Time is Money: A Pseudo-Panel Analysis of The Relationship Between Income and Commute Time

March 2019

Roman Sielewicz

Department of Economics
University of California, Santa Barbara
romansielewicz@ucsb.edu

Advisor: Daniel Cullen, PhD Candidate

Professor: Heather Royer, PhD

Abstract

This paper investigates the relationship between income and commute time, both at the individual and community levels. I use data from Integrated Public Use Microdata Series (IPUMS) sourced from the American Community Survey to show that a 1% increase in average income in a community is associated with a 0.09% increase in average commute time. In addition, I examine data from Los Angeles County to show that between individuals of the same demographic characteristics, 1% higher income is associated with 0.068% longer commute time. Across the United States, I find that commute times have been increasing in the 13 years sampled, but the increase is consistent across all races, education levels, and income deciles. Though increases in income are associated with increases in commute time for individual communities, I find that between communities within Los Angeles, a 1% difference in income is associated with a 0.828% *shorter* commute time, meaning that residents of poor communities in Los Angeles take longer to get to work than their wealthy counterparts.

Introduction

The commute is universal. Despite the advent of new technologies allowing us to be productive from nearly anywhere, the vast majority of humans travel to their workplace each day. The commute serves as a divide between home and work, and a barrier between personal and professional life.

Although there are a large number of factors that determine an individual's commute time, and an equally large number of variables that decide commute times for a given community, I investigate the relationship between income and commute time, both at the individual and community levels. It is important to note that commute time is not necessarily a source of disutility. In fact, it has been found that people have an ideal commute time, and the utility of the commute depends on how far travel times to work vary from this ideal time. However, achieving an ideal commute time is not possible for all members of society. According to a sample of San Francisco Bay Area residents in 2001 (Redmond and Mokhtarian), most surveyed have a longer commute time than they would like.

In a 2015 study from Harvard University (Chetty and Hendren), commute time was listed as the single biggest factor in an individual's chances of escaping poverty. Commute time, therefore, can serve as a major barrier to success for those less fortunate, and a major inconvenience for the rest of us. In this paper, I will address the following question: What is the relationship between income and commute time, both at the individual and the community levels? Ultimately, I find that average income and average commute time are positively related for communities and individuals, but within cities, specifically Los Angeles, low income and minority communities have higher average commute times than their wealthier counterparts.

Data

In this paper I use data from IPUMS: Integrated Public Use Microdata Series. Individual samples are collected by the American Community Survey from 2005 to 2017 and include observations of over 40 million individuals in the United States during that time period. Only employed individuals with a positive yearly income are included in my analysis, bringing down the total observations to just over 18 million individuals. The main focus of this paper is the relationship between income, and commute time.

Each individual in the sample has various demographic characteristics, including race, sex, education level, and means of transportation to work. Importantly for this analysis, the sample year and Public Use Microdata Area (PUMA) in which a person resides are also included.

To construct the pseudo-panel of PUMAs, I divide the sample by PUMA (of which there are 1,078) and sample year. Each observation in the pseudo-panel includes the average commute time, average income, proportion of White residents (as a percentage), proportion of residents who have obtained a bachelor's degree or higher (as a percentage), and the PUMA's Gini Coefficient. The proportion of White residents ranges from 1% to 99%, with a mean of 78%, while the proportion of residents with a bachelor's degree or above ranges from 3.5% to 91%, with a mean of 33%. Gini Coefficient in all PUMAs in all years ranges from 0.33 to 0.66, with a mean of 0.45.

Both income and commute time vary widely across individuals in the sample, but commute times vary greatly for individuals earning below the median yearly income of $38,000, while very few high-income earners have extremely high commute times. Individuals in the sample have an average yearly income of $54,000 and an average commute time of 25 minutes. An important statistic to note is that Americans outside of metropolitan areas have an average commute time of just 13.5 minutes, meaning that commute time depends largely on an individual's location

Methodology

In the absence of panel data, to analyze the effect of income on commute time I construct a pseudo-panel. Observations of individuals are grouped by the Public Use Microdata Area or *PUMA*[1] to which they belong and the year in which the sample was taken. To assess the effects of a variety of factors, the average income, average commute time, and Gini Coefficient of each PUMA are calculated along with the percentage of that area that identifies as White, and the percentage of that area that has obtained a bachelor's degree or higher. Then, to control for unobserved factors that differ depending on geography but remain consistent in a single PUMA throughout the sampling period, I add fixed effects for each PUMA, and additional fixed effects for each year to control for shocks that happen in specific years, such as the great recession in 2008 and 2009.

The analysis will revolve around the following model and variations of it.

$$log(T_{it}) = \alpha_i + \gamma_t + \beta_1 log(I_{it}) + \beta_2 W_{it} + \beta_3 C_{it} + \beta_4 W_{it} C_{it} + u_{it}$$

Where each symbol or letter signifies a variable defined by the following:

- T_{it} : Average Commute Time of PUMA i in year t
- α_i : Fixed Effect of Public Use Microdata Area i
- γ_t : Fixed Effect of year t
- I_{it} : Average income of PUMA i in year t

[1] In reality, the geographic delineations for each PUMA change with each new census, so another variable, *CPUMA* or Consistent PUMA is used for statistical analysis because its delineations remain consistent from 2000-2017. 2010 PUMA delineations are used for maps in this paper.

- W_{it} : Percentage of PUMA i that identifies as White in year t
- C_{it} : Percentage of PUMA i that has a College degree or higher in year t
- u_{it} : Error Term of PUMA i in year t

In addition to the PUMA and year fixed-effects models used to analyze the country's commute time trends as a whole, this paper will also employ a model to investigate individual cities and examine the relationship between commute times, income, and other demographic factors for individuals within Los Angeles, California.

$$log(T_i) = \alpha_i + \gamma_t + \beta_1 log(I_i) + \beta_2 R_i + \beta_3 E_i + \beta_4 R_i E_i + u_i$$

The specifications are similar to the previous fixed effects model:

- T_i : Commute Time of person i
- α_i : Intercept for PUMA of person i
- γ_t : Intercept for Year t
- I_i : Income of person i
- R_i : Race of person i
- E_i : Education Level of person i

Ideally, there would exist panel data on commute times for individuals, but due to limitations in the data, constructing a pseudo-panel and using cross sectional data at a city level is the best way to analyze the relationship between income and commute times.

Results

Commute Times have been increasing across the United States. Examining five of the country's Largest Metropolitan areas it is easy to see steady increases in average commute time across all five, but this trend is not limited to the cities shown, and in fact holds for each major city in the United States. Although commute time in a given area and population are undoubtedly related, the fixed effects model used to analyze commute time increases across the U.S. will account for the changes brought about by time period (through the year fixed effects variables) and assesses the parameters for each Public Use Microdata area individually.

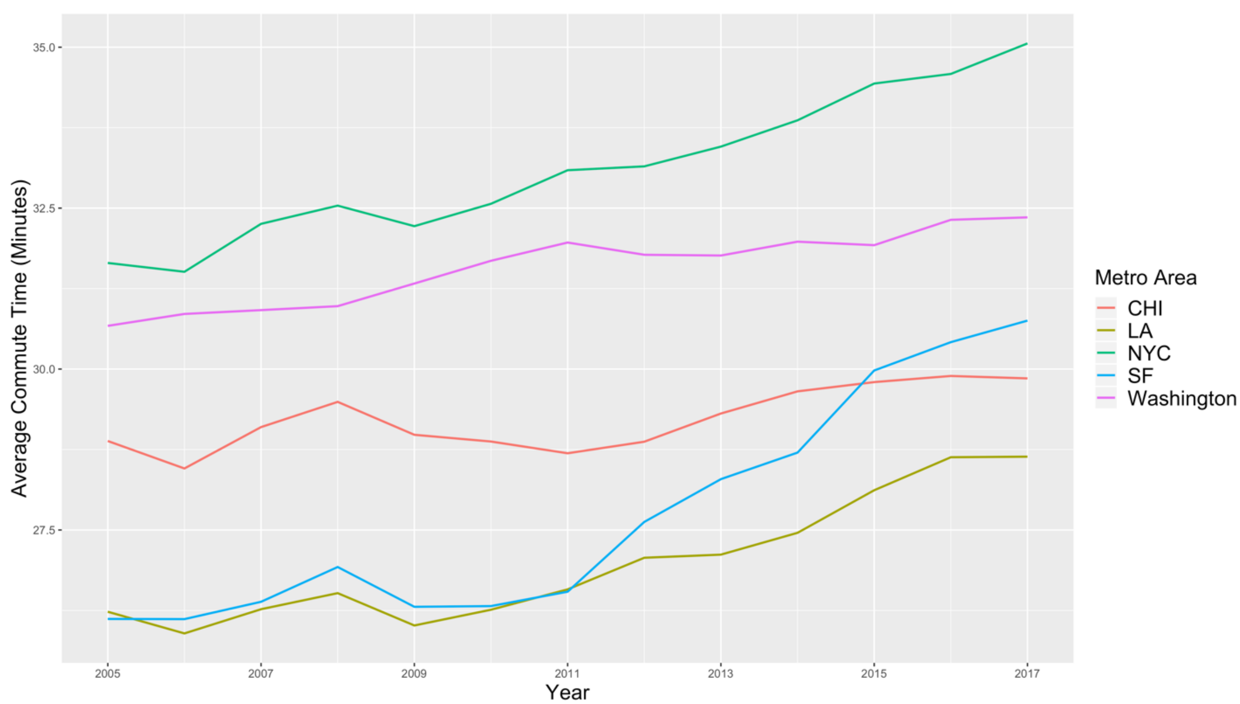

Figure 1: Average commute time by year in a selection of cities

Although commute times are increasing for every income decile, there does not appear to be any systematic difference in which income decile has experienced the greatest increase in commute time (Figure 2). Here we can see nearly identical trends in commute time for every income decile, with the first and second decile of income earners in America in each year experiencing nearly identical commute times on average. The main trend to notice among income deciles is that invariably, individuals in higher income deciles have longer commutes on average. This could be due to a variety of factors, including greater wealth concentration in cities, whose residents have an average commute time of 25 minutes, compared to 13.5 minutes for non-metropolitan areas.

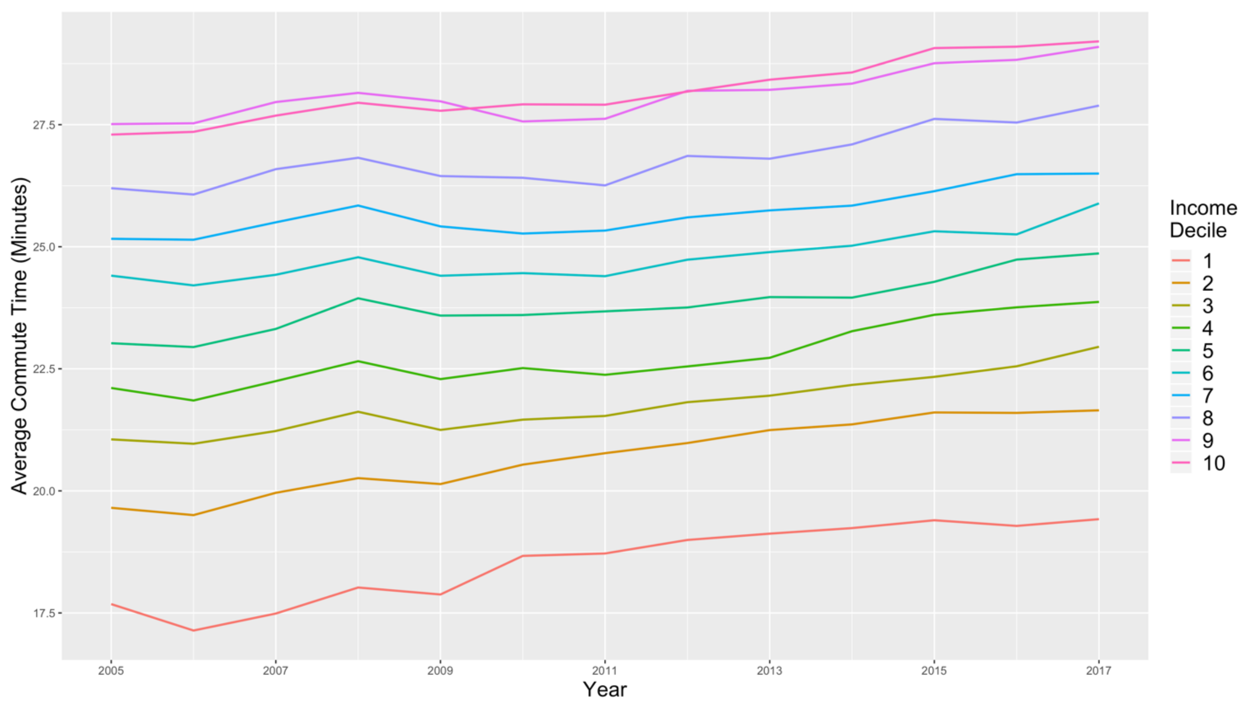

Figure 2: Average commute times of each income decile from 2005-2017

Among major racial groups in the United States, the story is much the same, with no one race increasing its average commute time by more than any other (Figure 3). However, there are clear differences in overall average commute times for each race. Most notably, White Americans have the second lowest average commute time of any race. This fact can likely be attributed to more whites living in rural or suburban areas where commute times are low.

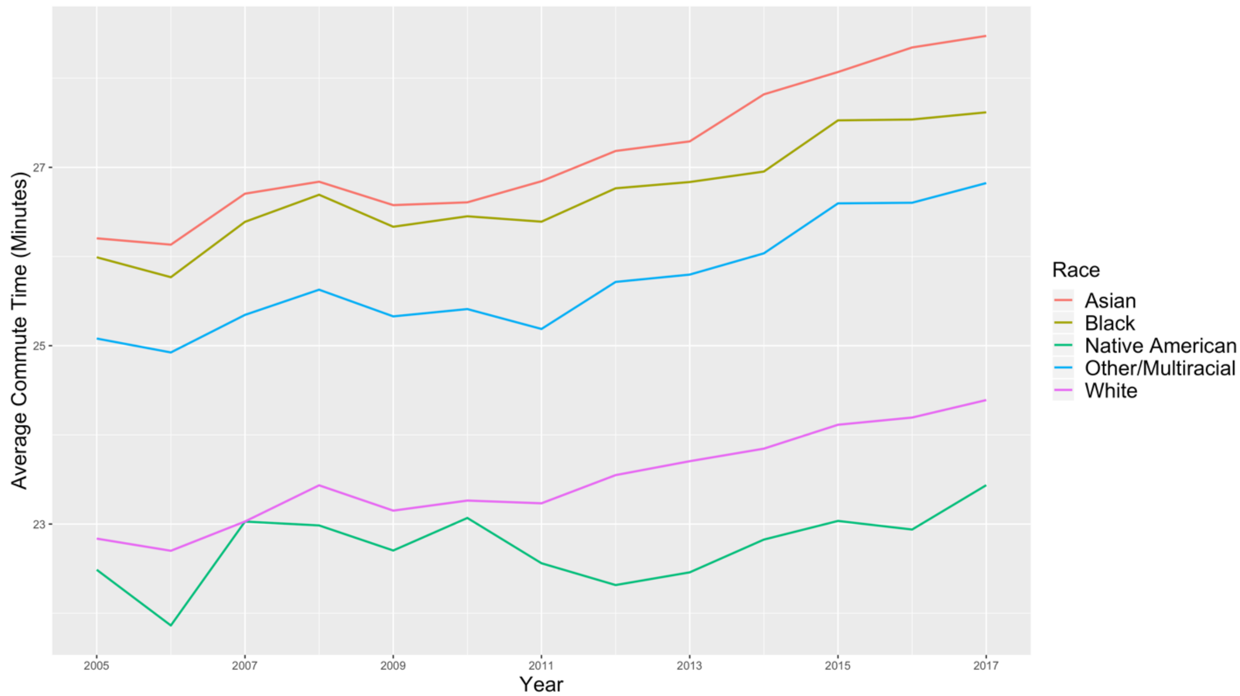

Figure 3: Average Commute time of major racial groups from 2005-2017

When breaking down average commute times by mode of transport to work, Americans who take public transport have the longest commutes at approximately 52 minutes, with Motor Vehicle drivers taking an average of 27 minutes to get to work (both as of 2017). Importantly, since 2005, average commute times have increased by about 4 minutes for all modes of transport to work and will continue to do so if these trends are maintained.

Table 1: US PUMAs OLS and Fixed Effects Regressions

	Dependent variable:		
	log(Commute Time)		
	(1)	(2)	(Fixed Effects)
log(Avg Income)	0.197***	0.424***	0.090***
	(0.006)	(0.011)	(0.009)
% White		−0.005***	−0.001**
		(0.0002)	(0.0002)
% College		−0.004***	0.002***
		(0.001)	(0.0004)
% White · % College		−0.00001	−0.00003***
		(0.00001)	(0.00001)
Observations	14,014	14,014	14,014
R^2	0.070	0.306	1.000
Note:			*p<0.1; **p<0.05; ***p<0.01

Figure 4: Regression with PUMA and year fixed effects. Bolded Coefficients are significant at the 1% level

The initial regressions reveal a strong positive association between income and commute time, even when tracking individual PUMAs (which explains the unusually high R-Squared) and controlling for yearly variation (Figure 4). Data suggests that as a community gets wealthier, its average commute time in turn increases.

The variables "% College" and "% White" represent the percentage of a PUMA that has a bachelor's degree or above and the percentage of a PUMA that identifies as White respectively. When these two variables are added in Regression 2 along with an interaction term, the coefficient on average income increases, suggesting that the result from Regression 1 suffers from omitted variable bias. With an R-Squared of 0.306, Regression 2 already fits the data quite well, as seen in the scatterplot below (Figure 5).

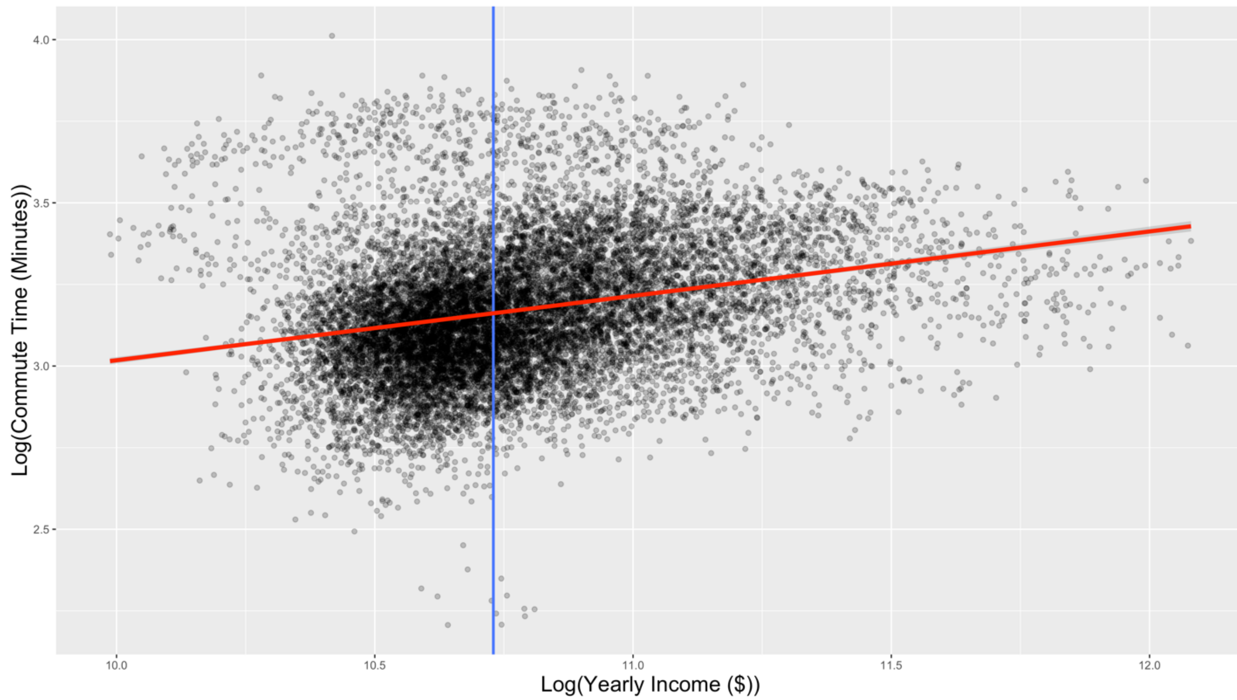

Figure 5: Log(Avg Commute Time) plotted against Log(Avg Income) at the PUMA level. Red Line is OLS regression with controls for education and racial composition

The Fixed Effects regression includes both PUMA and year fixed effects, controlling for both the consistent differences between regions and communities, and the variation due to yearly shocks. Using this model, it is possible to follow each Public Use Mircodata Area from 2005-2017 and observe the changes associated with variation in income, % White, and % College educated (Figure 6)

The data shows that a 1% increase in average income in a community is associated with a 0.09% increase in average commute time for that community. This means that if a community were to hypothetically double its average income, its commute time would increase by 9%. Because the model accounts for the effects of individual years in the regression, changes in population are controlled for as population varies year to year in each Public Use Microdata Area.

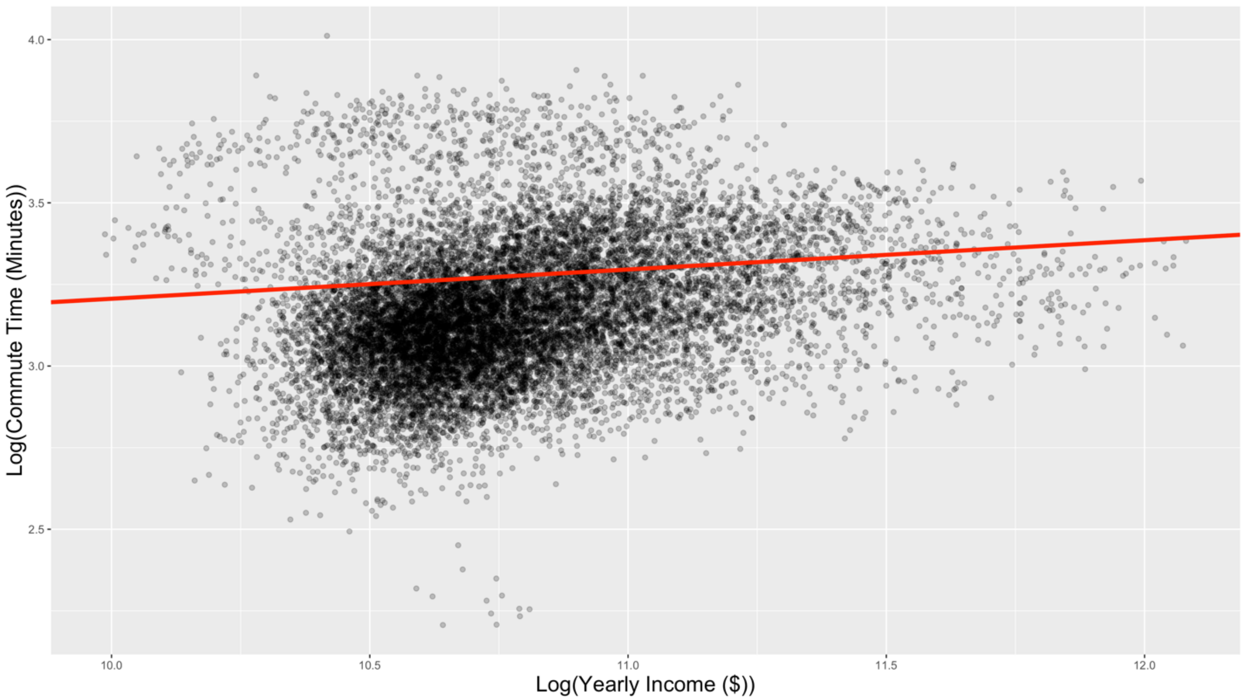

Figure 6: Scatterplot of Log(Avg Commute Time) Against Log(Avg Income)

Red line represents fixed-effect regression with average year fixed-effect as intercept.

Interestingly, the Gini coefficient, a measure of inequality, also seems to have an effect on a community's average commute time, though this effect becomes statistically insignificant when interacted with Log(Avg. income) (Figure 7). The reason for this negative association likely goes back to the commute times of each income decile. A community with a high Gini coefficient will have a few very wealthy individuals who tend to have long commutes, and a large chunk of low-income individuals who work locally and have shorter commutes. Since the Gini coefficient can only vary from 0 (perfect equality) to 1 (one person earning all the income), the Gini Coefficient of a community has little effect on average commute time in a community. This becomes even more apparent when accounting for the fact that Gini coefficients only range from 0.33 to 0.63 among all the PUMAs surveyed. Importantly, the association between income and commute time does not vary greatly even when adding the Gini coefficient to the regression.

Table 2: US PUMAs Fixed Effect Regression with Gini Coeffcient

	Dependent variable:		
	log(Avg Commute Time)		
	(1)	(2)	(3)
log(Avg Income)	0.090***	0.132***	0.123***
	(0.009)	(0.009)	(0.030)
% White	−0.001**	−0.0005**	−0.0005**
	(0.0002)	(0.0002)	(0.0002)
% College	0.002***	0.002***	0.002***
	(0.0004)	(0.0004)	(0.0004)
% White · % College	−0.00003***	−0.00003***	−0.00003***
	(0.00001)	(0.00001)	(0.00001)
Gini		−0.280***	−0.486
		(0.025)	(0.644)
log(Avg Income)· Gini			0.019
			(0.060)
Observations	14,014	14,014	14,014
R^2	0.99	0.99	0.99
Note:			*p<0.1; **p<0.05; ***p<0.01

Figure 7: OLS Regressions using PUMA and Year Fixed Effects

Los Angeles is well known throughout the U.S. for its traffic issues and brutal commutes, which increased by 9.2% on average from 2005-2017. Using cross sectional data for more than 500,000 individuals in LA from 2005-2017, I regress income on commute time, adding controls for year, Public Use Microdata area and mode of transport as well as interaction terms for education and race (Figure 8).

Table 3: LA Cross Sectional Regression

	Dependent variable:
	log(Commute Time)
log(Income)	0.068***
	(0.001)
Female	−0.072***
	(0.019)
Hispanic	0.057***
	(0.003)
Asian	0.153***
	(0.030)
Black	0.247***
	(0.045)
Native American	0.135
	(0.137)
Other/Multiracial	0.072**
	(0.029)
Graduate	−0.035***
	(0.005)
High School	−0.060***
	(0.004)
Less Than High School	−0.047***
	(0.005)
Observations	523,784
R^2	0.952
Note:	*p<0.1; **p<0.05; ***p<0.01

Figure 8: Cross Sectional regression of commute time on income for Los Angeles residents

PUMA and year dummies are also included in the regression, but not listed.

Much like in the pseudo-panel constructed for the United States, the data shows a positive relationship between income and commute time. An interesting result to note is that female Angelenos have on average a 0.07% shorter commute than their male counterparts of the same race and education level. In addition, there are statistically significant differences between commute times by racial group, with White individuals having the lowest commute times compared to individuals of other races who share the same demographic characteristics. However, the biggest difference is between White and Black individuals, with Black individuals having a commute time of just 0.25% longer than Whites, implying that an individual's income, location, and mode of transport to work are the biggest factors in determining their commute time.

Table 4: LA PUMA and Year Fixed Effects Regression

	Dependent variable:
	log(Avg Commute Time)
log(Avg Income)	0.028
	(0.045)
% White	0.001
	(0.001)
%College	−0.001
	(0.002)
% White % College	−0.00001
	(0.00003)
Observations	377
R^2	0.99
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Figure 9: OLS Analysis of Los Angeles County using PUMA and Year Fixed Effects

Analyzing Los Angeles County using the same PUMA and Year fixed effects model used for the United States, we see that change in the average income of an area in LA is not a strong predictor of its average commute time (Figure 9). In fact, none of the previous regressors are statistically significant, so average commute time for each PUMA in LA county is likely determined mostly by unobserved factors in that area, such as proximity to freeways or number of jobs in the local vicinity.

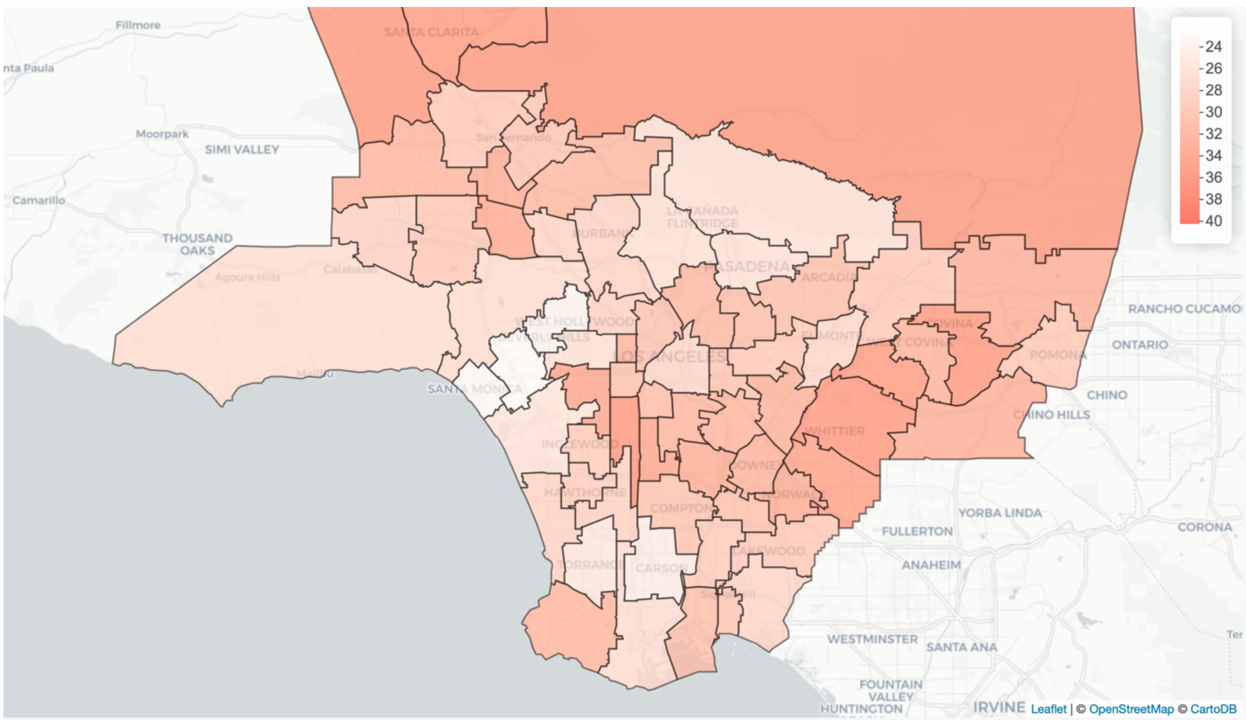

Figure 10: Average Commute Times in Los Angeles County by Public Use Microdata Area as of 2017

Darker shading signifies longer average commute time

Examining a map of Los Angeles County showing Average Commute times in each PUMA, there is a noticeable hotspot of low commute times from Santa Monica to West Hollywood (Figure 10). Low average commute times in these areas suggests that residents work nearby and do not need to travel far to reach their place of work.

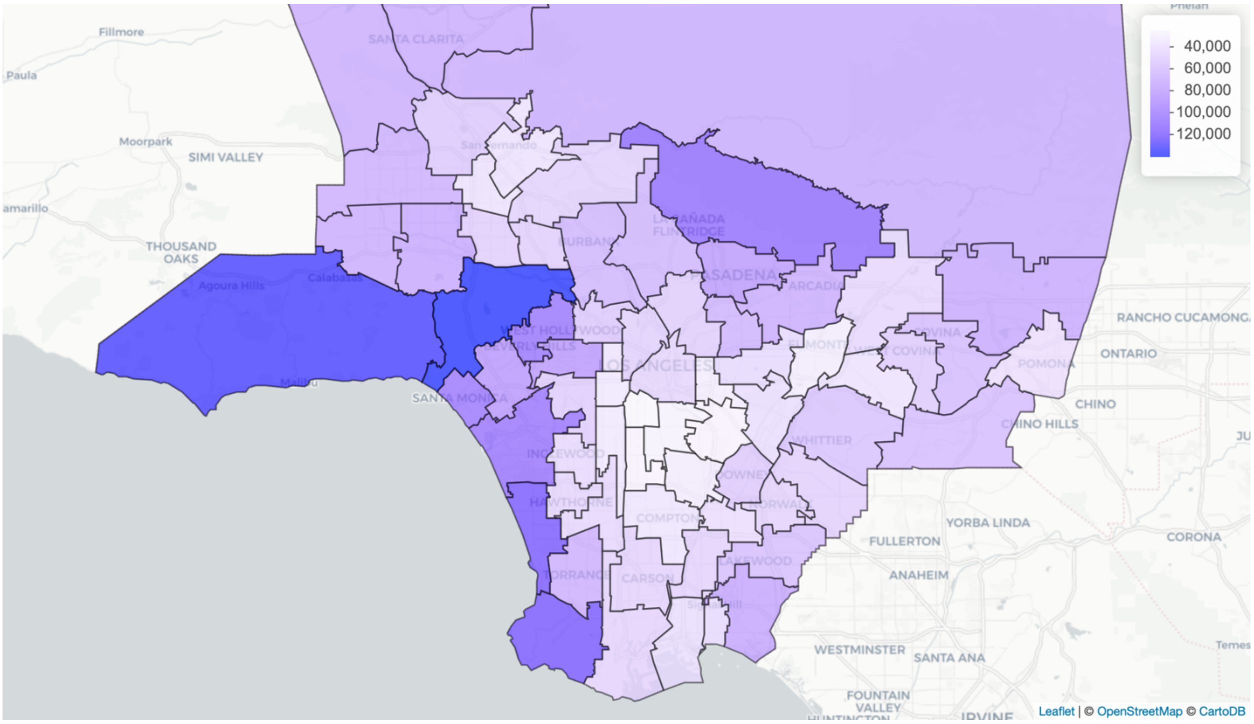

Figure 11: Map of Los Angeles County PUMAS by Average Yearly Income as of 2017

Darker Shading Signifies Higher Average Income

Comparing a map of Los Angeles County PUMAs by Average Yearly Income, many of the areas that have longer average commutes are also the areas with the lowest average incomes (Figure 11). This suggests the opposite of the findings from both the cross-sectional analysis of Los Angeles and panel analysis of United States PUMAs. However, this relationship does not tell the whole story, as the racial composition and education level of the PUMAs also play a role in determining the average commute time. In fact, adding covariates for racial composition and average education level, there is a large and statistically significant negative association between average income and average commute time in LA PUMAs (Figure 12). This suggests that although increase in income is associated with increase in commute time for similar individuals, areas with lower average income in Los Angeles tend to have higher average commute time and from the fixed effects model, (Figure 9) increases in average income in an LA PUMA have no statistically significant effect on its average commute time.

This important finding suggests that low-income communities in Los Angeles have systematic disadvantages in getting to work, and their residents may be geographically far removed from their job or forced to live far from work due to a variety of factors.

Table 5: LA PUMAs Regression

	Dependent variable:		
	log(avg Commute Time)		
	(1)	(2)	(3)
log(avg Income)	−0.078***	0.013	−0.828***
	(0.013)	(0.032)	(0.077)
% White		0.002***	−0.176***
		(0.001)	(0.015)
% College		0.003**	0.022***
		(0.001)	(0.008)
% White · % College		−0.0001***	−0.0004***
		(0.00002)	(0.00004)
log(avg Income) · % White			0.018***
			(0.002)
log(avg Income)· % College			−0.0002
			(0.001)
Observations	377	377	377
R^2	0.089	0.179	0.403
Note:			*p<0.1; **p<0.05; ***p<0.01

Figure 12: Regression of Average Commute time on Average Income for LA PUMAs

Conclusion

Across the United States, the main finding is that a 1% increase in a community's average income is associated with a 0.09% increase in its average commute time. From cross-sectional data in Los Angeles with controls for race, education level, year and place of residence, results are similar, showing that an individual with 1% higher yearly income commutes 0.068% longer on average than someone of the same race, education level and neighborhood. Though income and commute time are positively related in the U.S. as a whole, this is due in large part to low commute times and incomes outside of urban areas. In cities, this trend is reversed, and the results show that within Los Angeles, average commute times for a community *decrease* by 0.83% with a 1% increase in average income, suggesting there are structural factors at play preventing residents of low-income and minority areas from getting to work quickly.

References

Brownstone, David, et al. “Drivers' Willingness-to-Pay to Reduce Travel Time: Evidence from the San Diego I-15 Congestion Pricing Project.” Transportation Research Part A: Policy and Practice, Pergamon, 31 Dec. 2002, www.sciencedirect.com/science/article/pii/S0965856402000216.

Chetty, Raj, and Nathaniel Hendren. “The Impacts of Neighborhoods on Intergenerational Mobility.” *Harvard University*, April 2015.

Redmond, Lothlorien S., and Patricia L. Mokhtarian. The Positive Utility of the Commute: Modeling Ideal Commute Time and Relative Desired Commute Amount. Kluwer Academic Publishers, 2001, https://link.springer.com/article/10.1023/A:1010366321778.

Severen, Christopher. “Commuting, Labor, and Housing Market Effects of Mass Transportation: Welfare and Identification.” Christopher Severen, PhD, Oct. 2018, cseveren.github.io/files/Severen_JMP_2018.pdf.

Wachs, Martin, et al. “The Changing Commute: A Case Study of the Jobs/Housing Relationship over Time.” EScholarship, University of California, Apr. 1993, escholarship.org/uc/item/7424635r.