The Demand for Dollar Vans in New York City

Max Feinglass, Jiheng Huang, Jeremy Neiman, and Vipassana Vijayarangan

# Introduction

New York City has one of the most extensively used transit systems in the country. Over 30% of workers in the New York metro area take transit. Despite the substantial use of transit in NYC, the city has seen alternate forms of transportation gaining ground with the residents of the city. One that is frequently in the media is Uber, however, another mode has existed for decades. These are the commuter vans of New York. Informally called dollar vans or jitneys, they gained prominence during the transit strike of 1980. Due to the unavailability of transit, people took to biking and walking to work. And it was during this strike that the vans played a major role in getting people around and they have stayed relevant ever since.

The vans operated without any form of regulation until 1993, when they first came under the City Council. In practice the council rarely approved requests for licenses and it wasn’t until the Giuliani administration assigned the Taxi and Limousine Commission with the oversight of the vans, did regulation have a meaningful role to play in their operation.

Under existing New York City policy, the Taxi and Limousine Commission (TLC) and the Department of Transportation (NYDOT) both actively participate in the dollar van permitting process. Alexander Keating, the Director of the Special Projects Unit at NYDOT, stressed that NYDOT places significant weight on what regions a proposed commuter van intends to serve when considering granting a permit to an applicant. It is therefore of significance for both regulators and van operators to understand existing demand profiles for dollar van services. DOT and TLC use domain knowledge of the industry as well as rational decision making to approve or deny commuter van permits but are operating largely in the dark. A more data driven understanding of who are using existing dollar vans and why they are riding in them would inform the current regulatory effort and provide drivers a more nuanced understanding of their business.

This study took a host of publically available demographic data from the American Community Survey, economic and spatial records from the LEHD Origin-Destination Employment Statistics (LODES) database, as well as custom calculated commute times to search for correlations between New York City Census Tracts and known van routes.

The study then used a Support Vector Machines (SVM) machine learning algorithm to construct a predictive model that identified census tracts in New York that shared data profiles similar to those found in areas with existing dollar vans. This information allows regulators to assess the validity of new proposed routes and provides dollar van drivers with insight into potential expansion plans.

# Literature

The jitneys, when regulated, often operate differently from their informal counterparts. In New York City, regulation has increasingly favored larger van operators with demonstrated service quality. Regulation has also not addressed the burgeoning number of unlicensed vans. In other parts of the country, regulation has harmed the functioning of the jitney services though it is unclear whether the population they served was integrated into formal transit.

In Miami, the Metro-Dade Transit Agency's bus restructuring plan to introduce jitneys into the formal transit environment was met with wide spread opposition as passengers would have to have paid two fares when transferring between the buses and jitneys. This led to the agency reintroducing buses on lanes that were initially operated by buses, placing jitneys in direct competition with subsidized buses. Similarly, in Los Angeles, while Southern California Rapid Transit District was developing a roadmap to integrate jitneys into transit the California Supreme court instituted a 50c sales tax that was dedicated to transit funding. This allowed transit services to reduce fares and increase frequencies. The jitneys reduced their fares in response to the reduced transit fares, however without the transit fares this proved to be economically unsustainable and the eventually ceased operation.

The Taxi & Limousine Commission announced the Group Ride Vehicle Pilot Project, which intended to replace a defunct bus line with commuter vans. The advantage of the pilot was that it would not affect all commuter van operators and simultaneously provide a controlled environment to test many of the hypothesis about the operation of dollar vans. The pilot was not successful and six months after its initiation, all five of the van operators involved discontinued their service. The failure is attributed to a lag between the suspension of bus services and the introduction of commuter vans. The target passenger for the commuter vans was also unfamiliar with the service and hence apprehensive to try it. (King and Goldwyn, 2014).

The existing research highlights the assumptions that have been made about both commuter vans and the community they serve and this project aims to evaluate the validity of those assumptions.

# Data Description

## LEHD Origin-Destination Employment Statistics (LODES)

The LODES data provides census block level information about employee flows between census blocks. It gives a count of how many people commute from each block to each other. It also provides details on the characteristics of the workers, such as age and income. Our data is separated into information about the tracts subjects reside in and subjects they work in. The particular file used is the 2013 Origin-Destination for New York State for all jobs.

<http://lehd.ces.census.gov/data/lodes/LODES7/ny/od/ny_od_main_JT00_2013.csv.gz>

## American Community Survey (ACS) Demographic Data

The American Community Survey provides demographic data on census block level. This data provides information on various aspects such as age, population ethnicity, income and number of foreign born population in each census block in New York City. The particular file used is Selected Characteristics of the Total and Native Populations in the United States 2010-2014 American Community Survey 5-Year Estimates.

<http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_14_5YR_S0601&prodType=table>

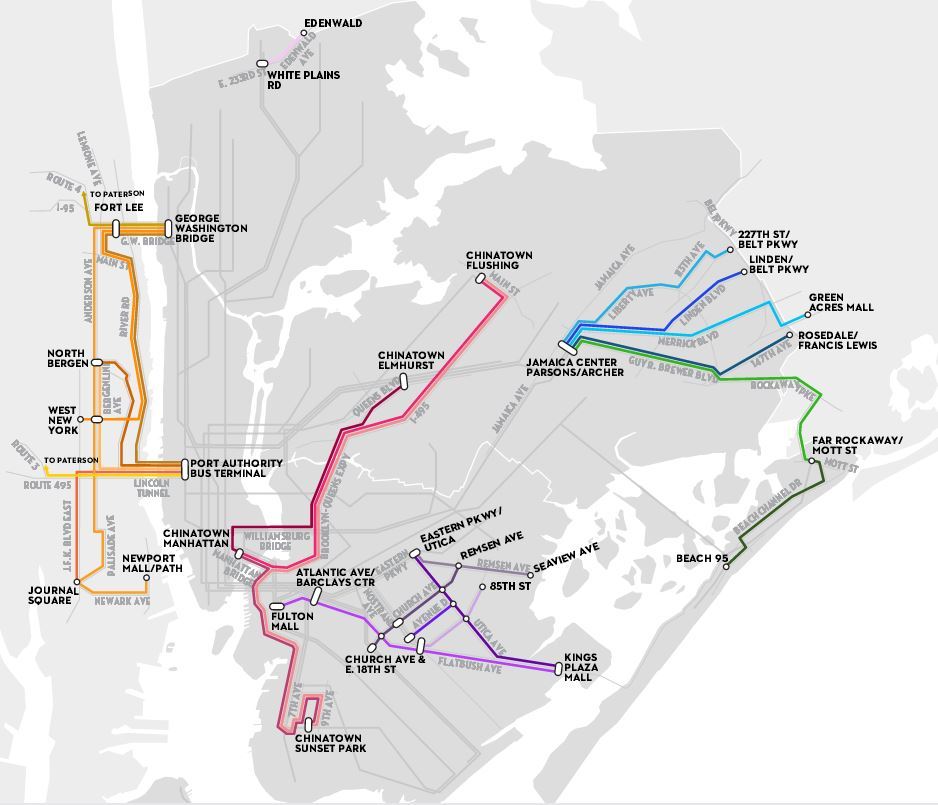


Figure 1: Aaron Reiss' map of dollar van routes

## Current Dollar Van Routes

The data on dollar van routes was compiled by Aaron Reiss in 2014 by riding the vans and documenting their routes. The authors of this report would like to extend our thanks to Aaron Reiss for consenting to be interviewed to provide background for this project. His help proved invaluable.

## Dollar Van Ride

On Feb 25, 2016 Jeremy Neiman and Jiheng Huang took a dollar van ride to Sunset Park, Brooklyn from Chinatown, Manhattan to get firsthand experience with dollar vans. From interviews with passengers during the ride, they found that one of the reasons that passengers choose dollar van is to save time as it is quicker to get to home in Brooklyn than taking the subway during off-peak hours. The dollar van followed its designated route and drops people along 7th Avenue and 9th Avenue in the Sunset Park neighborhood of Brooklyn. The driver told Jiheng, in Chinese, that it is not allowed to drive on 8th Avenue due to high traffic volume on that street’s commercial district. The dollar van driver also talked about current city regulation regarding safety. The company he is working for has a fleet that is examined regularly by NYDOT. The City routinely checks the number of seats installed and the van’s insurance status.

Figure : The backseat of a dollar van

## Geographic Information Systems Analysis

Aaron Reiss provided this project with detailed GPS coordinates of routes of various dollar vans throughout the Tri-State Area. These routes were personally ridden by Mr. Reiss and represent some of the best data available in a data poor domain. Its important to note that Mr. Reiss’s lines do no represent all the dollar vans in New York City. By their very nature, dollar vans are informal and their existence is spread only via word of mouth. There is strong anecdotal evidence that several of New York’s more insulated communities, Orthodox Jews for example, use dollar vans that were not captured in this study.

Of the lines that Mr. Reiss provided, only those completely within the borders of New York City were used. This excludes a large number of lines that originate in New Jersey and on suburban Long Island. Three major routes were examined. The first consisted of a route connecting Chinatowns in Elmwood, Queens, Flushing, Queens, and Sunset Park, Brooklyn with Chinatown in Manhattan. Another connected Downtown Brooklyn with predominantly Caribbean neighborhoods in South Brooklyn along Flatbush Avenue. Finally, a network of van lines that connect far Eastern Queens and the Rockaways with Jamaica Station, a transit hub in Queens, were examined.

For each line, origins and destinations were identified. Destinations were determined as places where people reported wanting to get off vans that were not their homes. These places included transfer points to more formal transportation, such as Jamaica Station, or commercial centers like Downtown Brooklyn. Chinatowns were considered both origins and destinations. Destinations were identified as points surrounded by a half-mile buffer. All census tracts contained within this buffer were flagged as a destination tract. Similarly, origins consisted of streets where dollar vans could reasonably pick up and/or drop off passengers. For the purposes of this study these areas were considered to be all areas within a half mile of the dollar van route that was not a grade separated expressway. All tracts within this buffer were flagged as origin tracts.

These areas were then overlaid on top of 2014 census tracts for New York City. The results are census tracts that are endowed with the properties of either being a dollar van origin, a dollar van destination, both, or none. The resulting map of dollar van origins and destinations shows the areas of heavy dollar van use in New York.

## Commute Time Data

Commute time data was created by using Open Trip Planner, New York City transportation network data and LODES data. Open Trip Planner is an open source project which can calculate travel times between two points for a variety of modes. It was configured with GTFS data for New York City buses, subways, railroad and ferries as well as road network data from Open Street Maps. Once configured, the public transit commute time between the center of each census tract in New York City was calculated.

Next, LODES data was used to calculate a single weighted average commute time for each census tract, X, by multiplying the commute time between X and every other census tract, Y, by the number of commuters who commute between X and Y, then divided by the total number of workers who live in X. This metric takes into account the level of transit service as well as the proximity to jobs.

# Methodology

## Support Vector Machines Feature Selection

A machine learning technique know as Support Vector Machines (SVM) was used to identify the most salient demographic variables that correlate with the location of dollar van origins and destinations. In multidimensional space, SVM creates a boundary between data points. If a data point falls on the wrong side a boundary, a loss is imposed proportional to the distance between that miscategorized point and the boundary. Minimizing this loss yields an optimal boundary line. Here, our boundary line delimits a tract containing a dollar van from a tract without a dollar van and our data points are all our demographic data. The minimization process yields what data points contribute most to this boundary. These data points can be taken to be the defining characteristics, amongst the variables considered, of the neighborhoods that support dollar vans.

Table 1: Most Salient Features of dollar van origins

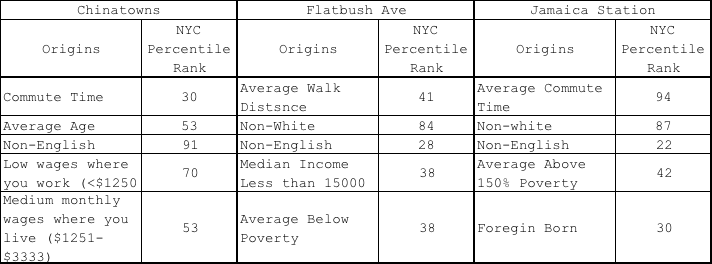
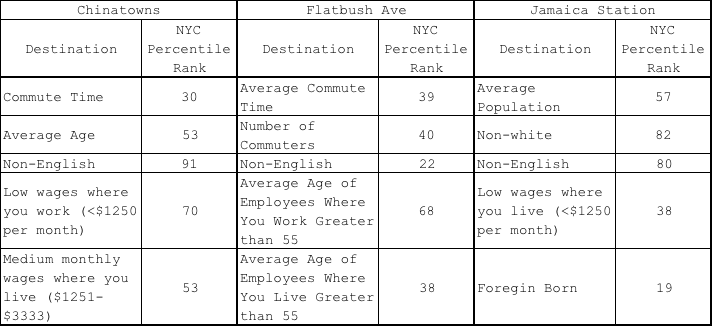


Table 2: Most salient features of dollar van destinations

****

The two tables above show the variables that most correlate to a census tract containing that dollar van origin or destination. The ‘NYC Percentile Rank’ is the relative value of that variable ranked against others in New York City. For example, based on the variables included in our study, the Jamaica Station destination tracts are most characterized by average populations, very high nonwhite populations, very high non-English speaking populations, relatively high wages and very low foreign born tracts.

It is worthy of note that the Chinatown route did not have distinct origins and destinations. Each tract was both an origin and destination. Therefore, their origin/destination features are identical.

## City-wide Predictive Model

A predictive model using support vector machines was used to learn how to classify dollar van origins and destinations. The algorithm was provided tract level commute time, demographic data, and jobs data, as inputs. It output a flag indicating whether the tract was either an origin or destination for dollar vans. We divided the 2160 census tracts in New York City into a 66% training and 33% test set.

For origins, the model correctly classified 94% of actual dollar van origins in the training set and 80% in the test set. For places that are not actually dollar van origins, the model identified 20% of the training set and 24% of the test set, which we could interpret as places where there could potentially be demand for dollar vans. The map below shows the results when the model is run on all the census tracts.

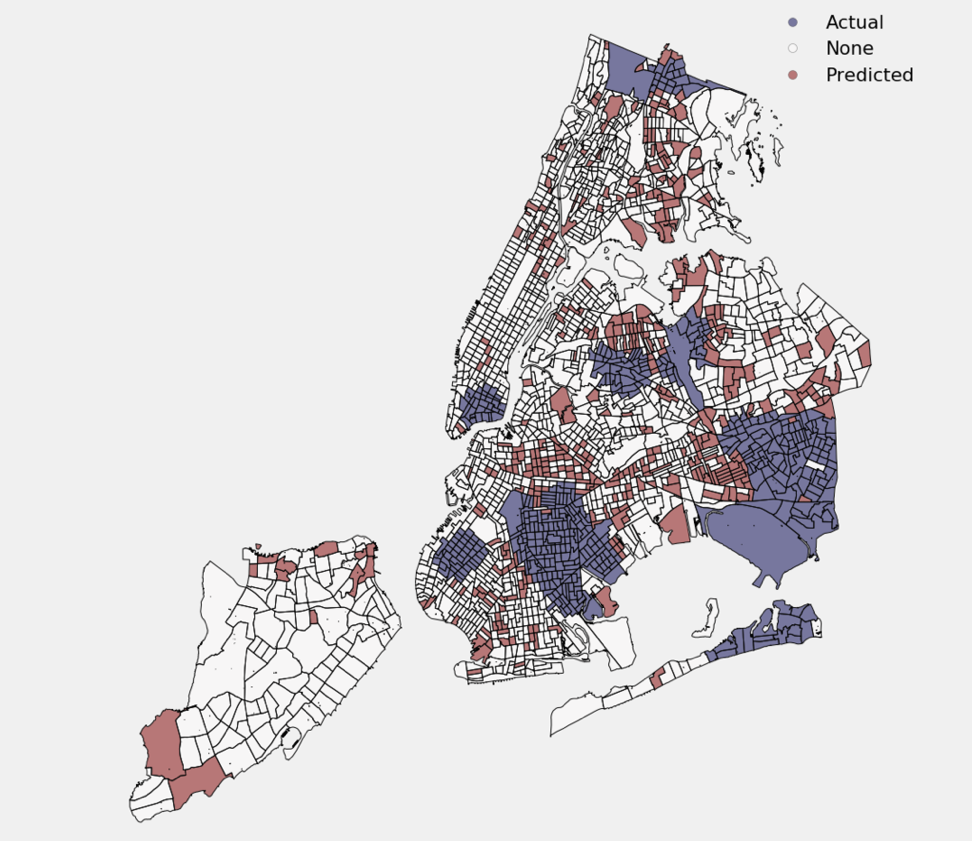


Figure 3: Actual and predicted dollar van origins

For destinations, the model correctly classified 100% of the actual dollar van destinations in the training set and 74% for the test set. For places that are not actually dollar van destinations, the model identified 23% of the training set and 25% of the test set, which we could interpret as places where there could potentially be demand for dollar vans to run to. The map below shows the results when the model is run on all the census tracts:

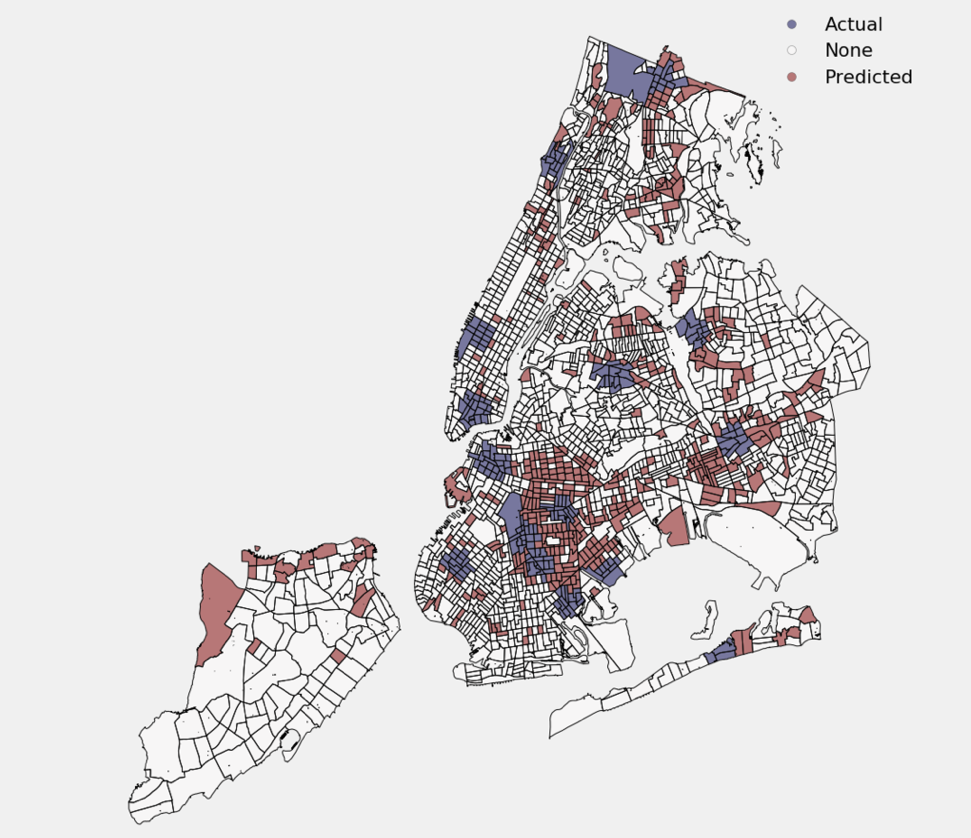


Figure 4: Actual and predicted dollar van destinations

# Ethical Considerations

The Dollar Van system exists outside of any planning mechanism, with no regulation, no academic assistance, no technocratic know-how, and no publically financed support. The forces that shape this system therefore escape institutional biases and spring forth directly from community need.

The organic nature of Dollar Vans has two distinct implications. Because Dollar Van’s appeared without transit's traditional planning process, the forces that shape the van’s routes, the van’s behavioral patterns, and the van’s effectiveness can be understood to be a case study in what certain communities look for in cheap transportation. Second, it is imperative that any conclusions made using this study do not make normative judgments about the effectiveness or organizational structure of the dollar van industry. People in government or the academy do not necessarily know how to ‘do it better’ and should not assume any superiority in any aspect of existing transit service.

Finally, it is important to note that Dollar Vans are owned, operated, and patronized overwhelmingly by people of color. While bringing dollar van’s out of the shadows and into harmony with New York’s other transportation options is the right thing to do, one must ensure that these systems continue to be creatures of, by, and for, the communities they serve.

# Next Steps

This dollar van study only covered a handful of the already discerned dollar van routes in New York City. Future investigations could incorporate lines that go beyond the New York City limits. This is especially interesting in New Jersey as Aaron Reiss has documented several van routes that travel north south corridors across the Hudson and then use the Lincoln Tunnel and George Washington Bridge to drop people at transit hubs in Manhattan. Another area of fruitful study could be a more comprehensive examination of dollar van routes that have not been documented.

An omission that should be corrected in a future study was the lack of a measure of transit richness in each census tract. While this is somewhat captured by our commute time data, simply including the number of subway stations or bus stops in a tract would have potentially produced interesting results.

A weakness of this study’s framework is how it ‘bakes in’ existing features that predict demand. We are presenting a snapshot of the dollar vans under study in a particular time and place. It well understood that transit demand is highly dynamic and fluctuates with economic conditions, technological change, and demographic shifts. This model takes none of these important ideas into account.

# Conclusions

This study demonstrates that the underlying reasons people ride dollar vans vary from place to place. In this study of only three routes, important distinctions arise between the lines that serve New York’s Chinatowns and those that serve Jamaica Station and Flatbush. For the successful regulation of dollar vans, the New York Department of Transportation (NYDOT) should recognize that the riders of dollars cannot be generalized into a monolithic constituency.

The Chinatown lines are characterized by large constellations of non-English speaking, relatively low income workers, who live in middle income neighborhoods. By and large, the dollar van users of these neighborhoods spend much less time commuting than most New Yorkers. This stands in stark contrast to the Flatbush and Jamaica Station lines that serve very high English speaking, non-white, middle income neighborhoods. These outer borough dollar van riders would otherwise have to spend more time commuting than nine out of ten New Yorkers. These divergent demographic profiles demonstrate that different lines can serve different purposes. Chinatowns connect to one another and provide a cultural link to those who do not speak English. This was anecdotally confirmed by our own van rides where Chinese New Yorkers were travelling to receive medical services in the Chinese language. The other lines serve more of an economic function as their destinations are job rich and their origins are not. Both justifications support robust dollar van lines.

This discrepancy presents a challenge to TLC and DOT. The reasons that the Chinatown dollar van lines thrive are not traditionally thought of as a core mission of public transit. The subway, bus, and road infrastructure was not designed to facilitate cultural and community connections. The dollar van system fills a need that would be very hard for the City to addresses due to scale, inequalities of information, and constantly changing neighborhood patterns. Chinatown vans can be seen as the community’s response to lack of mobility within the broader New York Chinese speaking community. Contrast that to the salient features of the Jamaica and Flatbush lines. These features are right in the wheel house of transit agencies, connecting residents to economic opportunity. Here, dollar vans are compensating for the lack of traditional transit as their non-dollar van calculated commute times are some of the highest within New York City.

When granting permits to new dollar vans, DOT should take this discrepancy into account. If the van proposes to connect cultural areas of the city with one another, there is no real viable alternative the city can provide. In this sense, the van route is solving a need that fits well within the city’s existing transit landscape. When permitting vans to the other lines, the city should recognize that dollar van demand in these places may signify a lack of transit infrastructure investment. It is not a reason to reject permits, as the residents of these neighborhoods need dollar vans to meet basic economic needs, but it should signal that these neighborhoods deserve a more permanent solution to reduce their commutes.

Finally, the other takeaway for the agencies is that it appears there are large swaths of New York that could support dollar vans. The predicative model demonstrates that lots of census tracts outside of the lines within the study have broadly similar characteristics to those that already support dollar vans. This includes broad swaths of Central Brooklyn, Eastern Queens, and the far north of the Bronx. Any permit requests for areas outside of these areas, for example a proposed route between Union Square and Williamsburg, would benefit from additional scrutiny as they would be meeting fundamentally different needs than the lines in this study.