	147	134	131
Highland Park	33	47	53	64	70	69	74	77	79	74	80	86	92	99	113	114	120	108	104	94
South Pasadena	21	29	36	35	44	41	43	45	51	52	54	57	69	88	105	111	111	108	99	90
Fillmore	23	37	37	40	41	37	45	52	46	47	56	59	65	68	79	78	84	79	73	65
Del Mar	82	106	111	110	111	126	138	140	151	150	160	178	201	225	248	250	249	235	221	208
Memorial Park	71	106	117	120	128	136	139	146	141	139	141	146	151	156	159	157	155	157	146	142
Lake	42	62	63	67	67	68	71	69	68	70	69	72	85	86	89	81	86	87	87	80
Allen	24	43	56	63	64	61	69	78	84	79	80	84	88	97	102	97	99	95	94	90
Sierra Madre Villa	45	65	70	80	82	79	82	98	97	98	95	103	108	122	140	133	132	132	133	121
Arcadia	17	23	24	25	26	27	30	31	32	34	33	35	35	37	40	32	36	32	28	27
Monrovia			1	1	1	1	1	1	1	1	1	1								
Duarte / City of Hope	84	106	113	108	117	123	128	140	139	141	141	135	146	148	161	157	153	148	143	139
Irwindale	19	27	26	31	31	33	33	38	38	34	35	38	37	40	42	41	43	41	42	40
Azusa Downtown	8	11	12	12	15	17	18	19	18	18	22	23	35	44	59	60	70	64	61	58
APU / Citrus College	5	4	5	5	7	7	7	7	10	13	15	17	16	19	25	27	30	26	21	18
Little Tokyo / Arts District	56	74	79	71	75	83	90	91	88	84	87	95	99	106	114	114	109	103	101	98
Pico / Aliso	137	165	169	172	173	186	190	192	201	195	204	218	232	237	247	248	259	248	234	224
Mariachi Plaza / Boyle Heights	166	182	188	193	194	195	199	203	200	199	196	199	215	213	228	225	226	219	221	209
Soto	76	83	84	85	88	91	91	92	93	91	95	98	102	100	99	99	102	97	95	96
Indiana	48	50	52	53	53	58	61	61	58	58	57	57	62	69	69	71	69	66	65	65
Maravilla	52	68	75	77	82	82	89	93	92	85	91	91	96	100	100	95	98	97	94	95
East LA Civic Center	10	12	17	17	18	19	20	23	23	28	30	36	54	68	77	93	98	89	79	77
Atlantic	18	21	24	29	23	21	23	21	31	31	37	41	44	55	58	56	50	45	46	41
Pico	8	11	11	10	12	13	14	13	16	15	18	21	30	29	32	28	25	25	26	25
LATTC / Ortho Institute	63	80	79	83	90	89	95	96	100	100	101	105	113	125	135	134	135	122	119	115
Jefferson / USC	132	157	161	164	169	170	177	185	190	189	189	193	189	194	194	194	195	190	189	187

Station Name	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Expo Park / USC	63	101	110	118	121	127	130	133	142	136	133	141	134	134	140	139	144	141	134	130
Expo / Vermont	53	78	80	81	80	83	87	87	93	87	87	88	86	91	96	89	93	89	88	87
Expo / Western	75	92	98	100	102	113	115	110	111	109	106	111	122	130	133	131	132	131	127	121
Expo / Crenshaw	49	68	79	80	84	85	90	93	87	90	98	97	83	85	84	89	96	89	87	82
Farmdale	73	89	101	103	107	109	110	122	118	119	127	128	132	143	161	148	164	153	144	137
Expo / La Brea	74	143	162	181	183	205	225	235	228	233	244	245	268	315	358	337	354	313	293	261
La Cienega / Jefferson	75	97	105	110	113	115	119	120	123	122	121	120	109	109	113	111	114	113	108	105
Culver City	47	59	64	69	73	73	76	80	83	82	87	84	88	98	102	98	101	99	92	91
Palms	37	75	78	83	90	97	100	109	104	96	99	107	118	123	132	129	122	122	116	109
Westwood / Rancho Park	33	59	63	62	63	74	88	93	89	92	95	92	97	124	130	119	127	113	94	88
Expo / Sepulveda	25	39	48	55	58	72	77	79	73	70	67	74	81	90	105	84	88	77	72	63
Expo / Bundy	22	26	27	31	31	32	31	44	49	50	57	61	68	84	93	94	94	85	79	68

Source: FTB

Summary Statistics by L.A. Metro Rail Line

Table 16. Number of Households by L.A. Metro Line by Year

Year	Blue	Expo	Gold	Green	Red
1994	12,392	9,142	13,494	6,441	9,933
1995	14,024	11,129	15,953	7,845	11,944
1996	15,392	12,523	17,595	8,654	13,146
1997	16,998	13,499	19,137	9,382	15,067
1998	18,316	14,576	20,695	10,256	17,209
1999	19,690	15,755	22,497	11,176	18,840
2000	20,990	16,580	24,373	11,866	19,665
2001	22,344	17,045	25,293	12,266	21,097
2002	21,998	16,747	24,415	11,899	20,386
2003	22,931	17,253	24,858	12,145	20,612
2004	25,119	18,117	26,531	12,904	22,454
2005	25,937	18,746	29,241	13,133	24,634
2006	27,618	20,060	31,186	13,999	27,049
2007	28,761	20,645	32,840	14,259	28,160
2008	29,089	20,167	32,367	14,359	27,062
2009	29,656	20,425	32,606	14,489	27,646
2010	30,012	20,638	32,841	14,558	28,201
2011	29,806	20,235	32,338	14,400	27,113
2012	29,677	19,786	31,945	14,373	26,166
Average	**23,197**	**17,004**	**25,800**	**12,021**	**21,389**
18 year change in population	139%	116%	137%	123%	163%

Source: FTB

Table 17. By L.A. Metro Transit Line, Households Moving IN and OUT of Neighborhoods near L.A. Metro Station Areas by Year and Income

		Households Moving IN							Households Moving OUT						
Line	Year	Negative Earners (<0% AMI)	Extremely Poor (0-30% AMI)	Very Poor (30-50% AMI)	Poor (50-80% AMI)	Middle Income (80-100% AMI)	Upper Middle Income (100-200% AMI)	Upper Income (200-300% AMI)	Negative Earners (<0% AMI)	Extremely Poor (0-30% AMI)	Very Poor (30-50% AMI)	Poor (50-80% AMI)	Middle Income (80-100% AMI)	Upper Middle Income (100-200% AMI)	Upper Income (200-300% AMI)
Blue	1994	--*	477	251	97	21	26	--	--*	1002	497	217	47	53	--
Blue	2003	--	602 (26%)	469 (87%)	306 (215%)	98 (367%)	125 (381%)	--	--	657 (-34%)	518 (4%)	368 (70%)	127 (170%)	142 (168%)	22
Blue	2012	--	903 (50%)	578 (23%)	355 (16%)	99 (1%)	205 (64%)	43	--	1035 (58%)	633 (22%)	423 (15%)	127 (0%)	298 (110%)	68 (209%)
Blue	Annual Standard Deviation	--	108	69	49	16	16	--	--	114	64	40	27	23	8
Expo	1994	--	139	119	69	29	28	-	--	363	181	126	69	88	--
Expo	2003	--	313 (125%)	220 (85%)	195 (183%)	82 (183%)	143 (411%)	49	--	297 (-18%)	287 (59%)	239 (90%)	100 (45%)	192 (118%)	48 (182%)
Expo	2012	--	310 (-1%)	214 (-3%)	192 (-2%)	76 (-7%)	163 (14%)	40	--	380 (28%)	244 (-15%)	197 (-18%)	92 (-8%)	201 (5%)	56 (17%)
Expo	Annual Standard Deviation	--	54	44	35	20	32	--	--	45	40	34	19	35	9
Gold	1994	--	310	207	130	46	59	--	--	747	380	221	70	90	--
Gold	2003	--	534 (72%)	417 (101%)	337 (159%)	117 (154%)	220 (273%)	55	--	590 (-21%)	443 (17%)	401 (81%)	150 (114%)	272 (202%)	54 (391%)
Gold	2012	--	664 (24%)	486 (17%)	351 (4%)	148 (26%)	292 (33%)	79 (44%)	--	769 (30%)	454 (2%)	362 (-10%)	170 (13%)	388 (43%)	112 (107%)
Gold	Annual Standard Deviation	--	94	48	51	26	38	--	--	47	40	36	24	40	9
Green	1994	--	170	110	63	--	23	--	--	340	202	98	--	28	--
Green	2003	--	257 (51%)	222 (102%)	166 (163%)	60	99 (330%)	--	--	233 (-31%)	193 (-4%)	194 (98%)	63	73 (161%)	--
Green	2012	--	299 (16%)	212 (-5%)	160 (-4%)	48 (-20%)	105 (6%)	22	--	330 (42%)	211 (9%)	149 (-23%)	56 (-11%)	78 (7%)	28 (75%)
Green	Annual Standard Deviation	--	61	35	22	13	22	5	--	45	28	35	9	16	7
Red	1994	--	376	195	106	27	35	--	--	794	363	204	58	75	--
Red	2003	24	519 (38%)	396 (103%)	277 (161%)	102 (278%)	143 (309%)	29	21 (133%)	689 (-13%)	527 (45%)	410 (101%)	152 (162%)	281 (275%)	43 (378%)
Red	2012	43	747 (44%)	531 (34%)	408 (47%)	176 (73%)	329 (130%)	55 (90%)	43 (105%)	1094 (59%)	734 (39%)	641 (56%)	231 (52%)	502 (79%)	90 (109%)
Red	Annual Standard Deviation	--	89	60	44	21	23	9	13	93	44	51	26	39	11

Source: FTB; Percent change between year shown and prior year in parenthesis; * -- number suppressed confidentiality reasons

Table 18. By L.A. Metro Line, NET Change in Number of Households in Neighborhoods near L.A. Metro Station Areas by Year and Income (Moving IN minus Moving OUT)

Net Change in Number of Households (Households Moving IN minus Households Moving Out)								
Year	Line	Negative Earners (<0% AMI)	Extremely Poor (0-30% AMI)	Very Poor (30-50% AMI)	Poor (50-80% AMI)	Middle Income (80-100% AMI)	Upper Middle Income (100-200% AMI)	Upper Income (200-300% AMI)
1994	Blue	--*	-525	-246	-120	-26	-27	--*
1995	Blue	--	-100	-38	-63	-27	-21	--
1996	Blue	--	-150	-59	-28	--	--	--
1997	Blue	-159	-322	--	-54	--	--	--
1998	Blue	--	-150	-62	--	--	--	--
1999	Blue	--	-163	-48	---	--	--	--
2000	Blue	--	-126	---	---	---	--	--
2001	Blue	--	-128	-85	-42	--	--	-21
2002	Blue	--	-48	41	--	--	--	--
2003	Blue	--	-55	-49	-62	-29	--	-22
2004	Blue	--	-276	-133	-58	-27	-26	--
2005	Blue	--	-286	-165	-121	-26	--	--
2006	Blue	--	-163	--	--	--	--	--
2007	Blue	--	-226	-83	-77	--	-28	--
2008	Blue	--	-136	-136	--	-32	--	--
2009	Blue	--	-131	-57	-59	--	-27	--
2010	Blue	-22	-84	-25	--	-23	--	-26
2011	Blue	--	-69	--	--	--	-70	--
2012	Blue	--	-132	-55	-68	-28	-93	-25
1994	Expo	--	-224	-62	-57	-40	-60	--
1995	Expo	--	-36	--	-34	--	--	--
1996	Expo	--	-84	--	--	--	--	-20
1997	Expo	--	-47	21	--	--	--	--
1998	Expo	--	-63	--	-37	--	-55	--
1999	Expo	--	-121	--	-47	--	20	--
2000	Expo	--	-61	29	--	--	-48	--
2001	Expo	--	-29	--	-26	--	-57	--
2002	Expo	--	-28	55	36	--	-40	--
2003	Expo	--	--	-67	-44	--	-49	--
2004	Expo	--	-79	-36	-22	--	--	--
2005	Expo	--	-101	-88	-26	-38	-46	--
2006	Expo	--	-42	22	26	27	--	--
2007	Expo	--	-118	-63	-61	--	-	--

Net Change in Number of Households (Households Moving IN minus Households Moving Out)								
Year	Line	Negative Earners (<0% AMI)	Extremely Poor (0-30% AMI)	Very Poor (30-50% AMI)	Poor (50-80% AMI)	Middle Income (80-100% AMI)	Upper Middle Income (100-200% AMI)	Upper Income (200-300% AMI)
2008	Expo	-21	-62	-76	-51	-29	-84	-20
2009	Expo	-25	-71	-46	--	--	20	-26
2010	Expo	-22	-32	-42	-52	--	--	--
2011	Expo	-21	-35	--	--	--	--	-31
2012	Expo	21	-70	-30	--	--	-38	--
1994	Gold	--	-437	-173	-91	-24	-31	--
1995	Gold	--	-133	-32	-65	--	-47	-20
1996	Gold	--	-82	-56	-71	-22	-33	-24
1997	Gold	--	-128	-44	-51	--	-51	--
1998	Gold	--	-103	---	--	--	-21	--
1999	Gold	--	-120	-91	-45	--	--	--
2000	Gold	--	-188	-46	--	-37	-43	--
2001	Gold	--	-133	-67	-48	29	--	-20
2002	Gold	--	--	--	24	--	--	--
2003	Gold	--	-56	-26	-64	-33	-52	--
2004	Gold	--	-163	-56	-52	--	--	15
2005	Gold	--	-181	-88	-62	--	--	--
2006	Gold	25	-160	-35	37	--	28	--
2007	Gold	--	-152	-106	-96	-20	-40	--
2008	Gold	--	-208	-79	-83	-61	-43	--
2009	Gold	--	-72	-61	-35	--	48	--
2010	Gold	--	-115	25	-23	--	-84	-21
2011	Gold	23	-73	--	--	--	-44	-25
2012	Gold	25	-105	32	--	-22	-96	-33
1994	Green	--	-170	-92	-35	--	--	--
1995	Green	--	-31	-31	--	--	--	--
1996	Green	--	-63	-29	--	--	--	--
1997	Green	--	-69	36	--	--	--	--
1998	Green	--	-114	--	33	--	20	--
1999	Green	--	44	--	--	--	--	--
2000	Green	--	-29	--	--	--	--	--
2001	Green	--	-76	--	--	--	--	--
2002	Green	--	-73	-26	46	--	--	--
2003	Green	--	24	29	-28	--	26	--
2004	Green	--	-92	-36	--	--	--	--

Net Change in Number of Households (Households Moving IN minus Households Moving Out)								
Year	Line	Negative Earners (<0% AMI)	Extremely Poor (0-30% AMI)	Very Poor (30-50% AMI)	Poor (50-80% AMI)	Middle Income (80-100% AMI)	Upper Middle Income (100-200% AMI)	Upper Income (200-300% AMI)
2005	Green	--	-56	-35	--	--	--	--
2006	Green	--	-36	--	21	--	20	--
2007	Green	--	-45	-21	--	--	--	--
2008	Green	--	-31	-44	-53	--	--	--
2009	Green	--	-85	21	--	--	31	-29
2010	Green	--	-90	--	-41	--	--	--
2011	Green	--	--	31	--	--	--	--
2012	Green	--	-31	--	--	--	27	--
1994	Red	--	-418	-168	-98	-31	-40	--
1995	Red	--	-170	-57	-65	--	-52	--
1996	Red	-29	-157	-32	-10	--	-44	-22
1997	Red	20	-80	--	-54	--	-29	--
1998	Red	-20	-116	--	--	--	--	--
1999	Red	-29	-188	-35	-46	-24	-29	--
2000	Red	24	-239	-117	-138	-61	-56	--
2001	Red	--	-345	-139	-114	-47	-40	--
2002	Red	--	-326	-149	-77	-27	-92	--
2003	Red	--	-170	-131	-133	-50	-138	--
2004	Red	--	-341	-128	-85	-24	-66	--
2005	Red	--	-323	-113	-132	-29	-21	--
2006	Red	--	-244	-161	-57	-48	-21	--
2007	Red	--	-288	-209	-131	-83	-110	-24
2008	Red	--	-273	-178	-132	-119	-134	-20
2009	Red	--	-236	-204	-115	-42	-125	--
2010	Red	--	-219	-77	-99	-37	-86	-22
2011	Red	--	-351	-210	-137	-52	-127	-38
2012	Red	--	-347	-203	-233	-55	-173	-35

Source: FTB; *-- number suppressed for confidentiality reasons

Robustness and Sensitivity

High Filer Count ZIP Codes

In our data, some 9-digit zip codes had a high number of filers (200+) in any given year. To explore this potential anomaly, and to understand where these 9 digit zip codes are and why they have so many filers, we filtered out 9 digit zip codes whose number of filers was above the .999th percentile in any given year and mapped these zip codes using GIS software. The map on

Figure 4 reveals that there were a total of three 9-digit zip codes with such high counts near station areas each affecting a separate station. The maps in Figure 5 display these zip codes zoomed in on top of a street-level base map. As can be seen in the Figure 5, these zip codes are located in commercial areas, suggesting that households that filed taxes using these zip codes filed under the employment address as opposed to their residential address. Because households filing in these locations are clearly not representative of residential patterns, we exclude these 9-digit zip codes from our analysis on residents in proximity of L.A. Metro rail stations.

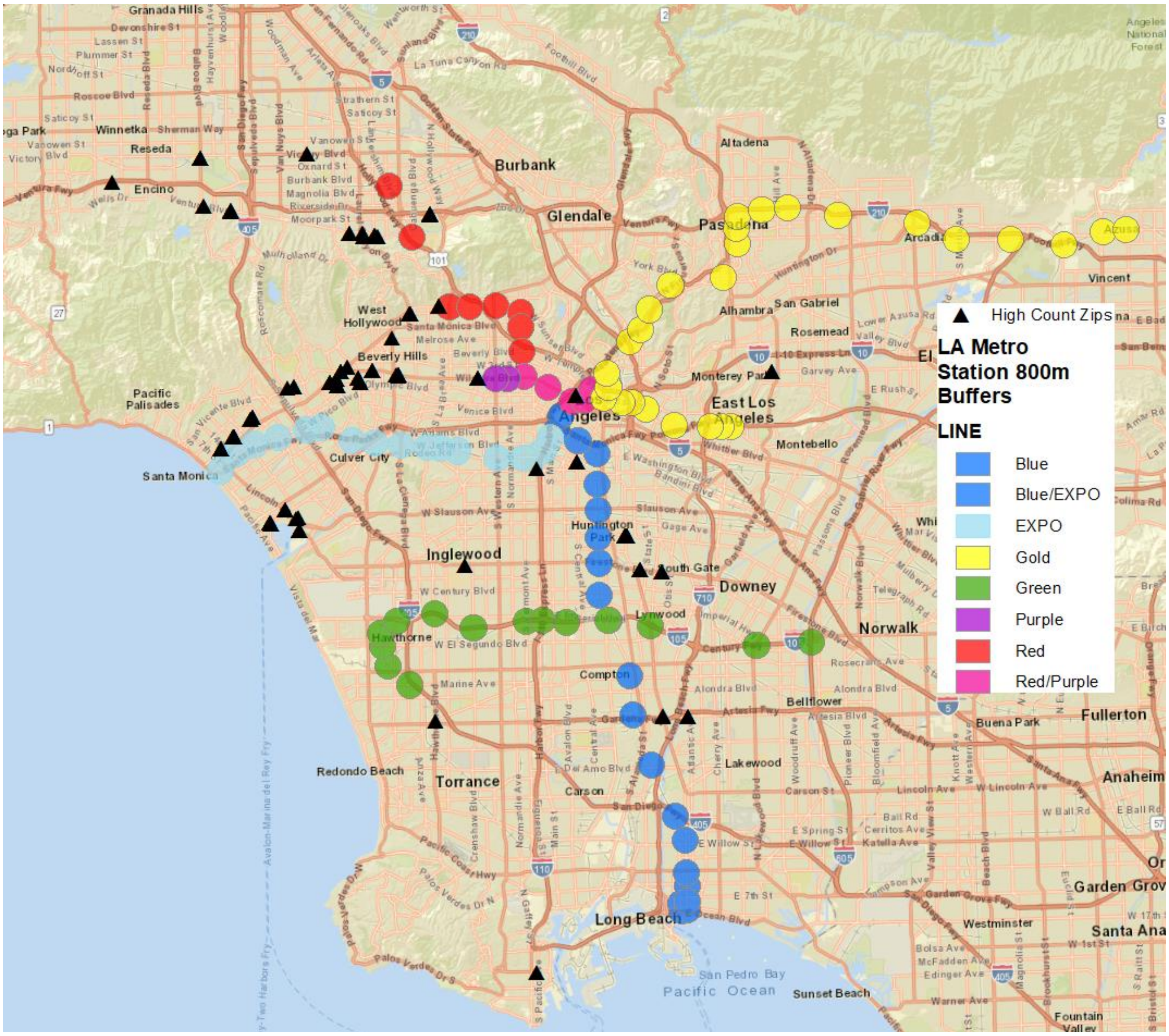

Figure 4. Map of High Filer County 9-digit Zip Codes

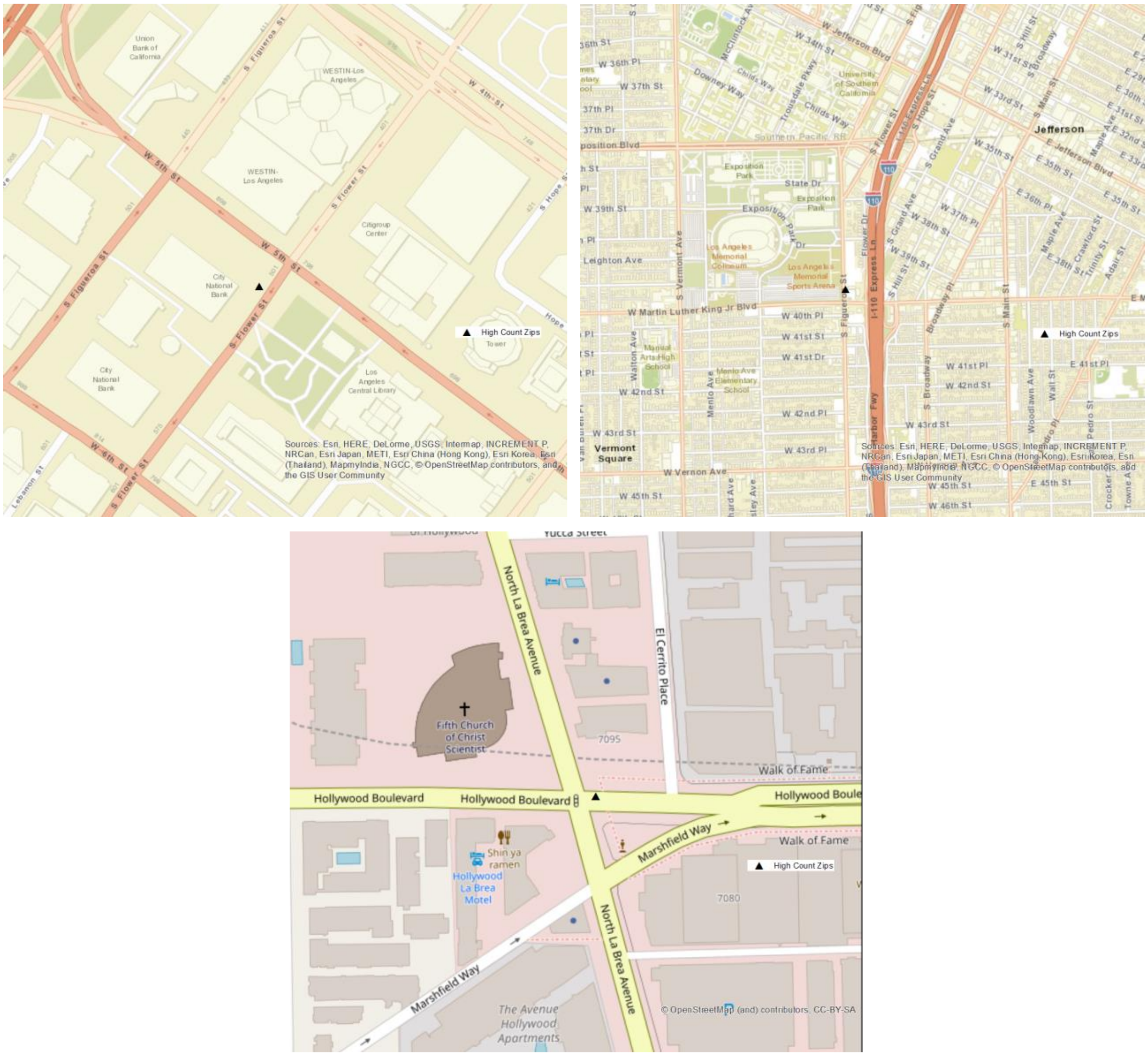

Figure 5. Location of High Count 9-digit Zip Codes near L.A. Metro rail stations (3 figures)
NOTE: Black triangles represent the geographic coordinates of the centroid of the high-count 9-digit zip codes

Robustness Regressions

We test the robustness of our regression models by running the regression described in Equation 1 above on different samples and comparing to the results in the main portion of this report.

The regression described in Equation 1 measures the effect of station openings on mobility, controlling for year fixed effects and station fixed effects. In Robustness Model 1, we remove station fixed effects to understand how coefficients change when we do not account for factors common to station areas that do not change over time. In Robustness Model 2, we account for the fact that the Green and Expo lines opened toward the beginning and end of our available data time period. In this robustness check, we exclude the observations from these two lines to see if this materially changes the analyses results. In Robustness Model 3, we restrict our sample to only stations that opened within the span of our dataset by dropping certain Expo and Gold line stations that opened after 2012. This checks whether our regression results are sensitive to the selection of untreated stations. In Robustness Model 4, we include stations with less than two 9-digit zip codes to see whether that exclusion materially affected regression results. In Robustness Model 5, we include the high-count outlier 9-digit zip codes in non-residential areas to check whether their exclusion materially affected regression results. Appendix Tables 19 and 20 report the results of these robustness checks using the pooled sample and using samples of households within a single income category, respectively.

Table 19. Robustness Regression Results for All Income Groups pooled

Model	Mobility Rate	Number of Observations	R-Squared	Rail Station Opening Effect (β)
Main Model	In-move	1216	0.39	-0.0098***
	Out-move	1216	0.65	-0.0036***
	Net-mobility	1216	0.13	-0.0063
Robustness Model 1	In-move	1216	0.12	0.01***
	Out-move	1216	0.10	0.0106***
	Net-mobility	1216	0.03	-0.0006
Robustness Model 2	In-move	741	0.49	-0.0019
	Out-move	741	0.63	-0.0047***
	Net-mobility	741	0.15	0.0028
Robustness Model 3	In-move	988	0.39	-0.0085***
	Out-move	988	0.73	-0.0005
	Net-mobility	988	0.16	-0.008
Robustness Model 4	In-move	1260	0.24	-0.0066
	Out-move	1261	0.32	-0.0073**
	Net-mobility	1259	0.04	-0.0002
Robustness Model 5	In-move	1216	0.39	-0.0098***
	Out-move	1216	0.65	-0.0036***
	Net-mobility	1216	0.13	-0.0063

NOTE: Statistical significance: *** p>0.01, ** p>.05

Table 20. Robustness Regression Results by Income Group

Mobility Rate	Income Group	Number of Observations	R-Squared	Rail Station Opening Effect (β)
Main Model				
In-move	**Negative Earners**	1180	0.07	0.0022
	Extremely Poor	1215	0.40	-0.0095**
	Very Poor	1215	0.29	-0.0015
	Poor	1215	0.39	-0.003
	Middle Income	1212	0.17	-0.0058
	Upper Middle Income	1214	0.29	-0.0109***
	Upper Income	1173	0.15	-0.0143
Out-move	**Negative Earners**	1188	0.04	0.006
	Extremely Poor	1216	0.43	-0.0009
	Very Poor	1216	0.37	-0.011***
	Poor	1216	0.31	-0.007***
	Middle Income	1214	0.31	0.0024
	Upper Middle Income	1214	0.41	0.0035
	Upper Income	1173	0.18	-0.0073
Net mobility	**Negative Earners**	1173	-0.01	-0.0006
	Extremely Poor	1215	0.07	-0.0087***
	Very Poor	1215	0.10	0.0091***
	Poor	1215	0.09	0.0049
	Middle Income	1211	0.04	-0.007
	Upper Middle Income	1214	0.07	-0.0144***
	Upper Income	1160	0.03	-0.001
Robustness Model 1				
In-Move	**Negative Earners**	1180	0.00	0.0151**
	Extremely Poor	1215	0.15	0.0105***
	Very Poor	1215	0.09	0.0126***
	Poor	1215	0.07	0.0122***
	Middle Income	1212	0.02	0.0108***
	Upper Middle Income	1214	0.00	0.0084**
	Upper Income	1173	0.01	0.002
Out-Move	**Negative Earners**	1188	0.00	0.0104
	Extremely Poor	1216	0.10	0.012***
	Very Poor	1216	0.08	0.009***
	Poor	1216	0.03	0.0121***
	Middle Income	1214	0.01	0.0135***
	Upper Middle Income	1214	0.01	0.0148***
	Upper Income	1173	0.00	0.0084
Net Mobility	**Negative Earners**	1173	-0.01	0.0051
	Extremely Poor	1215	0.03	-0.0015
	Very Poor	1215	0.01	0.0036
	Poor	1215	0.02	0.0001
	Middle Income	1211	0.00	-0.0031
	Upper Middle Income	1214	0.01	-0.0064
	Upper Income	1160	0.02	-0.0059

Mobility Rate	Income Group	Number of Observations	R-Squared	Rail Station Opening Effect (β)
Robustness Model 2				
In-Move	**Negative Earners**	720	0.09	-0.0165
	Extremely Poor	741	0.43	-0.0016
	Very Poor	741	0.26	-0.002
	Poor	741	0.40	-0.0059
	Middle Income	738	0.14	-0.0083
	Upper Middle Income	739	0.30	-0.0068
	Upper Income	704	0.12	0.0013
Out-Move	**Negative Earners**	726	0.04	0.0112
	Extremely Poor	741	0.45	-0.0022
	Very Poor	741	0.36	-0.0131***
	Poor	741	0.30	-0.0156***
	Middle Income	739	0.28	-0.0033
	Upper Middle Income	739	0.35	0.0065
	Upper Income	701	0.16	0.0025
Net Mobility	**Negative Earners**	714	-0.01	-0.018
	Extremely Poor	741	0.10	0.0006
	Very Poor	741	0.07	0.0111
	Poor	741	0.08	0.0097
	Middle Income	737	0.02	-0.0029
	Upper Middle Income	739	0.05	-0.0133
	Upper Income	691	0.01	0.0098
Robustness Model 3				
In-Move	**Negative Earners**	964	0.07	0.0017
	Extremely Poor	987	0.39	-0.0099***
	Very Poor	987	0.32	-0.0001
	Poor	987	0.41	0.0014
	Middle Income	985	0.18	-0.001
	Upper Middle Income	988	0.30	-0.0101
	Upper Income	961	0.15	-0.0165
Out-Move	**Negative Earners**	966	0.07	0.0039
	Extremely Poor	988	0.46	0.0037
	Very Poor	988	0.45	-0.0087***
	Poor	988	0.36	-0.0022
	Middle Income	987	0.36	0.0015
	Upper Middle Income	988	0.45	0.0026
	Upper Income	961	0.17	-0.0101
Net Mobility	**Negative Earners**	958	-0.01	-0.001
	Extremely Poor	987	0.06	-0.0137***
	Very Poor	987	0.13	0.0082***
	Poor	987	0.11	0.0044
	Middle Income	984	0.08	-0.0005
	Upper Middle Income	988	0.07	-0.0127
	Upper Income	952	0.03	-0.0067

Mobility Rate	Income Group	Number of Observations	R-Squared	Rail Station Opening Effect (β)
Robustness Model 4				
In-Move	**Negative Earners**	1185	0.09	0.0031
	Extremely Poor	1245	0.27	-0.0106**
	Very Poor	1239	0.30	-0.0012
	Poor	1239	0.42	-0.0023
	Middle Income	1231	0.17	-0.007
	Upper Middle Income	1240	0.31	-0.0112***
	Upper Income	1189	0.19	-0.0101
Out-Move	**Negative Earners**	1196	0.07	0.0077
	Extremely Poor	1247	0.23	-0.002
	Very Poor	1242	0.37	-0.0127***
	Poor	1244	0.21	-0.0059
	Middle Income	1235	0.29	0.0015
	Upper Middle Income	1242	0.44	0.0024
	Upper Income	1188	0.18	-0.0086
Net Mobility	**Negative Earners**	1177	-0.01	-0.0006
	Extremely Poor	1241	0.02	-0.0083***
	Very Poor	1237	0.11	0.0101***
	Poor	1237	0.14	0.0023
	Middle Income	1227	0.02	-0.0082
	Upper Middle Income	1235	0.19	-0.014***
	Upper Income	1173	0.03	0.0009
Robustness Model 5				
In-Move	**Negative Earners**	1180	0.07	0.0022
	Extremely Poor	1215	0.40	-0.0094**
	Very Poor	1215	0.29	-0.0015
	Poor	1215	0.39	-0.0029
	Middle Income	1212	0.17	-0.0058
	Upper Middle Income	1214	0.29	-0.011***
	Upper Income	1173	0.15	-0.0145
Out-Move	**Negative Earners**	1188	0.04	0.0056
	Extremely Poor	1216	0.43	-0.0008
	Very Poor	1216	0.37	-0.0111***
	Poor	1216	0.31	-0.0069***
	Middle Income	1214	0.31	0.0024
	Upper Middle Income	1214	0.41	0.0034
	Upper Income	1173	0.18	-0.0074
Net Mobility	**Negative Earners**	1173	-0.01	-0.0003
	Extremely Poor	1215	0.06	-0.0086***
	Very Poor	1215	0.10	0.0091***
	Poor	1215	0.09	0.0049
	Middle Income	1211	0.04	-0.0071
	Upper Middle Income	1214	0.06	-0.0144***
	Upper Income	1160	0.03	-0.0012

NOTE: Statistical significance: *** p>0.01, ** p>.05

References

1. Atkinson-Palombo, C. (2010). Comparing the capitalisation benefits of light-rail transit and overlay zoning for single-family houses and condos by neighbourhood type in metropolitan Phoenix, Arizona. Urban Studies, 47(11), 2409-2426.
2. Baker, D. M., & Lee, B. (2017). How Does Light Rail Transit (LRT) Impact Gentrification? Evidence from Fourteen US Urbanized Areas. Journal of Planning Education and Research, 0739456X17713619.
3. Bartholomew, K., & Ewing, R. (2011). Hedonic price effects of pedestrian-and transit-oriented development. Journal of Planning Literature, 26(1), 18-34.
4. Boarnet, M. G., Bostic, R. W., Rodnyansky, S., Prohofsky, A., Eisenlohr, A., and Webb Jamme, H.T. (2018). "The Effects of Rail Stations on Neighborhood Displacement in Los Angeles County, 1993-2013". California Community Foundation.
5. Boarnet, M. G., Bostic, R. W., Rodnyansky, S., Santiago-Bartolomei, R., Leslie, C. Environmentally Sustainable and Affordable Housing Near Transit in Los Angeles. *California Department of Transportation*. March 14, 2015. https://www.metrans.org/sites/default/files/research-project/15-13%20Final%20Report_Bostic.pdf
6. California Franchise Tax Board (FTB). (2006). *Tax Gap Plan: A Strategic Approach to Reducing California's Tax Gap*. Retrieved from https://www.ftb.ca.gov/aboutftb/taxgapstratplan.pdf
7. California Franchise Tax Board (FTB). (2017). *2016 California Tax Rate Schedules*. Retrieved from https://www.ftb.ca.gov/forms/2016-california-tax-rates-and-exemptions.shtml
8. Chapple, Karen. Mapping susceptibility to gentrification: The early warning toolkit. *Berkeley, CA. University of California-Berkeley, Center for Community Innovation*, 2009. Retrieved from http://communityinnovation.berkeley.edu/reports/Gentrification-Report.pdf.
9. Dong, H. (2017). Rail-transit-induced gentrification and the affordability paradox of TOD. Journal of Transport Geography, 63, 1-10.
10. Duncan, M. (2011). The impact of transit-oriented development on housing prices in San Diego, CA. Urban studies, 48(1), 101-127.
11. Golub, A., Guhathakurta, S., & Sollapuram, B. (2012). Spatial and temporal capitalization effects of light rail in Phoenix: From conception, planning, and construction to operation. Journal of Planning Education and Research, 32(4), 415-429.
12. Grube-Cavers, A., & Patterson, Z. (2015). Urban rapid rail transit and gentrification in Canadian urban centres: A survival analysis approach. Urban Studies, 52(1), 178-194.
13. Heilmann, Killian. Transit Access and Neighborhood Segregation. A Study of the Dallas Light Rail System". Working Paper. http://econweb.ucsd.edu/~kheilman/personalResearch.html
14. Hess, D. B., & Almeida, T. M. (2007). Impact of proximity to light rail rapid transit on station-area property values in Buffalo, New York. Urban studies, 44(5-6), 1041-1068.

15. Higgins, C.D., & Kanaroglou, P.S. (2016). Forty years of modelling rapid transit's land value uplift in North America: Moving beyond the tip of the iceberg. Transport Reviews 36(5): 610-634.

16. Immergluck, D. (2009). Large redevelopment initiatives, housing values and gentrification: the case of the Atlanta Beltline. Urban Studies 46(8), 1723-1745.

17. Kahn, M. E. (2007). Gentrification trends in new transit-oriented communities: Evidence from 14 cities that expanded and built rail transit Systems. Real Estate Economics, 35(2), 155-182.

18. Kahn, Matthew E. Gentrification Trends in New Transit-Oriented Communities: Evidence from 14 Cities That Expanded and Built Rail Transit Systems. *Real Estate Economics*, 2007. 35, (2): 155–82.

19. Knaap, G. J., Ding, C., & Hopkins, L. D. (2001). Do plans matter? The effects of light rail plans on land values in station areas. Journal of Planning Education and Research, 21(1), 32-39.

20. Lin, Jeffrey. Gentrification and Transit in Northwest Chicago. *Journal of the Transportation Research Forum*, 2002. 56: 175-191.

21. Los Angeles County and Metropolitan Transportation Authority (L.A. Metro). Long Range Transportation Plan, 2009. http://media.metro.net/projects_studies/images/final-2009-LRTP.pdf

22. Pilgram, C. A., & West, S. E. (2017). Fading premiums: The effect of light rail on residential property values in Minneapolis, Minnesota. Regional Science and Urban Economics.

23. Pollack, S., Bluestone, B., Billingham, C. Maintaining Diversity In America's Transit-Rich Neighborhoods: Tools for Equitable Neighborhood Change. *Dukakis Center for Urban and Regional Policy,* 2010.

24. Redfearn, C. L. (2009). How informative are average effects? Hedonic regression and amenity capitalization in complex urban housing markets. Regional Science and Urban Economics, 39(3), 297-306.

25. Rodnyansky, S. Boarnet, M.G., Bostic, R.W., Santiago-Bartolomei, R., Williams, D. & Prohofsky, A. (2018). New Ways to Measure Residential Mobility: Assessing the Relationship of Rail Transit on Neighborhood Mobility Rates in Los Angeles County in Household Mobility and Neighborhood Impacts (Unpublished Doctoral Dissertation). University of Southern California, Los Angeles, CA.

26. Southern California Association of Governments (SCAG) 2012-2035 Regional Transportation Plan/Sustainable Communities Strategy, 2012. http://rtpscs.scag.ca.gov/Pages/default.aspx

27. Steven Manson, Jonathan Schroeder, David Van Riper, and Steven Ruggles. *IPUMS National Historical Geographic Information System: Version 13.0* [Database] (NHGIS). Minneapolis: University of Minnesota. 2018. http://doi.org/10.18128/D050.V13.0

28. Zuk, M., & Chapple, K. Urban Displacement Project. *Mapping Neighborhood Change in Los Angeles County*, 2015a. http://www.urbandisplacement.org/map/la

29. Zuk, M., Bierbaum, A., Chapple, K., Gorska, K., Loukaitou-Sideris, A., Ong, P., Thomas, T. "Gentrification, Displacement and the Role of Public Investment: A Literature Review". *Working papers of the Federal Reserve Bank of San Francisco*, 2015. http://www.frbsf.org/community-development/publications/working-papers/2015/august/gentrification-displacement-role-of-public-investment/

30. Zuk, M., Bierbaum, A., Chapple, K., Gorska, K., Loukaitou-Sideris. Gentrification, Displacement, and the Role of Public Investment. *Journal of Planning Literature*, 2017. DOI: 10.1177/0885412217716439