The marginal congestion of a taxi in New York City

Daniel Mangrum*
Federal Reserve Bank of New York

Alejandro Molnar*
The World Bank

December 9, 2020

Abstract

We exploit the partial deregulation of New York City taxi medallions to provide a causal estimate of the impact of vehicle density on congestion. We employ taxi trip records to measure historical street-level speed. We find the taxi deregulation policy caused a local 8-9% decrease in speed. We estimate an empirical congestion elasticity curve from heterogeneity in vehicle density, counted from aerial orthoimagery. Additionally, we document a substantial traffic slowdown since 2013 using novel urban sensor data. Most of the slowdown in midtown Manhattan is accounted for by new supply from ridehail applications. We also assess current congestion pricing policies.

Keywords: Urban transportation. Congestion. Externalities. Regulation. Taxi. Rideshare. Ridehail. Transportation Network Companies. JEL: D22, D62, L91, L92, R41, R42

*We are grateful to Gilles Duranton, Federico Gutierrez, Christian Hansen, Tom Hubbard, Wang Jin, Robin Lindsey, Andrea Moro, Tobias Salz, Pedro Sant'Anna, Florian Scheuer, Paulo Somaini, Matt Turner and Arthur Van Benthem for comments as well as seminar participants at the National University of Singapore, the Dept. of Economics and the Transportation Center at Northwestern, San Andrés, Stanford GSB, Vanderbilt, the Kleinman Center at Wharton, the Econometric Society 2017 Winter Meeting, IIOC 2017, TRF 2017, INFORMS 2018 and UEA 2018. Moustafa Awad Ismaail provided excellent research assistance digitizing aerial imagery. Mangrum is thankful for support from the Kirk Dornbush Summer Research Grant, and Molnar for support from the Research Scholar Grant and the "Big Data Architecture" trans-institutional grant, both from Vanderbilt University. Emails: daniel.mangrum@ny.frb.org and aimolnar@worldbank.org

The congestion of transportation infrastructure has both macro and microeconomic costs. Transportation affects the organization and scale of economic activity in cities (Redding and Turner, 2015; Ahlfeldt *et al.*, 2015): while agglomeration saves transportation costs, these savings diminish as density congests infrastructure. When transportation infrastructure becomes congested, residents both bear and impose external costs. In the context of road congestion, costs include the value of time (Anderson, 2014), pollution and health (Currie and Walker, 2011; Knittel *et al.*, 2016; Simeonova *et al.*, 2018), carbon emissions (Barth and Boriboonsomsin, 2008) and subjective well-being (Kahneman and Krueger, 2006; Anderson *et al.*, 2016).

We document that traffic in New York City, one of the most productive cities in the United States (Hsieh and Moretti, 2017), has slowed down substantially since 2013. Taxi trips within midtown Manhattan have slowed down on the order of 15%, but the slowdown is widespread: it shows up in taxi trip records throughout the day and over the entire city, on highways – as measured though a novel dataset that we assemble from EZ-pass sensors – and in responses to the American Community Survey.

Cities such as New York are undergoing technological and policy changes with uncertain costs and benefits: roads may be under increasing demand from transportation network companies (also known as ridehail) and parcel delivery services, while space in the city is increasingly allocated to bikeshare, bike lanes and pedestrians (Sadik-Khan and Solomonow, 2016).

We evaluate a policy change that is unique to New York City: starting in August 2013, the city partially deregulated its medallion taxi industry by authorizing a new class of taxi. "Boro" taxis are almost identical to the city's 13,237 traditional yellow medallion taxis, but are painted green and are restricted from picking up passengers at airports or in Manhattan south of East 96th street and West 110th Street. We exploit the unique natural experiment provided by the roll-out of up to 6,539 new taxis searching for passengers under a well-defined spatial constraint to causally estimate the impact of taxi supply on traffic congestion in New York City.

We employ a dataset with 1.3 billion taxi trip records from 2009 to June 2016, which contain precise times and geographic coordinates of pickups and dropoffs, to construct a dataset of street-level speed that spans both sides of the exclusion boundary in northern Manhattan and the period prior to and after the roll-out of boro taxis. Our measures of speed are constructed at a high spatial resolution by selecting the trips that involved a pickup and a dropoff along the same avenue (about 69,300 trips per month in our area of interest) and either averaging or projecting travel times onto street segments or flexible functions.[1]

[1]Our analysis focuses on traffic flow along the north-south avenues because crosstown trips are

Avenues flow faster toward the north of Manhattan, as density and traffic decline, and all avenues have slowed down over time. Following the introduction of boro taxis, however, a stark speed gap opened up north of the boro taxi exclusion boundary (see Figures B.2 and B.3). In our baseline estimates we find that the boro program's roll-out caused an overall 8-9% slowdown in the speed of traffic in northern Manhattan.

Our estimates are robust to controlling for observable changes in road use (such as the deployment of bike lanes and bikeshare, among others), as well as a large set of specifications to address unobservable confounders. Unobservables that could be a cause for concern include differential trends over the city's geography (for instance, due to economic recovery or gentrification) or a differential impact from the arrival and growth of ridehail as a transportation alternative. Estimates of the boro program's impact track the timing of the roll-out of taxi supply, which ramped up over the last quarter of 2013, and are identified under narrow time windows around the roll-out. Our estimates are also robust to trends by avenue and treatment status that are linear in time but are also allowed to kink and track the rapid growth in ridehail. We also find that our results are not sensitive to differential time trends over space in the measurement error of geographic coordinates, a potential confounder in research that employs data from GPS units within dense urban settings. We address spatial spillovers, which could arise naturally as boro taxis transport passengers south of the pickup exclusion boundary, and congestion could be transmitted across the boundary through gridlock: alternative estimation strategies are consistent with the upper range (9%) of our impact estimates.

We next exploit the effects of the policy to estimate congestion externality parameters in terms of vehicle density, either for taxis or all vehicles on the road. To quantify the density of taxis and other vehicles, we collect a sample of 29 aerial orthoimagery scenes going back to 2010. We digitize the location of 132 thousand yellow taxis, boro taxis and other vehicles in the imagery. Our data allows us to measure changes to the patterns of vehicle density over space, comparing the period before and after the roll-out of boro taxis as well as increased vehicle density in midtown Manhattan due to growth in ridehail services.[2] We employ this data to directly measure changes in taxi and other vehicle density at a high degree of spatial resolution, which in turn allows us to study

interrupted by Central Park in our area of interest. Trips on either side of the park are scarce: 1345.5 trips per month across all streets on the East side area of interest, and essentially no trips on the West side.

[2]To our knowledge, ours is the first application of vehicle counts from aerial imagery in economics, or to the study of the long-run impact of regulation on vehicle density and congestion in any discipline. See Donaldson and Storeygard (2016) for a discussion of satellite imagery in economics. Traffic engineers have employed aerial imagery for decades (e.g. Johnson, 1928; Greenshields, 1948) for infrastructure design studies and more recently to calibrate agent-based models of traffic flow. A recent literature in machine learning addresses the detection of vehicles from aerial images (e.g., Reilly *et al.*, 2010, Mundhenk *et al.*, 2016; Razakarivony and Jurie, 2016) but its output has yet to be applied in empirical research.

the heterogeneity of the boro program's impact over the boro zone and project it onto location attributes such as measures of urban density. We estimate changes in travel time to changes in vehicle density (i.e., congestion elasticities) flexibly and as a function of underlying urban attributes.

In a first application, we apply our estimated congestion elasticity curve with a study of New York City's slowdown since 2013. We document that traffic in midtown Manhattan slowed down by a median 15.2% from June 2013 to June 2016, and corroborate this finding by contributing two new sets of time series to the state of knowledge on traffic speed and volumes in New York City: i) archived highway speeds reported from a network of EZ-pass sensors going back to 2013, and ii) vehicle miles traveled by vehicle class, estimated from car inspection odometer data.

We digitize aerial imagery samples in an area of midtown Manhattan and find a 14.1% increase in vehicle density over the three years before and after the summer of 2013. We also detect a substantial change in the observed color composition of (non-taxi, non-truck) cars in midtown Manhattan: whereas 29.3% of cars counted in overhead imagery prior to the summer of 2013 were black, that share has since increased to 45.7%. VIN-level records from the New York state Department of Motor Vehicles indicate that while 20.9% of privately owned cars in New York City are black, 73.0% of For-hire vehicles (FHVs, a category that includes ridehail) are black. Employing color-share data, we estimate that the density of for-hire vehicles in midtown Manhattan has increased by 222.9% since the summer of 2013.[3]

Our estimate of the location-specific elasticity of congestion to taxi density is 0.358 for midtown Manhattan. Combined with the net 30.5% increase in aerial counts for both taxis and FHVs, we find that the initial phase of ridehail growth in midtown Manhattan can account for 71.7% of the contemporaneous slowdown in traffic in the area. Overall, evidence from aerial counts and the spatial and time series patterns of New York City's slowdown do not appear to be consistent with rival explanations such as bike lanes, construction or use by cyclists, pedestrians or trucks. Estimates on social costs from ridehail's deployment to a new market are relevant to thousands of cities that continue to face regulation decisions on how to regulate transportation network companies.[4]

[3]Taxis were down 11.8%. Aerial images combined with color-share data allow us to separately estimate the density of vehicles other than taxis, such as ridehail and private cars (for which we estimate a 6.8% decline). An advantage of our approach is that we can measure vehicle density in dense, downtown city streets. Previous studies using counts from loop detectors, stationary traffic cameras or gantries (e.g. Xie and Olszewski, 2011; Li *et al.*, 2017) are usually limited to highways or arterial roads.

[4]As of 2020 Uber is present in 70 countries and 927 cities. Lyft is in 657 cities in North America, Bolt is in 289 cities worldwide, and Ola claims to be present in 150 cities in India and a further 100 worldwide. Didi is present in 7 countries, Grab in 8, and Yandex.Taxi in 17, among other providers. Uber and Lyft accounted for 446 million trips in the US in the last quarter of 2019 (estimated as 178.4 million over a 0.39 market share,

We apply our local estimates of congestion elasticities to evaluate recent and future congestion pricing policies in New York City. We estimate tract-level marginal social damages, and find that the social cost of driving anywhere in Manhattan south of 125th St is between twice and 16 times larger than the "Congestion Surcharge" of $2.50 (for taxis) or $2.75 (for ridehail) enacted by New York State in February 2019. We also explore the implications of our estimates to proposed fees for New York City's congestion pricing plan, which has been approved by the New York state legislature and awaits final fee-setting and implementation. Over a range of assumptions on demand we find that proposed congestion fees would save between 0.9 and 3.1 minutes on the average commute.

Section 1 below describes the boro taxi program. In Section 2 we estimate the impact of the boro program on traffic speed, for which we rely on historical street speed data constructed from taxi trip records (methods are in Appendix B). In Section 3 we derive a congestion elasticity curve for New York City in terms of underlying urban density. We employ our congestion externality estimates in two policy applications. In Section 4 we quantify the increase in midtown Manhattan vehicle density due to ridehail, and account for its impact on slower traffic speed. In Section 5 we evaluate two sets of congestion pricing policies, one of which was recently passed and another is currently under debate in New York City. Section 6 concludes.

1 Taxi regulation in New York City and the launch of "boro" taxis

Taxi service in New York City is regulated by the Taxi and Limousine Commission (TLC). Among other regulations, the TLC sets fares, licenses drivers and cars and issues medallions, which are a fixed number of permits that must be attached to a licensed car for it to operate as a taxi (i.e. accept street hails) in the five boroughs of New York City. The number of medallions was fixed at 11,787 for decades, with a few small increases after 1996, and remained at 13,237 between May 2008 and August 2013. The only type of medallion taxi available up to August 2013 was the traditional yellow taxi. Although yellow taxis are statutorily allowed to pick up passengers in all five boroughs of New York City, most taxi pickups occurred in Manhattan, particularly midtown, and at the city's two airports. The spatial distribution of yellow taxi pickups is illustrated in Figure G.10 using data for June 2013. Areas outside of northern Manhattan and the remaining four boroughs had limited taxi coverage, and throughout most of New York City residents

from Lyft, Inc. S1 statement to the SEC).

could not dependably hail a yellow taxi from the curb. Prior to ridehail apps the most common method for residents in the outer boroughs to hire car transportation was to call radio-dispatch bases for pre-arranged transportation from "for-hire vehicles," as well as unlicensed street-hails offered by some for-hire vehicles and other informal providers.[5]

In August 2013 a new class of taxi was allowed on the road: painted green and referred to in regulation as "street-hail livery" (SHL) taxis, or alternatively as "boro" or green taxis, these taxis and their drivers are regulated like yellow taxis with a major distinction: passenger pickups (whether pre-arranged or street-hails) are not allowed at the two city airports or south of East 96th Street and West 110th Street in Manhattan, an area referred to as the "hail-exclusion zone." The objective of the boro taxi program was to increase street-hail availability in the outer boroughs and provide a regulated alternative to unlicensed street-hails.[6] Following a multi-year political and regulatory process, the roll-out of the boro program involved an initial sale of 6,000 medallion permits, as well as a fleet roll-out and driver recruitment process carried out by medallion owners. Additional permits were put on sale in August 2014.[7]

2 Impact of the boro program on traffic congestion

In this section we describe how we exploit the roll-out of the Street-Hail Livery program to evaluate the impact of boro taxis on congestion in northern Manhattan. Our area of study spans 1 kilometer of the Upper East Side and 2 km of East Harlem on the East Side, and 2 km of Upper West Side and 1 km of Morningside Heights and Harlem on the West Side. Urban density and socioeconomic status decline towards the north, although not in a monotone or smooth manner. Prior to the launch of the boro program, the declining gradient in urban density was associated with lower yellow taxi pickup activity, with spikes around major intersections, which include subway stations.[8] The boro taxi program was intended to increase supply in areas undeserved by traditional yellow taxis, but its implementation through a "hail-exclusion" zone south of East 96th and West 110th streets created an incentive for boro taxi drivers to cruise for hails in the area just north of the boundary: whereas this area was of marginal cruising value for yellow

[5] See Gao Hodges (2012) for the history of the city's taxi cab industry, including unlicensed activity.

[6] See www1.nyc.gov/office-of-the-mayor/news/362-13/mayor-bloomberg-more-1-000-boro-taxis-now-new-york-city-street for a NYC government press release on the roll-out of the program and its intended goals. Other differences with yellow taxis are that 20% of boro medallions require the taxi to be wheelchair-accessible, and the taxis are required to undergo vehicle inspections at half the regularity of yellow taxis.

[7] Figure 4 contains a time series of the number of unique boro taxis that recorded at least one trip during each month.

[8] As shown in Fig. G.10 most yellow taxi activity is concentrated at airports as well as midtown and downtown Manhattan.

taxi drivers, it is a corner-solution for boro taxi drivers, leading to a localized shock to supply.[9] Our empirical strategy in this section consists of exploiting this localized shock, and its roll-out over time, to evaluate the impact of taxi supply on traffic speed.

A first challenge is that there is no publicly available dataset that contains information on street speed at a high level of spatial resolution in New York City and over a historical period going back prior to the roll-out of the boro taxi program.[10] We employ historical records of taxi trips from the Taxi and Limousine Commission (collected from GPS-enabled taxi meters under the "Taxicab Passenger Enhancement Program") to construct measures of historical street speeds. Geographical coordinates are present in the data from January 2009 through June 2016, which defines our sample period. We describe and assess three alternative methods for constructing retrospective speed data in Appendix B. Bias-variance tradoffs lead us to prefer a simple method that consists of measuring street travel time along a city block by taking the average travel time from all taxi trips that cross that block. We present robustness results to the alternative methods.

Historical data on street speed at a high level of spatial resolution allows us to evaluate the impact of policies that impact road use in a well-defined spatial manner. Figure 2 summarizes our average travel time data (plotted here as speed) across all avenues.[11]

We employ a difference-in-differences specification to estimate an average reduced-form impact from the boro taxi program on avenue speed in northern Manhattan. On avenues with two directions of traffic, such as Park Avenue, we classify trips according to direction. We refer to each avenue and direction of traffic as a "run." Let s_{rbm} be a travel time outcome, which in most specifications is the log seconds per meter at the level of an

[9]This notion holds under a spatial search equilibrium as in Buchholz (2020): areas with high baseline demand support higher search in equilibrium, although the endogenous continuation value for above-boundary pickups may be shaded down by a higher transition probability into the hail-exclusion zone.

[10]Recent studies of traffic congestion have employed probe data from cell-phones, obtained from Google Maps (e.g. Hanna *et al.*, 2017; Akbar and Duranton, 2017; Akbar *et al.*, 2018), but historical data was not available to us over our period of study. Additionally, travel time APIs are developed with the aim of prediction and their methodologies may changed in ways that render them not comparable over long periods or over space. Another alternative, the INRIX dataset, is constructed from probes on fleets of commercial vehicles and is not as attractive as our methodology for dense urban settings such as New York City.

[11]Avenue-level figures are available in Figures B.2 and B.3. Inspection of these figures reveals the identifying variation in the data that drives our main difference-in-differences estimates below: historical speeds increase in the northern direction along most avenues, along with declining urban density and traffic, and as time goes on all streets become slower, but particularly so north of the hail-exclusion boundary, an area we refer to as "boro-zone." Particularly stark examples include 5th Ave (Fig. B.3a) and Lexington Ave (Fig. B.3e), but this outcome can be observed on most avenues. This is always the case on the East side, but not on the West side where some avenues both narrow and become two-way north of the hail exclusion boundary. The time-invariant attributes of streets, such as changes in width, the direction of traffic, or bends in the roadway will be absorbed in specifications described below by fixed effects at the level of the avenue run and block.

avenue run (r), block (b) and month (m) from January 2009 to June 2016. In a standard difference-in-differences specification we would define the post dummy as August 2013 and all months thereafter:

$$s_{rbm} = \beta_{DID}\text{post}_m \times \text{boro zone}_{rb} + \beta x_{rbm} + \text{fixed effects} + \varepsilon_{rbm} \quad (1)$$

However, the supply of boro taxis was rolled-out gradually: only 114 boro taxis carried out at least one trip during the program's first month of August 2013, and this number rose to 445 taxis in September, 1,128 in October, 1,754 in November and 2,736 in December. Although this ramp-up was relatively fast, it took until June 2015 for the program to reach the highest number of boro taxis ever observed to be active in a given month, which was 6,539. The time series for the number of unique boro taxis is plotted as the solid green line in Figure 4. We define the $\textit{boro roll-out}_m$ treatment intensity measure as the ratio of boro taxis active each month (zero prior to August 2013) to the maximum value of 6,539, and our preferred specification is to employ this intensity measure in difference-in-differences specifications of the form:

$$s_{rbm} = \beta_{DID}\text{boro roll-out}_m \times \text{boro zone}_{rb} + \beta x_{rbm} + \text{fixed effects} + \varepsilon_{rbm} \quad (2)$$

Our baseline specifications includes fixed effects δ_{rb} at the level of each run and block, absorbing time invariant features of the road, as well as a set of monthly time-effects δ_{rm} at the level of each run, which absorb the variation in speed along each avenue over time that is common across blocks. The main difference-in-differences coefficient under this fixed effect specification is estimated from the differences within each avenue between the travel time north vs. south of the hail-exclusion boundary as the boro program was rolled out. For regressions based on eqs. (1) or (2) to provide an estimate of the causal impact of the boro program we require the parallel trend assumption that street speed in the hail-exclusion "control" zone and boro "treatment" zone are not trending apart from each other for reasons other than the impact of the program. We will discuss below the robustness of our estimates to trends in unobservables, along with alternative fixed-effect structures and specifications, but first we consider a number of observable features of the urban environment that vary at the level of road segments over time and could act as confounders for our estimate of the boro program's impact on traffic speed.

A first source of potential, observable confounders consists of changes in the use of road space, for example for transportation alternatives. Bike lanes were expanded substantially throughout NYC during the period of our study, so we use time-stamped

maps of the bike lane network to control for bike lane expansion in our area of study.[12] A second change in transportation infrastructure with a potential impact on road use and traffic speed was the deployment and northward expansion of the Citibike bikeshare system. Citibike launched south of Central Park - outside our area of interest - in May 2013, but expanded into the upper east and west sides over the summer of 2015. We map the deployment of Citibike stations over time as well as, in an alternative specification, all areas along an avenue south of any newly opened Citibike station, to control for both stations and cyclists.[13] Additional potential confounders include changes to road conditions. We construct count measures at the road-segment level for complaints reported to the city's "311" municipal services hotline, for conditions that could impinge traffic. We also construct count measures for outstanding potholes. Additionally, two re-zoning events took place during our period of study, the establishment in June 2012 of Special Enhanced Commercial Districts 2 and 3 along a number of avenues on the Upper West Side. Although this re-zoning was entirely commercial in character, intended to regulate "ground floor frontages" to preserve commercial activity and "multi-store character," in some specifications we include a control for these zoning changes. Details on the construction of these controls are in Appendix A.

We report the results for our baseline specifications on Table 2. The table is organized to report results for our three alternative measures of travel time: average speeds as in Appendix B.1 in Cols. 1 and 2, B-spline projected speeds as in Appendix B.3 in Cols. 3 and 4, and dummy projected speeds as in Appendix B.2 in Cols. 5 and 6. Within each of these, we report the results for Eqs. (1) and (2).

Comparing either Cols. 1, 3 and 5 or 2, 4 and 6 we can see that results for a given specification are similar across alternative estimation methods for the travel time outcome.[14] In all specifications it is also the case that the DID coefficient estimate on the $boro \times post$ dummy is lower than on the $boro \times roll\text{-}out$ intensity measure, which is what we would expect if a partially-rolled out boro program has a less than full impact on traffic speed. Once the full roll-out of the program is allowed for, we find travel time increases of 8-10% in the boro zone relative to the hail-exclusion zone.[15] To further

[12] See Sadik-Khan and Solomonow (2016) for the former NYC Department of Transportation commissioner's account of the substantial expansions in bike lanes and pedestrian plazas over this period.

[13] Hamilton and Wichman (2018) find the bikeshare infrastructure reduced traffic in Washington D.C.

[14] We employ an exhaustive set of fixed effects: this results in an R^2 well above 0.9 for specifications (1) and (2), but declines to around 0.5-0.6 for projection-based travel times due to higher variance in that outcome variable.

[15] If we had defined "full roll-out" as a ratio relative to a lower denominator (e.g. the average number of boro taxis on the road after a certain period, rather than the maximum of 6,539 taxis) then the estimated impact from the program on traffic speed would be slightly higher, but there would be no effect on the congestion estimates reported in Section 3 below.

evaluate the role of timing and pre-trends in our specification, we re-run a version of Eq. (2) in Col. 2 in which the pre-treatment period is defined to start 24 months prior to the actual introduction of the boro program (i.e in August 2011) and we estimate a month-specific $\hat{\beta}_{DID,m}$ coefficient for every month thereafter. Figure 4 plots (in blue) the point estimates and 95% confidence intervals for these coefficients: we can see that there are no pre-trends, as monthly coefficients are flat at zero prior to the actual roll-out of the program. The speed difference between the hail-exclusion and boro zones starts to develop shortly after August 2013, quickly trending upward over the first six months, and then more slowly over the next year. The trend in $\hat{\beta}_{DID,m}$ coefficients tracks the trend in the roll-out of boro taxis: the measure of this roll-out is not used in the estimates plotted in Figure 4, but it is overlaid on the figure for comparison.[16]

In Table 3 we evaluate some alternative specifications to assess our claim that we have estimated an 8-10% causal impact of the boro program on traffic speed (estimates of $\hat{\beta}_{DID}$ can be read equivalently as log increases in travel time or log decreases in speed). Col. 1 presents an estimate in levels, rather than logs, which is statistically significant and consistent with our baseline results: at the sample average of 0.182 seconds per meter (i.e. 12.32 miles per hour) this estimate represents a 7.2% decrease in speed (to 11.44 miles per hour). Col. 2 presents estimates with common time effects (i.e. fully saturated at the run-block and at the monthly level, but not interacted with runs). Relative to our baseline, this specification allows variation in travel time between runs over time to identify the difference-in-difference estimates.

The major threat to a causal interpretation of the estimates from our research design is that unobservable components of traffic speed could have different trends in the hail-exclusion "control" and boro "treatment" zones, for reasons other than the program's impact. We check against any such linear trends in Col. 3, and find that the estimated impact decreases by only 17% after the inclusion of a linear trend interacted by run and treatment status. Robustness to a difference in unobserved linear trends is a strong test for the validity of a difference-in-differences design (since identification is provided by the timing of the boro program's launch in August 2013 and its subsequent roll-out relative to any linear trend), although this is less the case in a setting such as ours in which all treated units change status at the same time.

[16]Close attention to Figure 4 reveals that the "dynamic DID" coefficient estimate spikes in August and September of each year. These spikes are explained by a seasonal violation of the parallel trends assumption in our DID specification: every August (and to a lesser extent, September) there is a vacation-driven decline in both the population of taxi users and drivers in NYC (see Figure 1 for a monthly time-series of taxi trips that displays negative seasonality in August). This vacation-induced reduction in traffic and increase in speed is stronger south of the hail-exclusion boundary during the "post" period (whereas the opposite occurred in the summer of 2010), resulting in the seasonal variation in travel time differences displayed in Figure 4.

We cannot check for robustness against arbitrary, non-linear differences in trends between the treatment and control zones, since the set of these trends contains the roll-out of the boro program that we rely on for identification. We can, however, check the robustness of our estimates against an alternative, well-defined trend that may be a cause for concern. The major recent change to NYC transportation that we have not yet addressed in detail was the growth in ridehail (plotted in Fig. 1), starting with Uber's entry as a black-car service in May 2011. Effects from ridehail's growth on the slowdown of NYC's streets (as documented since 2013 in Appendix D, and for our area of interest by the downward, parallel translation over time of the speed curves shown in Figures B.2 and B.3) could potentially bias our estimates. Although our research design controls for ridehail's potential effect on traffic speed to the extent that this effect was proportional both north and south of the hail-exclusion boundary, differential impacts across zones remain a potential confounder. To address this concern we construct an "aggregate ridehail trend" from monthly ridehail trips, and interact this trend with runs and the treatment zone.[17] In Col. 4 we jointly include the ridehail trend along with a linear trend: we find that our reduced-form estimate of the impact of the boro program is robust to a trending differential unobservable that is not only linear, but is allowed to accelerate after May 2011 at any rate proportional to the growth of ridehail in New York City.

We present Cols 5 and 6 as checks on statistical inference.[18] In Col 5 we present the results from averaging, for each spatial unit, all time periods into a before and an after period. This is clearly not a preferred specification, as it eliminates variation in observables over time (from controls, such as bike lanes, or the identification provided by the gradual roll-out of boro-taxis that was documented in Fig. 4), but it is useful as a check that statistical significance is not an artifact of time series dependence. We then collapse all variation between road intervals into a single segment average for the treatment or control zone, as a further check that inference does not depend on unmodeled spatial dependence along each avenue. Col 6 presents the difference-in-differences estimate from this further step, which averages each avenue into four cells: both time periods vs. the avenue's treatment and control segments. Under this specification, which contains only 62 observations (and a fixed effect for each of 17 avenue runs, of which 16 are included in the control zone and 15 in the treatment zone) we still find a statistically significant

[17] Trip-level reporting by ridehail providers to the TLC started with the second and third quarters of 2014. There is a gap in reporting during the fourth quarter of 2014, and reporting has resumed from 2015 onwards. We interpolate the monthly growth rate for the fourth quarter of 2014, and extrapolate backwards to Uber's launch in May 2011 using the 10.1% monthly growth rate that has been observed for all ridehail providers since the second quarter of 2014.

[18] Standard errors in all other specifications are two-way clustered at the level of run$\times$ block and run$\times$ month. Statistical significance holds at coarser clusters such as runs.

3.8% additional slowdown of the boro zone relative to the hail-exclusion zone control, which in turn slowed down by 5.8%.[19]

In Table 4 we present results under alternative data samples. Cols 1 and 2 look at narrow event windows around the introduction of boro taxis (6 and 12 months before and after, respectively). These results confirm that the boro program's impact is identified from variation in traffic speed around the roll-out period (as shown previously by the month-level "dynamic difference-in-differences" results plotted in Fig. 4), rather than long-run trends or events outside the period in which the program was introduced. In particular, results for both windows exclude a city wide-reduction in the speed limit from 30 mph to 25 mph, which took place on November 7, 2014.[20] Speed limits were changed along the entire length of the avenues in our study, so our baseline specifications already control for avenue-level effects through the monthly time-effects δ_{rm} at the level of each run. Results from Cols. 1 and 2 further confirm that the speed limit change does not introduce a violation of the parallel-trends assumption.

The remainder of Table 4 evaluates alternative samples over space. Cols 3 and 4 break out results into the east and west sides of northern Manhattan: we find a heterogeneous impact of the boro program, with double the travel time increases on the east than on the west side. We further investigate the variation in the program's impact over space by estimating an interaction of boro roll-out and an indicator for every street north of 86th (rather than a single indicator for the entire zone). In Figure 3 we plot the result of such specifications separately for east and west and for three separate outcomes: log average travel time, log taxi pickups (as an indicator of taxi activity or supply induced by the program) and (for the east side only) changes in the log of aerial counts of taxis obtained from digitized aerial orthoimagery.[21] Figure 3a shows that taxi pickup activity jumped sharply at the boro zone boundary, relative to the hail-exclusion zone where pickups were unchanged. We also find that taxi activity (as measured by pickups) increased more markedly on the east than the west side, which accounts for part of the difference in the

[19] We omit averaged controls, since at this level of aggregation they become an arbitrary set of free parameters. Our estimate of the $\hat{\beta}_{DID}$ coefficient is still significant and higher in magnitude if we include averages of our controls, or if we restrict our sample to the balanced panel of 14 runs that span both the treatment and control zones. Results employing the B-spline projection travel time outcome are consistent in sign and magnitude, but not significant at this level of aggregation.

[20] The speed limit change was part of the broader "Vision Zero" traffic safety program, announced in January 2014 by NYC mayor Bill de Blasio and rolled-out in stages thereafter. The program consists of a broad set of initiatives intended to eradicate traffic fatalities, including public service announcements, a focus on enforcement of traffic laws, street intersection redesign and the change in speed limits. We additionally control for a number of these policies in Appendix Table 6. See www1.nyc.gov/site/visionzero/initiatives/initiatives_archive.page for further details.

[21] Aerial counts form the basis for the empirical applications in the second half of the paper, and are introduced in detail in Section 3.1 and Appendix C. Here we note that they are broadly in agreement with taxi pickups as an indicator of local taxi activity.

area-wide reduced form policy impact estimates in Cols 3 and 4. Figure 3b shows that travel times also increased substantially north of the boro zone boundary (relative to an average for 84th and 85th streets) but that changes in speed have substantial spatial pre-trends.[22]

Spatial pre-trends of this nature imply that the difference-in-differences coefficient from estimating Eq. 2 is attenuated as an estimator of the program's overall impact on traffic speed (as areas attributed as policy controls are slowed down by the trend). Unlike pre-trends in time, which we should not expect the program to produce, it is reasonable to expect the boro program to produce some form of spatial trend due to spatial spillovers. The channels though which the boro program could produce endogenous spillovers include: i) the direct effect of boro taxis on traffic as they drive south of the boundary to drop off passengers; and the incidence of such travel should diminish with distance to the boundary, ii) the direct effect of yellow taxis (or ridehail providers), which may retrench south of the boundary in response to the increased, localized competition from boro taxis to the north, and iii) an indirect effect transmitted through traffic congestion: a taxi may not only have a direct effect that slows down traffic at its immediate location, but also down the road through the effect of queuing cars, or on adjacent roads through gridlock (Schwartz, 2015).

To evaluate the sensitivity of our results to spillovers, in Cols. 5-7 of Table 4 we drop the intervals in the control zone that are 3, 6 or 9 blocks immediately south of the boundary. The estimated effect rises as we drop adjacent intervals, and we estimate an effect of 0.086 (i.e. 11% higher) if we exclude the 500m adjacent to the treatment boundary, but the sensitivity to this exclusion does not change after this point.

Lastly, we evaluate spillovers through spatial autoregressive models in which we allow for dependence in speed between neighboring road segments. Our aim in estimating a spatial model is to decompose the impact of increased taxi activity due to the boro program into the direct effect on an affected road segment vs. an indirect effect, and in turn use this decomposition to quantify the extent to which spillovers may bias the difference-in-difference estimation of the program's impact. Road networks are an ideal example of the structure in a pure spatial autoregressive model for a couple reasons:[23]

[22]Direct evaluation of aggregate and avenue-level speed data also reveals the spatial pre-trends. Figures B.2 and B.3 (see, in particular, the northbound direction of Broadway in Fig. B.2a, as well as B.3a and B.3g) show that following the program's roll-out, the speed on some avenues starts to decline in areas neighboring the hail-exclusion to the south.

[23]Spatial models are subject to the concern (Pinkse and Slade, 2010; Gibbons *et al.*, 2015) that a single reduced form (i.e. the regression of a dependent variable on lagged values of regressors) can be mapped to different sets of structural parameters depending on assumptions on the nature of spatial interactions that may be hard to verify (for example, an exclusion restriction on direct effects of lagged regressors distinguishes a pure spatial autoregressive model from one with spatial lags). We do not take a stand on

First, adjacency links are well defined and directed. We consider up to three possible interactions: i) cars can slow down cars that are behind them, so $s_{r,b}$ depends on $s_{r,b-1}$ if $b-1$ is ahead of b in the direction of traffic, ii) s_{rb} depends on $s_{r'b}$ if r and r' are different directions of traffic on the same avenue (note that this interaction is subject to the reflection problem if estimated in isolation) and iii) s_{rb} depends on $s_{r'b}$ if r and r' are neighboring avenues with the same direction of traffic. Second, exclusion restrictions on the structure of spatial lags are reasonable: if a car is slowed down by a double-parked vehicle or a bike lane that is two blocks up the road, this happens because a bottleneck and a slowdown are created ahead that are transmitted down the road through traffic speed, and not through the direct action of double-parking on a car that is two blocks away.

Let $\mathbf{y}$ be the vector of every run-block-month travel time, $\mathbf{W}$ the matrix collecting all adjacency interactions between run-blocks (i.e., the Kronecker product of the road adjacency network and the sample's time dimension), and $\mathbf{X}$ the matrix with the boro roll-out and control variables that vary at the run-block-month level. We model travel time with the equation $\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, where ρ is a spatial dependence parameter and $\boldsymbol{\beta}$ is now interpreted as a direct impact. Following Gibbons *et al.* (2015), we estimate this equation by instrumenting the spatial autoregressive term with its reduced form in terms of series of distributed lags for the exogenous variables, i.e. $\mathbf{W}\mathbf{y} = \mathbf{W}\mathbf{X}\boldsymbol{\pi_1} + \mathbf{W^2}\mathbf{X}\boldsymbol{\pi_2} + \mathbf{W^3}\mathbf{X}\boldsymbol{\pi_3} + \mathbf{W^4}\mathbf{X}\boldsymbol{\pi_4} + \cdots + \boldsymbol{\nu}$, where $\boldsymbol{\pi_i} = \rho^{i-1}\boldsymbol{\beta}$ and $\boldsymbol{\nu} = (\mathbf{W} + \rho\mathbf{W^2} + \rho^2\mathbf{W^3} + \dots)\boldsymbol{\varepsilon}$.

Table 5 reports model estimates using four spatial lags in the first stage and under two alternative assumptions for the spatial weighting matrix. Col. 1 employs a matrix $\mathbf{W_a}$ that includes all the interactions listed above, while the matrix $\mathbf{W_b}$ used for Col. 2 includes only backward effects along a direction of traffic. The estimated coefficient for spatial dependence is precisely estimated and almost identical across both models. Patterns of spatial dependence vary slightly across models, however, as these depend on the assumed adjacency matrix. In both models, the estimated spillover effect from the program across the first block of the program boundary is high (3.5 percentage points in Col. 1, 2.5 p.p. in Col.2).[24] Spillover effects decay thereafter, slightly faster in the three-adjacency model. The estimated impact of the boro program on traffic speed

this issue, since our aim is not to estimate structural spatial interaction parameters, but rather to evaluate the sensitivity of our baseline estimates to allowing for spatial dependence.

[24] If we let b indicate here an individual run-block, the vector of estimated total (direct plus indirect) structural effects for every run-block is given by $\hat{\mathbf{t}} = \hat{\beta}_{DID}\hat{\mathbf{S}}\mathbf{e}$, where $\hat{\mathbf{S}} = (\mathbf{I} - \hat{\rho}\mathbf{W})^{-1}$ and $\mathbf{e}$ is an indicator vector for the blocks in the boro zone. The average direct structural effect is $\hat{\beta}_{DID}\frac{1}{|\mathcal{B}|}\sum_{b\in\mathcal{B}}\hat{\mathbf{s}}_{\mathbf{bb}}$ where $\hat{\mathbf{s}}_{\mathbf{bb}}$ are the diagonal elements of $\hat{\mathbf{S}}$ and $\mathcal{B}$ is the set of blocks in the boro zone. The average direct plus indirect effect on the boro zone is $\mathbf{e}'\hat{\mathbf{t}}$ and the indirect effect on the boro zone is the difference between the previous two. Indirect effects on the control zone (in its entirety, or average spillovers over specified segments) are defined analogously by selecting and averaging from the vector $\hat{\mathbf{t}}$.

throughout the entire hail-exclusion zone in the sample is comprable across models: 1.08 p.p. in Col. 1, 1.22 p.p. in Col. 2, with confidence intervals spanning a range from 0.78 to 1.50 p.p. across models. We employ these magnitudes below to evaluate the robustness of our baseline results to spatial spillovers.

We review the addition of further controls to our baseline specification in Table 6, including types of bike lane and "311" hotline complaints, as well as commercial zoning.[25] The bottom line from the table is that the inclusion of a large set of road level controls has a negligible effect on our estimate of the boro program's reduced-form impact: in Col 1 we present results for our baseline DID specification, but with no controls, and find $\hat{\beta}_{DID} = 0.080$. Column 2 includes a large set of road controls, Col 3 includes a commercial zoning change which came into effect in June 2012, restricting ground floor commercial frontages with the goal of "(maintaining), over time, the general multi-store character (and) promoting a varied and active retail environment" along several avenues in the Upper West Side. Column 4 includes a set of traffic calming measures introduced as part of the city's "Vision Zero" traffic fatality initiative. The addition of all controls jointly (presented in Col. 5) has no effect on our headline DID estimate.[26]

The specification in Col. 2 includes a set of additional, more disaggregated controls of the alternative road uses included in our baseline (Table 2, Col 2), such as four types of bike lane, six categories for "311" complaints and a control both for Citibike stations and the road segment to the south of these. This disaggregation has no effect on the DID estimate, and neither does a full set of interactions (not reported) for the two count variables of potholes and complaints.[27]

[25]New York City provides a wealth of urban data to the public, its online portal is opendata.cityofnewyork.us. See Glaeser *et al.* (2015) for a discussion of urban data and empirical research. A common roadblock to the use of urban data in causal research designs is that cities do not always store or maintain timestamped versions of the data that they produce, whether from administrative records or urban sensors.

[26]The commercial zoning control is a step function that changes in June 2012 over a large area in the Upper West Side. We include zoning only as a robustness check, rather than in our baseline estimates, since it has a small incidence that may be an artifact of overfitting.

[27]It is not surprising that some of these controls have no incidence on our estimates: Citibike, in particular, only expanded into our area of study after the launch of the boro program, and by June 2016 had not expanded north of W 110 St or E 96 St. The table also reports on a secondary empirical finding on the effect of bike lanes on speed: although our study is not intended to evaluate the effect of bike lanes on street speeds in New York, and our area of study does not exploit all available variation in the deployment of bike lanes, we find that the introduction of bike lanes in northern Manhattan was correlated with slower street speeds, contradicting a previous report by the NYC Department of Transportation (NYC-DOT, 2014). In our baseline specification we find that avenues become 2.4% slower following the introduction of any bike lane, whereas the breakout across bike lane types in Table 6, Col 2 finds that the effect appears to increase in the lane type's utilization of road space: signed routes (which do not employ road space, and consist of signage indicating how cyclists should navigate an intersection) do not have a significant effect. Wide parking lanes (designed to provide additional space for cyclists) and standard bike lanes (which cars can invade, in a traffic violation) are associated with a 2.1% and 1.7% speed decrease, and protected bike

3 The elasticity of congestion

In the previous section we quantified the boro taxi program's impact on traffic speed in northern Manhattan. The program's effect on congestion is a first-order cost, to be evaluated by New York transportation regulators along with the program's transportation benefits such as increased taxi hail availability throughout the five boroughs.[28] In this section we estimate the change in congestion in terms of changes in vehicle density. We measure vehicle density both for taxis and for all vehicles using counts from digitized aerial orthoimagery, and we further quantify the heterogeneity in this elasticity by projecting it onto measures of the city's urban density. Our goal is to provide a flexible congestion elasticity estimate that can be applied outside of northern Manhattan to regulation issues in the city more broadly, such as congestion pricing, medallion deregulation[29] and the expansion in car-based transportation from "disruption" by ridehail.

TLC trip record data can be used to measure taxi activity at a high resolution in space and time, but only by means of trip endpoints (pickups and dropoffs).[30] Taxi pickups are an imperfect measure of taxi supply or density because they arise endogenously from matches between vacant taxis searching for hails and potential passengers attempting to hail taxis from the street curb. Taxi supply levels and wait times may lead potential

lanes (in which, usually, a parking lane is moved inward to separate cars from cyclists) are associated to a 3.2% speed decrease. It might be expected that bike lanes should reduce traffic speed by reducing the maximum throughput capacity available to cars. Speed reductions may be small, however, if the flow of cars decreases due to substitution to other roads or transportation modes such as bicycles or transit. Reduced flows following lane reductions would be consistent with the finding by Duranton and Turner (2011) that vehicle-miles traveled are proportional to highway lanes, although this is a long-run finding at the level of metropolitan statistical areas. Further, bikelanes may improve traffic if they isolate prexisting cycling from other vehicles.

[28]The TLC has issued two reports on the boro program, containing information on changes in the availability of taxi service outside the "Manhattan Core" as well as some assessment of the program's impact on traffic speed. On congestion, the first report (NYC-TLC, 2013) employs data through October 2013 and provides a table with year-on-year changes in average taxi trip speed for trips between 28 broad subdivisions of the city, using data from both yellow and boro taxis. The results are heterogeneous and inconclusive, and as of October 2013 only 1,128 boro taxis (17.3% of the full roll-out) had provided a trip. The assessment in the second report (NYC-TLC, 2015) consists of the following statement: *"As to congestion throughout all five boroughs outside of the Manhattan Core, average Boro Taxi trip speeds have remained steady over the last year, even as more Boro Taxis have entered into service. Comparing trips completed in January through June across the five boroughs, the average trip speed was 11.2 mph in 2014, compared to 11.3 mph in 2015."* This assessment is i) restricted to trips from Boro taxis only, ii) an average for most of the city, rather than between areas where the program's incidence may be expected to vary, and iii) uses January to June 2014 as a baseline for comparison, although 5,069 boro taxis were already on the road by June 2014 (3,818 as of January 2014), which is 77.5% (58.4% respectively) of the program's peak roll-out of 6,539 boro taxis in June 2015.

[29]Buchholz (2020) and Fréchette *et al.* (2019) evaluate deregulation counterfactuals that do not account for endogenous changes to traffic congestion and related costs.

[30]Medallion and driver identifiers, which allow a researcher to study taxi movements between trips, are only available for yellow taxis and through 2013, so they do not cover the period of boro taxis roll-out.

passengers to choose alternative transportation modes or not to begin travel at all from a given location. Our interest lies in the effect on congestion from the long term increase in supply, as might result from medallion deregulation or the unregulated entry of a close taxi substitute.

3.1 Measuring changes in the spatial distribution of vehicle density with aerial orthoimagery

We measure vehicle density in a straight-forward manner by constructing a new dataset on the observed locations of taxis and other vehicle over space and time. We acquire aerial orthoimagery for northeastern Manhattan captured over 29 dates, 11 before and 18 after August 2013. We manually digitize, i.e. tag, the observed location of 131,921 vehicles in this imagery.[31] Figure 5 shows the densities of both taxi and all vehicle locations on the north-east side of Manhattan before and after the boro program.[32]

Aerial imagery reveals how taxis reallocated around the hail-exclusion boundary following the introduction of the boro program: yellow taxi density decreased slightly throughout most of the kilometer south of the hail-exclusion boundary. At the boundary itself there has been a spike in supply from boro taxis, but this is almost exactly offset by a retrenchment from yellow taxis. The density of yellow taxis has remained largely unchanged north of the boundary, while the additional supply of boro taxis has led to increases in the taxi density from both types of taxi to average 64.6% over the two kilometers north of the boundary (and ranging from 0.0 to 291.5% growth). The change in density for all vehicles was naturally more muted: we measure increased total density both at the boundary and most areas to the north, but do measure slight declines in density at a couple locations.

[31] In our analysis, however, we only employ 26 scenes through June 2016, because the TLC has anonymized geographic coordinates since the second half of 2016. Sources and methods are detailed in Appendix C. We employ all the aerial orthoimagery available for New York City from all known sources (public or commercial) but one, the exception being a second commercial vendor. Examples of imagery and tagging are displayed in Figures C.5a and C.5b. We evaluated and disregarded satellite imagery of the highest resolution available (Digital Globe WV2 and WV3, Airbus Pleiades, Kompsat-3), which is unsuitable for vehicle roadbed counts due to resolution and visual obstruction from buildings. See Donaldson and Storeygard (2016) for a discussion of satellite imagery in economics research.

[32] We have overloaded the term "density" here, in the sense that Figure 5 plots kernel density estimates of observed vehicle locations along the centerlines of Manhattan's north-south avenues, but we rescale the units on the y-axis to indicate the density (in vehicles per km^2) that any segment length over the x-axis would mean-integrate to over the width of our area of interest, which is between 1st and 5th Avenues. Rescaling allows us to compare changes in vehicle density along a north-south axis before and after the introduction of the boro program.

3.2 Estimation of the elasticity of congestion

We begin by estimating the elasticity of congestion to vehicle density for vehicles of class v (taxis, or all vehicles) over the entire treatment zone ($\bar{\eta}_{c,v}$), i.e. we employ the boro roll-out as an instrument for the effect of vehicle density on travel times. The first stage of our empirical specification is:

$$\text{log vehicle density}_{rbm} = \pi_{DID}\,\text{roll-out}_m \times \text{boro zone}_{rb} + \pi_x x_{rbm} + \delta_{rb} + \delta_{rm} + \nu_{rbm} \quad (3)$$

and the equation for the second stage is:

$$s_{rbm} = \bar{\eta}_{c,v}\,\text{log vehicle density}_{rbm} + \gamma x_{rbm} + \delta_{rb} + \delta_{rm} + \varepsilon_{rbm} \quad (4)$$

where $\bar{\eta}_{c,v}$ is the congestion elasticity of vehicle density, and we require the exclusion restriction that, conditional on covariates, any effect from the boro taxi program's roll-out on traffic speed occurred through its effect on the number of vehicles on the road rather than any other channel. We estimate an effect separately for taxis and for all vehicles, and report OLS, first stage and 2SLS results in Table 7. We find average congestion elasticities to vehicle density of 0.141 (taxis) and 0.725 (all vehicles).[33]

We next define the elasticity as a function of location attributes l with a flexible polynomial $\eta_{c,v}(l) = p(l,\theta)$, and estimate:

$$s_{rbm} = \eta_{c,v}(l)\,\text{log vehicle density}_{rbm} + \gamma x_{rbm} + \delta_{rb} + \delta_{rm} + \varepsilon_{rbm} \quad (5)$$

while instrumenting this equation through the set of interactions defined by $p(l) \times \text{roll-out}_m \times \text{boro zone}_{rb}$. We use three measures to project the variation in congestion elasticity onto location attributes: population density, "point-of-interest" density, and transit density.[34] We model $p(l)$ as a cubic polynomial in each of the logs of population and place-of-interest density, a cross-interaction term between these, and a categorical variable indicating zero, low or high density of transit facilities on a particular block

[33]The latter value indicates a subtantial overall impact on travel time, but as an area average is still short of "hypercongestion" (see Geroliminis and Daganzo (2008) and Anderson and Davis, 2020) and is also lower than calibration parameters developed for highway congestion in the transporation engineering literature (e.g. the Bureau of Public Roads (BPR) model is calibrated at an elasticity of 4, see Parry and Small (2009) and Basso and Silva (2014)).

[34]Population counts at the block level are from the 2010 Decennial Census. We use the geographic coordinates of points-of-interest (POI), queried from the Google Places API, as a measure of density in commercial and urban amenities. POI include commercial establishments, along with schools, churches, museums, and other common destinations. We also obtain coordinates for transit facilities from the Google Places API. For each measure, we compute density by dividing counts according to the total land area of the intersection of Census Block polygons and the polygons for each block-run.

(defined from natural breaks in the underlying distribution). We report the outcome of estimating these projections on Figure 6. Each panel reports the congestion elasticity to all vehicle density (in purple) and taxi density (in orange) as a function of one density variable while evaluating the other at its sample mean. We find that congestion elasticities to taxi density increase in both urban density measures. All-vehicle elasticities flatten out at higher values of urban density, but in areas where both population and place-of-interest density are high their effect aggregates well beyond unity. Cross-interactions between the densities are positive, but not quantitatively significant. Not reported in the figure, we also find that, conditional on the population and place-of-interest density at a location, the effect of an increase in vehicle density on congestion is lower at locations with some on-block transit.

4 Application: the role of ridehail in the slowdown of NYC traffic

In this section we apply our estimates of the congestion elasticity of taxi supply to explain the substantial slowdown of New York City since 2013. First, we provide descriptive evidence on the depth, timing and extent of the traffic slowdown throughout the city's streets and highways since 2013. Next, we discuss the three existing studies that evaluate the effect of ridehail on traffic congestion in the city, and we complement our assessment of these studies with evidence on traffic speed and vehicle use. We then answer a simple accounting question: given our estimates for the congestion elasticity of vehicle density, could the city's slowdown be accounted for by the additional supply of for-hire vehicles that was enabled by ridehail applications (e.g. Uber, Lyft, Via, Juno and Gett)? To answer this question, we employ aerial imagery counts to estimate the growth in the supply of for-hire vehicles within an area of midtown Manhattan, as well as measure substitution from private vehicles. We find that most of the observed slowdown in midtown Manhattan can be explained by growth in the supply of for-hire vehicles.[35]

Figure 1 shows that the travel time of the median yellow taxi trip has increased

[35]To our knowledge, our study is the first to quantify growth in the supply of vehicles on the road due to ridehail, and to provided direct (i.e. non-survey) evidence on substitution from private vehicles. Uber has provided academics with data for the purposes of studying the value of the platform to consumers and workers, e.g. Cramer and Krueger (2016), Cohen *et al.* (2016), Hall and Krueger (2016), Angrist *et al.*, Chen *et al.* (2017), Castillo *et al.* (2017) and Hall *et al.* (2017). In particular, Castillo *et al.* (2017) illustrate how dynamic pricing can allow for increased density of Uber vehicles, increasing Uber's efficiency by reducing travel to (and time to) pickup. The paper does not address the impact of Uber on road use or speed. To date, Uber has not provided access to data for the study of congestion externalities or the degree to which ridehail complements or substitutes transit or the use of private vehicles. Outside of a data agreement, Ge *et al.* (2016) find evidence of racial discrimination on the platform.

substantially since 2013: after remaining flat at an average 5.1 minutes per mile from June 2009 to June 2013, from 2013 onward the median travel time started to rise and has exceeded 6 minutes per mile since the second half of 2016.[36] Median or average travel time over millions of monthly trips could be subject to selection from changing patterns in the use of taxis: congestion could possibly remain constant while passengers choose to take taxis on more congested roads, driving up the median travel time.

To condition on possible changes in travel patterns over space, the plots in Figure D.6.(a)-(c) show time series for selected percentiles of taxi travel time per mile during the month of June, with each line tracking a percentile for a particular "taxi-zone" pair within midtown Manhattan.[37] Percentiles of the travel time across most origin-destination pairs remained flat until June 2013 and have increased thereafter. Figure D.6.(d) summarizes the variation between June 2013 and June 2016: across 342 origin-destination cell-pairs within midtown Manhattan every percentile has shifted upward on average; the median and 10th percentile of travel time increased in 98.9% of all pairs, whereas the 90th percentile increased in 95.3% of pairs. The median travel time has increased by 15.2% on average across all origin-destination cell pairs in midtown. We describe congestion in midtown for the purposes of our accounting exercise but while midtown slowed down by more than average, it is not unusual: extending the previous comparisons to all taxi-zone pairs with more than 5 trips in both June 2013 and June 2016, we find that the median taxi trip slowed down by more than it did on average in midtown across 34.4% of the cell-pairs in the city. Figure D.7 reports on speed changes using a data source that is novel to our paper: pair-linked event data from a network of EZ-pass sensors throughout highways and expressways in the city. There is a gap in the data around 2014, but there has been an evident slowdown since 2013 throughout the entire distribution of highway speeds (fig. a), throughout the day (b), and throughout the city's network of highways (contrasting (c) and (d)).[38]

The fact that the speed of traffic in New York City has declined substantially in recent years has been reported previously in the media and through aggregate statistics in city

[36]The same pattern shows up in survey responses from the US Census' American Community Survey (ACS). Manhattan residents reported that travel time to work by car remained approximately constant between 2009 and 2013, but increased by 2.75 minutes (7.6%) from 2013 to 2016.

[37]"Taxi-zones" are polygonal cells defined by the Taxi and Limousine Commission, and are the finest spatial unit at which trip origination data is provided by ridehail apps. From July 2016 onward the TLC has also started anonymizing individual taxi trip records to the taxi-zone level. There are 19 taxi-zones in midtown Manhattan and 263 in all of the city.

[38]While the NYC Department of Transportation exposes the data to the internet as a real-time feed (see Data Appendix A), the DOT has claimed in correspondence with us not to have archived it. We have archived the data since 2015, and data from three months including June 2013 was kindly made available by Prof. Tomonari Masada of Nagasaki University, who archived it for unrelated research.

transportation reports.[39] However, there is no prior academic research on the timing and scale of the slowdown, or credible empirical evidence on its causes. In January 2016 the Office of the Mayor issued the ad-hoc "*For-Hire-Vehicle Transportation Study*" (NYC, 2016), which concluded that "*reductions in vehicular speeds are driven primarily by increased freight movement, construction activity, and population growth. (...) E-dispatch (...) did not drive the recent increase in congestion in the CBD*." The report fails to meet modern standards for identification and there is evidence that it was produced under strong industry lobbying.[40] Two follow-up reports (Schaller, 2017b and Schaller, 2017a) have cautioned that vehicle miles traveled in the city are on the rise due to ridehail, and that an increasing share of taxi and ridehail vehicle miles are driven empty.[41] As corroborating evidence, we estimate time series for VMTs per day, making explicit methodological decisions to address censoring in the data but remaining agnostic on the distribution of VMTs between traditional car services and ridehail (see Appendix F). We plot time series for estimated VMTs in Figures D.8d and D.8e. We find that the sum of taxi (yellow plus boro) and FHV miles increased by 886 million miles (34.6%) between 2013 and 2016. Our results are broadly consistent with Schaller (2017b), although that report may understate growth in miles through 2016 as well as ridehail's impact through 2015.[42]

[39]Examples include Fitzsimmons and Hu (2017) and Aaron (2015). NYC-DOT (2016) reports a 12% slowdown between 2010 and 2015 in the annual average of taxi trip speed in Manhattan below 60th Street, and a 20% slowdown in a "Midtown Core" area.

[40]During the summer of 2015, NYC Mayor Bill de Blasio suggested capping growth in for-hire vehicles. In the words of former Uber lobbyist Bradley Tusk, "*we were able to, through a really aggressive campaign, beat back completely (...) some limitations on Uber's growth. (...) We ran $4 million in TV spots. We did radio ads. We did direct mail. We had digital ads. We mobilized our customers (...) We had five different lobbying firms. We worked every single editorial board in the city. We worked the columnists and the pundits. It was just an all-in, 24-hour-a-day campaign*" (quoted from Mehta, 2016). The claim that ridehail did not cause congestion is made by incorrectly equating trip counts to vehicle supply: "*increases in e-dispatch trips are largely substituting for yellow taxi trips in the CBD. Because these e-dispatch trips are substitutions and not new trips, they are not increasing VMT (Vehicle Miles Traveled).*" Congestion is attributed to construction and pedestrians by appealing to correlation with city-wide growth rates in these activities since 2009, a baseline year affected by economic recession.

[41]Using odometer inspection data from for-hire vehicles and taxis, and under some explicit assumptions on private car use and substitution, Schaller (2017b) estimates that ridehail led to an additional 600 million vehicle-miles traveled in 2016 relative to 2013, a 27.1% growth over yellow taxi and FHV miles in 2013. Schaller's analysis requires unstated assumptions to handle the fact that vehicle mileages are censored at the last available inspection date, and to allocate flows of miles at the VIN level to stocks of ridehail vs. traditional car services.

[42]Following an initial freedom-of-information request we found that the odometer data that was provided to Schaller contained gaps. Consultation with the TLC allowed us to obtain additional inspection data. More substantively, Schaller deducts the decline in yellow taxi VMTs from his gross estimate of ridehail VMTs, under the assumption that all losses in taxi miles were substitution and not additional miles created by ridehail. However, the report does not track VMTs by boro taxis, which as Figure D.8d shows more than offset the declines in yellow taxi VMTs though the first half of 2016. Since Schaller (2017b) implicitly assumes no substitution from yellow to boro taxis, any degree of actual substitution would imply that the estimate of ridehail's growth is biased downward. The report also claims to be consistent with the finding in NYC (2016) that ridehail was not a major source of congestion through the first half of 2015. Unlike

In addition to documenting the extent and timing of speed changes and vehicle use through data from taxi meters, EZ-pass and odometer inspections, our contribution to the study of the causes of New York City's slowdown consists of i) measuring the growth in the supply of for-hire vehicles on the road, and ii) applying our estimate of the marginal congestion of taxi denisty to evaluate whether increases in vehicle density are consistent with the observed slowdown.

There are challenges to the use of aerial imagery to count supply on the road. Due to the cost of digitizing the imagery we restrict our analysis to an area on the east side of midtown Manhattan, delimited by 1st and 6th avenues and by 38th and 57th streets. We must also rely on a small sample of historical images captured before and after the introduction of ridehail. As in Section 3.1, we compare periods before and after August 2013: comparisons between these two periods are differences in means over a sample of 26 images, 11 before and 15 after August 2013 (details in Appendix C). A further methodological challenge to employing our taxi density elasticity estimate is that aerial images do not reveal whether observed cars are private or for-hire vehicles. We address this shortfall in the data by counting (in addition to taxis and trucks) the cars that have a black exterior color separately from cars of any color other than black. Counting by color was motivated by a prior that black is the most popular color among for-hire vehicles in New York City, as well as a growth in the share of black cars over successive aerial images that is evident to the naked eye. We then obtained current and historical snapshots of the vehicle registration data for every vehicle in the state from the New York Department of Motor Vehicles. The data contains exterior paint color and county of registration for almost all vehicle identification numbers (VINs). We calculate the share of vehicles with a black exterior paint color among those with an active NY state registration and registered as for-hire vehicles with the Taxi and Limousine Commission ($\alpha_t^{b,F}$). Equivalently, we calculate $\alpha_t^{b,P}$ for private cars (i.e. non-TLC registered) in NYC. Black-paint shares averaged 20.9% for private cars and 73.0% for FHVs in our period.[43] For each image we obtain a pair (c_i^b, c_i^{nb}) of car counts by color. If we are willing to assume that the paint color shares of cars counted in our midtown area of interest are proportional to color shares in the vehicle registration data for New York City, then we can derive an estimate of the otherwise unobserved counts for private and for-hire

the rest of the report, however, this claim is based on growth in trip counts rather than VMTs. Analyzing VMTs for the same period as NYC (2016), we find that VMTs from yellow taxis plus FHVs increased 8.4% between June 2013 and June 2015, and by a further 17.8% between June 2015 and June 2016 (15.0% and 24.0%, respectively, if VMTs from boro taxis are included).

[43] Vehicle counts likely include commuters from outside the city. We cannot trace the origin of the vehicles in our midtown counts to neighboring states and counties, and lack vehicle registrations by color for neighboring states. However, black-paint shares for private vehicles barely change (average 19.2%) if we include registrations from neighboring Nassau and Westchester counties in NY state.

vehicles $(\widehat{c}_i^P, \widehat{c}_i^F)$ in midtown by inverting

$$\begin{bmatrix} c_i^b \\ c_i^{nb} \end{bmatrix} = \begin{bmatrix} \alpha_t^{b,F} & \alpha_t^{b,P} \\ 1-\alpha_t^{b,F} & 1-\alpha_t^{b,P} \end{bmatrix} \times \begin{bmatrix} \widehat{c}_i^F \\ \widehat{c}_i^P \end{bmatrix} \tag{6}$$

for each image.

Average vehicle densities per km^2 are reported in Table 8, where we refer to counts for midtown Manhattan. We find that the number of vehicles per km^2 increased by 14.1% from the 11 images prior to the 15 images after August 2013. We underscore that due to the small size of our sample and the variance over images, the estimates are imprecise: although 14.1% growth is an economically significant magnitude, it is not statistically significant. We do find a statistically significant 11.8% decline in the number of yellow taxis. We also find a 42.5% increase in the number of cars, which breaks down into a 122.7% increase in black cars and a 9.3% increase in non-black cars. Using the share of car colors in the general population of NYC vehicles, we estimate that private cars on the road decreased by 6.8%, whereas the density of FHVs increased by 222.9%. The increase in FHVs from 112.5 to 363.3 per km^2 more than offset the drop in 62.5 observed taxis per km^2: the sum of taxis plus FHVs on the road in midtown has increased 30.5% in the average of aerial images observed over the past three years, relative to the four years earlier.

In Section 3 we estimated a congestion elasticity curve for both taxi and all-vehicle density as a function of two urban density measures. Projecting congestion elasticity estimates onto two density measures helps us extrapolate a parameter estimated in northeastern Manhattan to midtown: whereas midtown has among the highest "point-of-interest" densities in the city (average 16,937 POI per km^2 in our area of study), it also has among the lowest population densities (average 16,898 people per km^2). Fitting our estimated congestion externality curves to these figures we find that midtown has a congestion elasticity of 0.813 for all vehicles, and .358 for taxis. At these values, the overall 14.1% measured increase in all vehicles (which we can account is entirely due to ridehail, with countervailing reductions in taxi and private car densities) can account for an 11.5% slowdown in traffic in midtown Manhattan. Using our estimates of both the net increase in both taxis and FHVs (up 30.5%) and our congestion elasticity to taxi density we find that ridehail can account for a 10.9% slowdown. Both of these estimates, which are in close agreement, suggest that ridehail can account for between 71.7% and 75.6% of the slowdown that we documented for midtown Manhattan traffic between June 2013 and June 2016 (taking the 15.2% median slowdown in speed across midtown taxi zones

as a representative value). [44]

Our application does not establish a causal effect from ridehail on traffic speed. However, it is reasonable to find (given our causal estimates of the congestion elasticity of supply) that the largest change to road transportation in New York City over the past three years can account for most of the observed slowdown. Our count (and other) data also shed light on some of the alternative causes that have been suggested for the slowdown. We cannot measure the use of the road by couriers and delivery services but, as elsewhere in the city, we find that truck counts have decreased. We count 7.6 boro taxis per km^2 after August 2013, which relative to a total count of 1301.8 vehicles suggests that boro taxis' impact on midtown congestion has been negligible. We do not adjust our estimate of the change in vehicle supply for changes in private car use, since adaptation in private use is part of the reduced form impact in our congestion elasticity estimates. However, our aerial counts find evidence of modest substitution. Our estimates of this substitution are not statistically significant, although we emphasize that this is due to our small sample, rather than an insignificant economic magnitude.[45] Bike lanes are also unlikely to play any role in the traffic slowdown within the area of our application in midtown: the area has 46.1 km of roads and 6.9 km of bike lanes, but almost all these bike lanes were already in place by June 2013. The only bike lane expansion over the next three years was the conversion of 489 meters along 1st Avenue from shared to protected bike lanes. This expansion implies a 1.1% change in road exposure (absent a correction for spatial spillovers) to a differential slowdown of which our best estimate is 1.6 percentage points, as reported in Table 6. The Citibike bikeshare system was also already in place in midtown by June 2013.

It is also worth noting that the slowdown in highways and expressways that we document using EZ-pass data is inconsistent with some of the suggested alternative explanations for congestion. Bike lanes, cycling and pedestrians are almost surely unrelated to slower highways. Although courier and delivery services may have led to increased highway use, their impact would be relative to baseline utilization for delivery to brick-and-mortar stores. The channel that NYC (2016) suggests for the impact of delivery services, double-parking, is not relevant on highways either. We

[44]The presented estimates do not adjust for the impact of spatial spillovers, which would result in higher congestion elasticity estimates. On the other hand, basline estimates with the B-spline methodology for speed measurement would result in marginally lower impacts. Congestion impact estimates using speed measurement only at the times that correspond to the capture times for the aerial imagery do not significantly affect results (see Table G.6).

[45]There is almost no prior research on substitution patterns between ridehail and other modes. Our results are consistent with an internet survey of ridehail users in seven cities conducted by Clewlow and Mishra (2017), who report that "49% to 61% of ride-hailing trips would have not been made at all, or by walking, biking, or transit."

have not undertaken a study of highway construction projects in New York City, but the widespread spatial distribution of the highway slowdown seen from Figure D.7c to D.7d suggests that individual construction projects are an unlikely source. On the other hand, we cannot rule out that increases in traffic due to boro taxis, which are a negligible cause of congestion in midtown Manhattan, may have contributed to highway congestion in the outer boroughs.

5 Application: travel time externalities and congestion pricing

As a second application, we use our location-specific estimates of the elasticity of travel time to vehicle density to quantify the marginal damage of an additional vehicle due to congestion across all of Manhattan south of 125th St.[46]

We begin by focusing on commuters into Manhattan and estimating the social cost of driving an additional vehicle within a particular census tract c. We use the Census Transportation Planning Products (CTPP) to derive counts of workers, C_c, who live within 50 miles of Manhattan and commute into c. We then calculate the mean hourly wage, w_c, for the workers who commute to c. We compute the value of each minute in traffic, W_c, by taking half of the average hourly wage (Anderson, 2014;Small and Verhoef, 2007) and dividing by 60. Next, we calculate the number of private vehicles commuting into c, V_c, either driving alone or in a carpool. We allocate the impact to commuters based on their workplace tract, and calculate a percentage change in vehicles at each tract c associated with the addition of one private commuting vehicle, $\psi_c = \frac{V_c+1}{V_c}$. We estimate the average transit time within Manhattan, T_c for all commuters from their origination census tract to their workplace census tract, c, using data from the CTPP.

We calculate the social damage due to excess travel time from an additional vehicle commuting to census tract c as

$$MD_c = C_c\, W_c \left(\frac{V_c + 1}{V_c}\, \eta_c^A\, \bar{T}_c \right) \tag{7}$$

where the first two terms represent the total value of a minute for all commuters into c and the term in parentheses represents the additional trip time associated with an additional vehicle driving in c. Excess travel time is calculated by applying the percentage change in vehicles from our local elasticity estimate for all vehicles to travel time and lastly, multiplying by the average transit time within Manhattan.

[46]The literature on congestion pricing and its effect on travel and related benefits includes Tang (2020) on London, Gibson and Carnovale (2015) on Milan, and Karlström and Franklin (2009) on London. Kreindler (2020) and Martin and Thornton (2017) perform congestion pricing experiments.

In Figure 7, we show how these values vary over space by workplace census tract, c. Panel (a) shows the average hourly wage, w_c across Manhattan. Panel (b) shows our estimated elasticity of vehicles to travel time, η_c^A, by c.[47] Lastly, Panel (c) shows the marginal damage for each c associated with an increase of one commuter.

These estimates range from \$5.49 to \$45 over census tracts. To place these estimates in perspective, we find that the social cost (to the value of commuter time only) of driving anywhere in Manhattan south of 125th St is anywhere between twice and 16 times larger than the "Congestion Surcharge" of \$2.50 (for taxis) or \$2.75 (for ridehail) enacted by New York State in February 2019.

We next discuss a cordon congestion pricing policy approved by the state of New York in April 2019, which was scheduled to go into effect at the beginning of 2021. When the policy takes effect, private vehicles entering into Manhattan south of 60th St. will be charged a congestion fee at most once per day. Fares are yet to be determined and values in the range of \$10 to \$14 have been discussed in the media in order to reach targeted revenues of \$1 billion per year.[48]

We conduct a back-of-the-envelope welfare analyisis of the revenue gains, reduction in marginal damage, and deadweight loss associated with this congestion pricing policy. To calculate the reduction of marginal damage, we again use data on commuters from the CTPP limiting the sample to only those commuters with workplaces within the proposed congestion zone. For this exercise, we calculate MD_c where c is the full proposed congestion zone. In addition to commuters, we also account for taxi and for-hire vehicle trips that travel into the congestion zone. Using data from the CTPP, we estimate a daily inflow of 142,301 private vehicles transporting 188,235 commuters into the congestion zone. We estimate an average hourly count of taxi and for-hire vehicles in the congestion zone by computing the average hourly number of pickups within the congestion zone using trip record data from the NYC TLC. Across weekdays in May 2019, there was an average of 26,960 taxis and for-hire vehicles carrying 33,847 passengers in the congestion zone.[49] We compute the average annual salary for all workers who commute into the congestion zone from the CTPP to be \$73.96 and the average Manhattan travel time to be 35.5 minutes. Lastly, our estimated congestion elasticity with respect to all cars for the congestion zone is 0.913. Using these values, we arrive at a marginal damage value of

[47] To compute η_c^A for each c across all of Manhattan, we estimate POI density using a random sampling procedure. Due to computational costs, we are unable to use the Google Places API to extract all POIs for Manhattan. Instead, we randomly sample 40 points per census tract each with a 20 meter search radius and collect the total number of POIs in the random sample to compute POI density for each census tract.

[48] www.nytimes.com20190326nyregionwhat-is-congestion-pricing.html

[49] TLC data only includes passenger counts for taxis. We assume one passenger per vehicle for FHV, which understates marginal damages.

$26.20 for each commuter that enters congestion zone.

We model the demand for commutes with a simple iso-elastic demand function $q = Ap^{\gamma}$ with price elasticity γ. Using the total number of commuters and calibrating vehicle use values (p_0), we solve for welfare quantities under alternative fees and a couple limiting elasticity assumptions. In Table 9, Panel A, we fix γ at -0.16 which we believe to be a lower bound on the price elasticity of demand for vehicle trips on the extensive margin.[50] We evaluate congestion charges at $12, a midpoint among proposals, and our estimate of the Pigouvian tax at $26.20. We also calibrate the daily use price of a private vehicle to either $24.64, the daily ownership cost of a large sedan according to the American Automobile Association (AAA), or $54.64, the ownership cost plus the daily cost of a monthly parking pass for midtown. Under this elasticity assumption the reduction in the number of vehicles is small. A proposed fee of $12 reduces private car commuters by between 4,500 and 9,000, and raises revenue of around $1.6 million per day. Congestion damages are reduced by $116-230 thousand per day. The Pigouvian tax roughly doubles the reduction in vehicles, the revenue, and the surplus under either baseline price assumption. In Panel B we repeat the analysis calibrating the price elasticity at -0.3.[51] We find an 80% larger consumer surplus under the elastic commuter assumption, driven by comparably larger reductions in congestion damages. Summarizing results in terms of commute time, we find that the $12 fee proposal would save between 0.9 to 3.1 minutes over the range of our assumptions (relative to an average within-Manhattan commute time of 35.5 minutes), whereas the fee set at average marginal damage would save between 1.7 and 5.3 minutes on the average commute.

6 Conclusions

In this paper we evaluate the congestion cost of a particular taxi deregulation episode in New York City. We also quantify a long-run traffic congestion externality that was previously unmeasured and yet is of first order importance to the regulation of car-based transportation providers in the city. Additionally, we document that there has been a substantial slowdown in NYC's traffic since 2013, and find that the increase in supply due to ridehail can account for most of this slowdown. This is an important result both for New York City and for other cities throughout the world that are assessing the costs and benefits of ridehail and evaluating strategies for its effective regulation.

[50] The estimate comes from a pilot study by Puget Sound Regional Council (2008) and is representative for commuters with high alternative transit options in the sample.

[51] A joint city agency study (NYC-DOT and NYC-TLC, 2019) assumes this elasticity for the own-fare demand elasticity for ridehail, which we expect is larger than the commuter elasticity.)

To execute this study we collect and analyze a large number of sources of urban data, of which some are "big", some have not been collected or analyzed previously, and some are entirely new. The City of New York is a leading example in the practice of making urban data publicly available. However, these efforts are at an early stage: standards have not fully developed, and are often geared toward mobile application developers rather than researchers. Some data remains siloed, and some valuable data is lost when it is deemed ephemeral by internal government users and is not archived. Additionally, changes in technology have led to private firms generating and owning data on their utilization of public goods and congestible resources. Regulators may want to develop frameworks so that businesses that are built on the use of public infrastructure report data on this use. The availability of urban data will increase in the near future, along with the use of internet-connected devices and other urban sensors. The use of methods from empirical economics in the analysis of such data can be expected to increase, motivated by changes to the economic activities that take place within cities (in transportation, these may include ridehail, bike lanes, bikeshare, dynamic tolling and autonomous vehicles). In the future, in-vehicle GPS may allow for first-best remedies such as road-specific congestion pricing – as long as regulators can quantify and price the relevant externalities.

References

Aaron, K. (2015). If Not Uber, Who's Responsible for This Traffic? *WNYC Broadcast. August 11, 2015*.

Ahlfeldt, G. M., Redding, S. J., Sturm, D. M. and Wolf, N. (2015). The Economics of Density: Evidence From the Berlin Wall. *Econometrica*, **83** (6), 2127–2189.

Akbar, P. and Duranton, G. (2017). Measuring the Cost of Congestion in Highly Congested City: Bogotá.

Akbar, P. A., Couture, V., Duranton, G. and Storeygard, A. (2018). *Mobility and Congestion in Urban India*. Tech. Rep. Working Paper 25218, National Bureau of Economic Research.

Anderson, M. L. (2014). Subways, Strikes, and Slowdowns: The Impacts of Public Transit on Traffic Congestion †. *American Economic Review*, **104** (9), 2763–2796.

— and Davis, L. W. (2020). An empirical test of hypercongestion in highway bottlenecks. *Journal of Public Economics*, **187**, 104197.

—, Lu, F., Zhang, Y., Yang, J. and Qin, P. (2016). Superstitions, street traffic, and subjective well-being. *Journal of Public Economics*, **142** (Supplement C), 1–10.

Angrist, J. D., Caldwell, S. and Hall, J. V. (). Uber vs. Taxi: A Driver's Eye View. *American Economic Journal: Applied Economics*, **Forthcoming**.

Barth, M. and Boriboonsomsin, K. (2008). Real-World Carbon Dioxide Impacts of Traffic Congestion. *Transportation Research Record: Journal of the Transportation Research Board*, **2058**, 163–171.

Basso, L. J. and Silva, H. E. (2014). Efficiency and Substitutability of Transit Subsidies and Other Urban Transport Policies. *American Economic Journal: Economic Policy*, **6** (4), 1–33.

Buchholz, N. (2020). Spatial Equilibrium, Search Frictions and Dynamic Efficiency in the Taxi Industry.

Castillo, J. C., Knoepfle, D. T. and Weyl, E. G. (2017). *Surge Pricing Solves the Wild Goose Chase*. SSRN Scholarly Paper ID 2890666, Social Science Research Network, Rochester, NY.

Chen, M. K., Chevalier, J. A., Rossi, P. E. and Oehlsen, E. (2017). *The Value of Flexible Work: Evidence from Uber Drivers*. Working Paper 23296, National Bureau of Economic Research.

Clewlow, R. R. and Mishra, G. S. (2017). *Disruptive Transportation: The Adoption, Utilization, and Impacts of Ride-Hailing in the United States*. Research Report UCD-ITS-RR-17-07, University of California, Davis.

Cohen, P., Hahn, R., Hall, J., Levitt, S. and Metcalfe, R. (2016). *Using Big Data to Estimate Consumer Surplus: The Case of Uber*. Working Paper 22627, National Bureau of Economic Research.

Cramer, J. and Krueger, A. B. (2016). Disruptive Change in the Taxi Business: The Case of Uber. *American Economic Review: Papers & Proceedings*, **106** (5), 177–182.

Currie, J. and Walker, R. (2011). Traffic Congestion and Infant Health: Evidence from E-ZPass. *American Economic Journal: Applied Economics*, **3** (1), 65–90.

Donaldson, D. and Storeygard, A. (2016). The View from Above: Applications of Satellite Data in Economics. *Journal of Economic Perspectives*, **30** (4), 171–198.

Duranton, G. and Turner, M. A. (2011). The Fundamental Law of Road Congestion: Evidence from US Cities. *American Economic Review*, **101** (6), 2616–2652.

Fitzsimmons, E. G. and Hu, W. (2017). The Downside of Ride-Hailing: More New York City Gridlock. *The New York Times. March 6, 2017*.

Fréchette, G. R., Lizzeri, A. and Salz, T. (2019). Frictions in a Competitive, Regulated Market: Evidence from Taxis. *American Economic Review*, **109** (8), 2954–2992.

Gao Hodges, G. R. (2012). *Taxi!: A Social History of the New York City Cabdriver*. Baltimore: NYU Press.

Ge, Y., Knittel, C. R., MacKenzie, D. and Zoepf, S. (2016). *Racial and Gender Discrimination in Transportation Network Companies*. Working Paper 22776, National Bureau of Economic Research.

Geroliminis, N. and Daganzo, C. F. (2008). Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B: Methodological*, **42** (9), 759–770.

Gibbons, S., Overman, H. G. and Patacchini, E. (2015). Chapter 3 - Spatial Methods. In J. V. H. a. W. C. S. Gilles Duranton (ed.), *Handbook of Regional and Urban Economics*, *Handbook of Regional and Urban Economics*, vol. 5, Elsevier, pp. 115–168.

Gibson, M. and Carnovale, M. (2015). The effects of road pricing on driver behavior and air pollution. *Journal of Urban Economics*, **89**, 62–73.

Glaeser, E. L., Kominers, S. D., Luca, M. and Naik, N. (2015). *Big Data and Big Cities: The Promises and Limitations of Improved Measures of Urban Life*. Working Paper 21778, National Bureau of Economic Research.

Greenshields, B. D. (1948). The Potential Use of Aerial Photographs in Traffic Analysis. In *Proceedings of the Highway Research Board*, vol. 27, pp. 291–297.

Hall, J. V., Horton, J. J. and Knoepfle, D. T. (2017). Labor Market Equilibration: Evidence from Uber.

— and Krueger, A. B. (2016). *An Analysis of the Labor Market for Uber's Driver-Partners in the United States*. Working Paper 22843, National Bureau of Economic Research.

Hamilton, T. L. and Wichman, C. J. (2018). Bicycle infrastructure and traffic congestion: Evidence from DC's Capital Bikeshare. *Journal of Environmental Economics and Management*, **87**, 72–93.

Hanna, R., Kreindler, G. and Olken, B. A. (2017). Citywide effects of high-occupancy vehicle restrictions: Evidence from "three-in-one" in Jakarta. *Science*, **357** (6346), 89–93.

Hsieh, C.-T. and Moretti, E. (2017). Housing Constraints and Spatial Misallocation.

Johnson, A. N. (1928). Maryland Aerial Survey of Highway Traffic Between Baltimore and Washington. In *Proceedings of the Highway Research Board*, vol. 8, pp. 106–115.

Kahneman, D. and Krueger, A. B. (2006). Developments in the measurement of subjective well-being. *The journal of economic perspectives*, **20** (1), 3–24.

Karlström, A. and Franklin, J. P. (2009). Behavioral adjustments and equity effects of congestion pricing: Analysis of morning commutes during the Stockholm Trial. *Transportation Research Part A: Policy and Practice*, **43** (3), 283–296.

Knittel, C. R., Miller, D. L. and Sanders, N. J. (2016). Caution, Drivers! Children Present: Traffic, Pollution, and Infant Health. *The Review of Economics and Statistics*, **98** (2), 350–366.

Kreindler, G. E. (2020). Peak-Hour Road Congestion Pricing: Experimental Evidence and Equilibrium Implications.

Li, S., Purevjav, A.-O. and Yang, J. (2017). *The Marginal Cost of Traffic Congestion and Road Pricing: Evidence from a Natural Experiment in Beijing*. SSRN Scholarly Paper ID 2948619, Social Science Research Network, Rochester, NY.

Liang, J., Gong, J., Sun, J., Zhou, J., Li, W., Li, Y., Liu, J. and Shen, S. (2017). Automatic Sky View Factor Estimation from Street View Photographs—A Big Data Approach. *Remote Sensing*, **9** (5), 411.

Martin, L. A. and Thornton, S. (2017). *Can Road Charges Alleviate Congestion?* SSRN Scholarly Paper ID 3055522, Social Science Research Network, Rochester, NY.

Mehta, S. (2016). Meet Uber's Political Genius. *The Vanity Fair Interview. June 17, 2016*.

Mundhenk, T. N., Konjevod, G., Sakla, W. A. and Boakye, K. (2016). A Large Contextual Dataset for Classification, Detection and Counting of Cars with Deep Learning. In *Computer Vision – ECCV 2016*, Lecture Notes in Computer Science, Springer, pp. 785–800.

NYC (2016). For-Hire Vehicle Transportation Study.

NYC-DOT (2014). Presentation on "Protected Bicycle Lanes in NYC". Available at nyc.gov/html/dot/downloads/pdf/2014-09-03-bicycle-path-data-analysis.pdf.

— (2016). Mobility Report. October 2016.

NYC-DOT and NYC-TLC (2019). *Improving Efficiency and Managing Growth in New York's For-Hire Vehicle Sector*. Tech. rep.

NYC-TLC (2013). Hail Market Analysis, available at www.nyc.gov/html/tlc/downloads/pdf/boro_taxi_market_study.pdf.

— (2015). Hail Market Analysis, available at www.nyc.gov/html/tlc/downloads/pdf/hail_market_analysis_2015.pdf.

Parry, I. W. H. and Small, K. A. (2009). Should Urban Transit Subsidies Be Reduced? *American Economic Review*, **99** (3), 700–724.

Pinkse, J. and Slade, M. E. (2010). The Future of Spatial Econometrics. *Journal of Regional Science*, **50** (1), 103–117.

Poco, J., Doraiswamy, H., Vo, H. T., Comba, J. L. D., Freire, J. and Silva, C. T. (2015). Exploring Traffic Dynamics in Urban Environments Using Vector-Valued Functions. *Computer Graphics Forum*, **34** (3), 161–170.

Puget Sound Regional Council (2008). *Traffic Choices Study: Summary Report*. Tech. rep.

Razakarivony, S. and Jurie, F. (2016). Vehicle detection in aerial imagery: A small target detection benchmark. *Journal of Visual Communication and Image Representation*, **34**, 187–203.

Redding, S. J. and Turner, M. A. (2015). Transportation Costs and the Spatial Organization of Economic Activity. In G. Duranton, J. V. Henderson and W. C. Strange (eds.), *Handbook of Regional and Urban Economics*, *Handbook of Regional and Urban Economics*, vol. 5, Elsevier, pp. 1339–1398.

Reilly, V., Idrees, H. and Shah, M. (2010). Detection and Tracking of Large Number of Targets in Wide Area Surveillance. In *Computer Vision – ECCV 2010*, Lecture Notes in Computer Science, Springer, pp. 186–199.

Sadik-Khan, J. and Solomonow, S. (2016). *Streetfight: Handbook for an Urban Revolution*. New York, New York: Viking, 1st edn.

Santi, P., Resta, G., Szell, M., Sobolevsky, S., Strogatz, S. H. and Ratti, C. (2014). Quantifying the benefits of vehicle pooling with shareability networks. *Proceedings of the National Academy of Sciences*, **111** (37), 13290–13294.

Schaller, B. (2017a). *Empty Seats, Full Streets: Fixing Manhattan's Traffic Problem*. Tech. rep., Schaller Consulting.

— (2017b). *Unsustainable? The Growth of App-Based Ride Services and Traffic, Travel and the Future of New York City*. Tech. rep., Schaller Consulting.

Schwartz, S. I. (2015). *Street Smart: The Rise of Cities and the Fall of Cars*. New York, New York: PublicAffairs, 1st edn.

Simeonova, E., Currie, J., Nilsson, P. and Walker, R. (2018). *Congestion Pricing, Air Pollution and Children's Health*. Working Paper 24410, National Bureau of Economic Research.

Small, K. A. and Verhoef, E. T. (2007). *The economics of urban transportation*. New York: Routledge.

Tang, C. K. (2020). The Cost of Traffic: Evidence from the London Congestion Charge. *Journal of Urban Economics*, p. 103302.

Xie, L. and Olszewski, P. (2011). Modelling the effects of road pricing on traffic using ERP traffic data. *Transportation Research Part A: Policy and Practice*, **45** (6), 512–522.

Zhan, X., Hasan, S., Ukkusuri, S. V. and Kamga, C. (2013). Urban link travel time estimation using large-scale taxi data with partial information. *Transportation Research Part C: Emerging Technologies*, **33**, 37–49.

Tables and figures

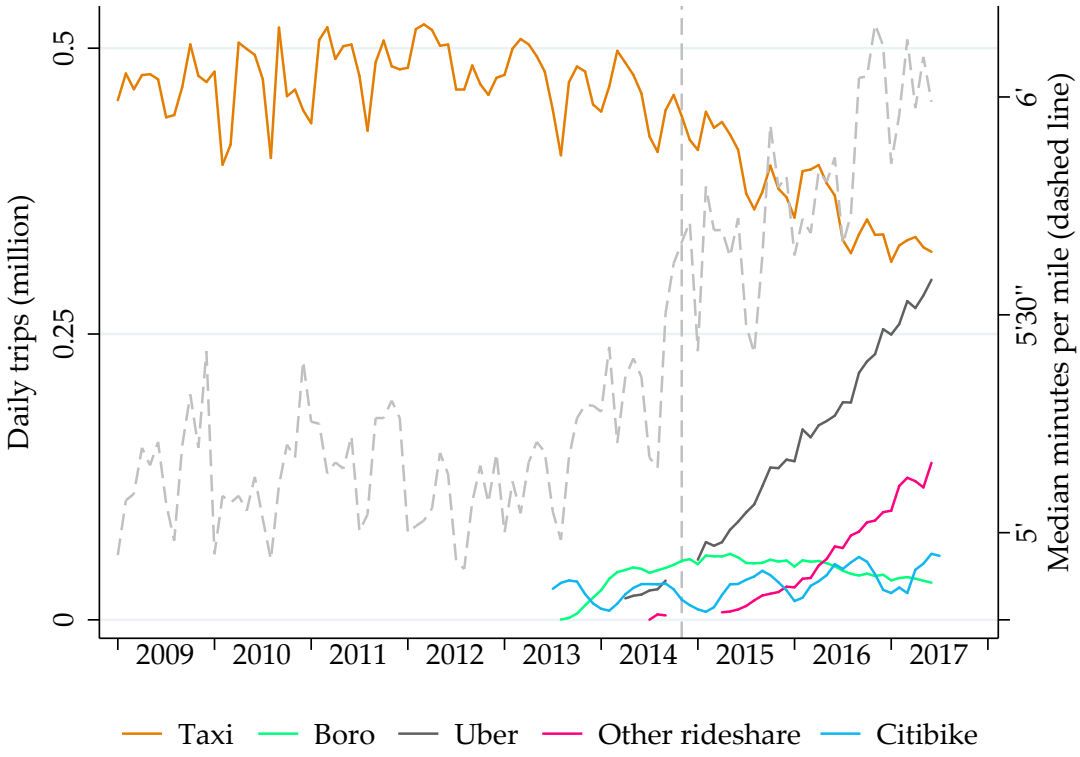


Figure 1: Daily trips in New York City by selected transportation modes, and yellow taxi median minutes per mile. Monthly statistics 2009-2017H1

Note: Trips from ridehail providers such as Uber and Lyft are available only for Q2 and Q3 in 2014, and from 2015 onwards. Dashed gray line plotted on second axis is the median minutes per mile among all yellow taxi trips in a given month. Vertical dashed line represents the Nov 7, 2014 date on which speed limits were lowered to 25 mph on most city streets under the "Vision Zero" traffic safety initiative.

Table 1: Descriptive statistics

	Mean	Median	SD	Min	Max
Seconds per meter (Avg.)	0.195	0.189	0.075	0.0924	3.103
Seconds per meter (B-spline)	0.127	0.101	0.271	0.0250	5
Seconds per meter (Projected)	0.127	0.123	0.054	0.0250	0.768
Taxi pickups	1873.965	961	2489.660	0	23636
311 complaints	0.024	0	0.188	0	7
311 complaint: blocked driveway (no access)	0.005	0	0.093	0	6
311 complaint: blocked driveway (partial access)	0.001	0	0.036	0	2
311 complaint: blocked roadway (construction)	0.001	0	0.027	0	2
311 complaint: failed roadway repair	0.003	0	0.060	0	5
311 complaint: pothole	0.014	0	0.134	0	5
311 complaint: rough roadway	0.001	0	0.024	0	2
Bikelane (any)	0.089	0	0.285	0	1
Bikelane: wide parking	0.026	0	0.159	0	1
Bikelane: protected	0.052	0	0.223	0	1
Bikelane: standard	0.004	0	0.065	0	1
Bikelane: signed route	0.008	0	0.089	0	1
Citibike station	0.002	0	0.056	0	2
Pothole outstanding	0.040	0	0.216	0	4
Zoning: EC-2/EC-3 districts	0.070	0	0.254	0	1
Traffic signal re-timing	0.041	0	0.199	0	1
Arterial slow zone	0.060	0	0.238	0	1
N/S speed bump	0.005	0	0.073	0	1
E/W speed bump	0.082	0	0.274	0	1

Note: Each observation is a run-block-month. Taxi pickups includes all pickups within a run-block geography, not only within-avenue trips. Bikelane (any) takes a value of 2 at three bikelane intersections. Seconds-per-meter values for both B-spline and Projected methods are bottom coded at 0.025 to prevent negative values and unusually large speed estimates (in excess of 90mph).

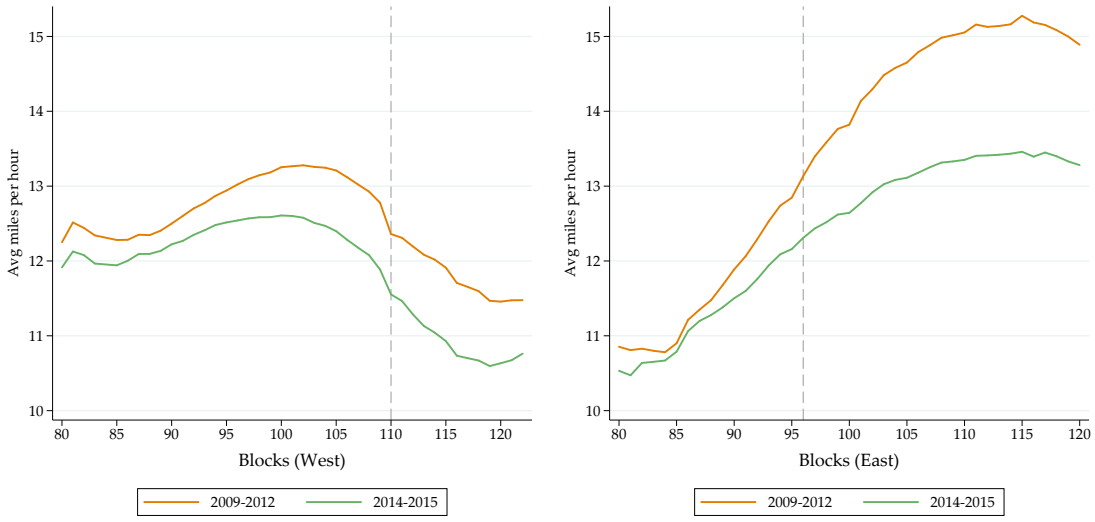
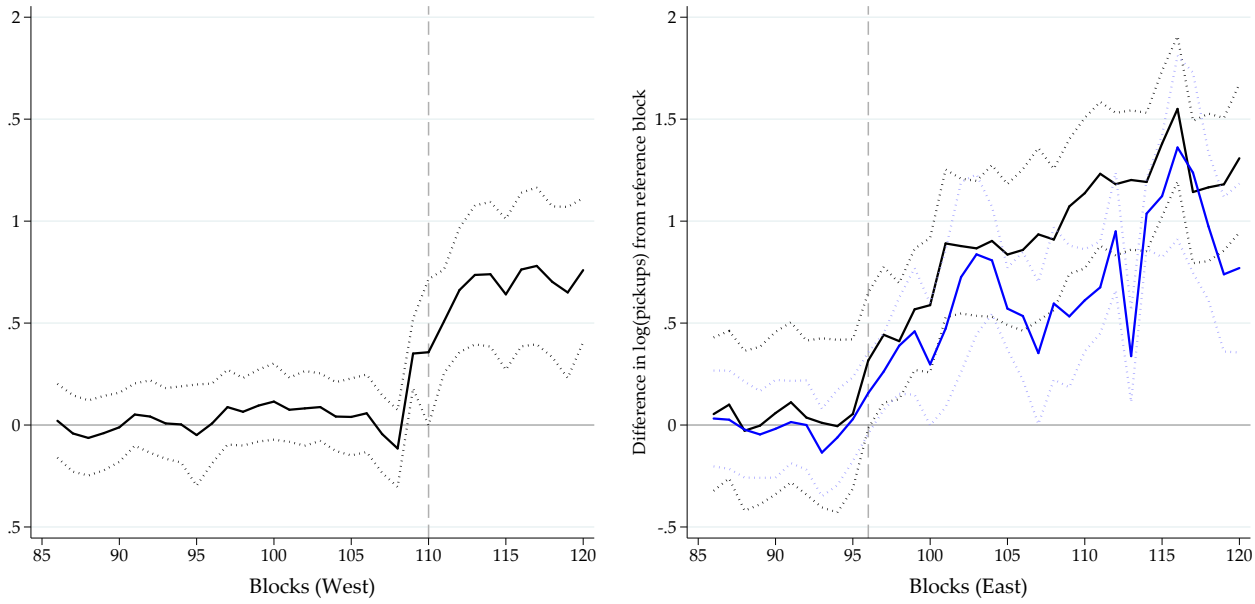
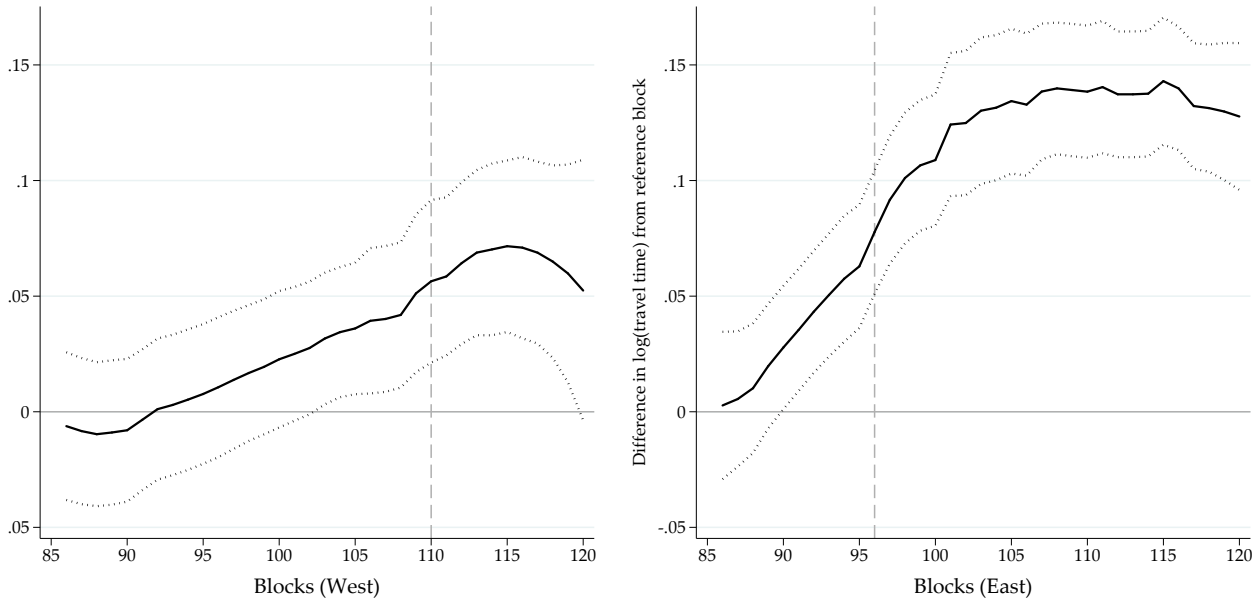


Figure 2: Average taxi speed

Note: Displays average travel time per unit of distance for all taxi trips traversing each city block (methodology detailed in Appendix B.1), for the subset of avenues where traffic flows across the boro program hail-exclusion boundary. Monthly panel data are averaged over both avenues on each side of Manhattan (left vs. right panel) and over all years before (orange) and after (green) the launch of the Boro taxi program. Data is excluded for 2013 (the program was implemented in August 2013) and 2016 (speed data is only available through June). Appendix Figures B.2 and B.3 display average travel times by year, individual avenue, and at a finer spatial resolution.



(a) Taxi pickups (West) and pickups and aerial counts (East)



(b) Travel time (avg. s/m)

Figure 3: Difference-in-differences estimates of the impacts of the boro program on speed, taxi pickups, and aerial counts(for East side only) by city block

Notes: We display difference-in-differences coefficients separately for the subsamples of West and East northern Manhattan. Dependent variables are the log taxi pickups in figure (a), and log travel time in figure (b), and for these an observation is defined by a city-block, run and month. Panel (e) additionally reports (in blue) on specifications where the dependent variable is log of aerial taxi counts, and where an observation is a city-block, run, and post August 2013 indicator. Aerial counts are smoothed across neighboring blocks with a triangular moving-average kernel, and further details on this data are in Appendix C. Confidence intervals are 95% point-wise for standard errors clustered at the run-block and run-month level (run-post for the aerial data). 84th and 85th St are the baseline omitted category in all regressions, and a separate coefficient is estimated for every in-sample block north of these. Vertical dashed line indicates the boundary of the "hail exclusion zone" for Boro taxis (96th St. E, 110th St. W).

Table 2: Reduced form impact of boro program on congestion.
Comparison across travel time estimation methodologies

	(1) Log avg. s/m (Treat dummy)	(2) Log avg. s/m (Baseline)	(3) Log Bsp. s/m (Treat dummy)	(4) Log Bsp. s/m (Baseline)	(5) Log est. s/m (Treat dummy)	(6) Log est. s/m (Baseline)
BZ x Post	0.061		0.049		0.062	
	(0.005)		(0.019)		(0.037)	
BZ x Roll-out		0.078		0.071		0.083
		(0.005)		(0.024)		(0.045)
Bikelane	0.024	0.024	-0.004	-0.006	0.084	0.084
	(0.005)	(0.005)	(0.048)	(0.048)	(0.036)	(0.036)
Citibike station	-0.015	-0.010	0.025	0.030	-0.085	-0.071
	(0.005)	(0.004)	(0.021)	(0.021)	(0.012)	(0.015)
311 complaints	0.001	0.000	-0.006	-0.006	-0.003	-0.003
	(0.001)	(0.001)	(0.010)	(0.010)	(0.003)	(0.003)
Potholes	-0.001	-0.001	0.001	0.002	-0.008	-0.008
	(0.001)	(0.001)	(0.007)	(0.007)	(0.008)	(0.008)
Run-block × Run-my FE	Y	Y	Y	Y	Y	Y
Observations	47430	47430	47430	47430	6849	6849
R^2	0.97	0.97	0.69	0.69	0.53	0.53
R^2-within	0.12	0.14	0.0012	0.0018	0.0052	0.0058

Note: Unless otherwise specified, robust standard errors are two-way clustered at the level of run× block and run× month. Each observation is a run-block-month. The difference in difference coefficient in the baseline regression (Col. 2, row 2) remains significant at 0.1% with coarser clustering schemes, such as two-way clustering at the level of runs and months (s.e. 0.0149), along any single dimension, or under alternative schemes (and bandwidths) with clustering on the time or time-run dimensions with Newey-West robustness to autocorrelation along the spatial dimension by block or run-block. Avg. speed are derived by pooling data from yellow and green taxis. We evaluate speed differences by taxi class in Table B.2 and find speed differences in some specifications. Speeding up boro taxi trips by 2.5% reduces the Col. 2 headline coefficient from 0.078 to 0.073. Further evaluation of inference is provided in Table 3, Cols 5 & 6.

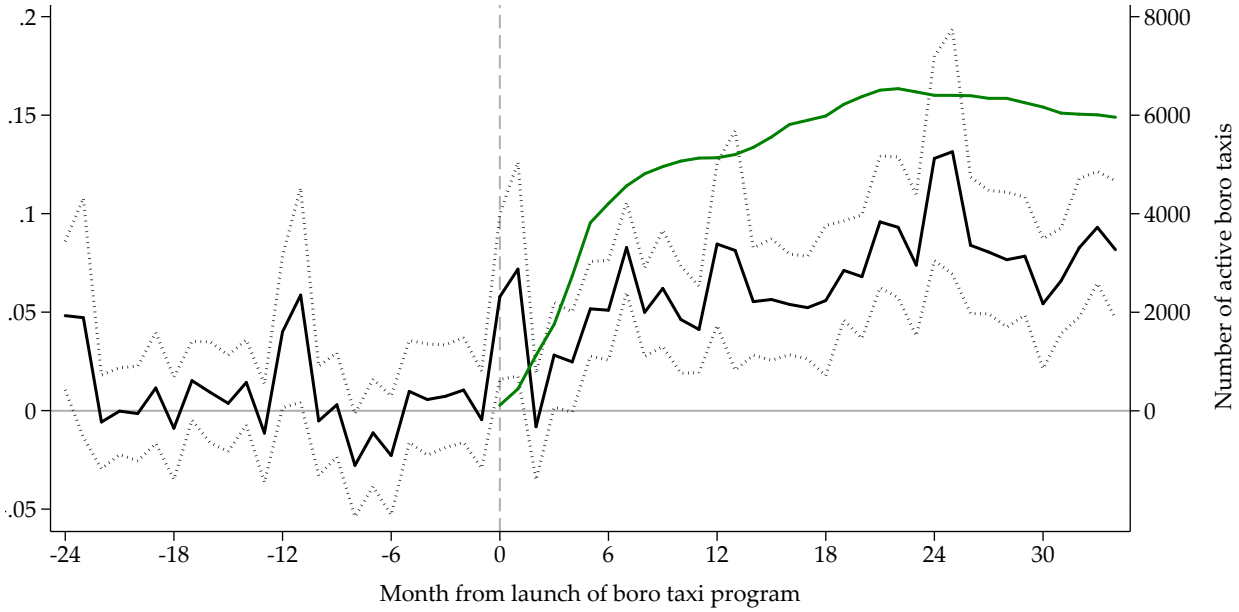


Figure 4: Difference-in-differences estimates of the impact of the boro program, by month

Note: In green, the number of unique boro taxis active every month, scale on the right axis. In blue, difference-in-differences coefficients and 95% point-wise confidence intervals for standard errors clustered at the run-block level. Each observation is a run-block-month. The baseline period is defined as the 31 months from Jan 2009 to Jul 2011, thus ending two years prior to the launch of the boro program. A separate coefficient is estimated for every month thereafter. Months are labelled on the plot relative to August 2013. Coefficient estimates are from Eq. (2), and include a fixed effect at the level of the run, block and calendar month to absorb block-specific seasonality.

Table 3: Reduced form impact of boro program on congestion.
Alternative specifications

	(1) Avg. s/m (Levels)	(2) Log avg. s/m (Rect. FE)	(3) Log avg. s/m (Linear trend)	(4) Log avg. s/m (Lin. & rideshare trend)	(5) Log avg. s/m (Aggr. T)	(6) Log avg. s/m (Aggr. T,Run)
BZ x Roll-out	0.014	0.059	0.071	0.066		
	(0.001)	(0.007)	(0.008)	(0.008)		
Bikelane	0.003	0.027	0.022	0.020	0.041	
	(0.001)	(0.007)	(0.004)	(0.004)	(0.025)	
Citibike station	-0.002	-0.015	-0.014	-0.008	0.011	
	(0.001)	(0.011)	(0.003)	(0.003)	(0.010)	
311 complaints	-0.000	0.002	0.001	0.001	0.009	
	(0.000)	(0.002)	(0.001)	(0.001)	(0.003)	
Potholes	-0.000	0.000	-0.000	-0.000	-0.002	
	(0.000)	(0.002)	(0.001)	(0.001)	(0.002)	
BZ x Post					0.047	0.046
					(0.016)	(0.017)
Post					0.051	0.059
					(0.011)	(0.012)
BZ dummy						-0.097
						(0.046)
Run-block × Run-my FE	Y		Y	Y		
Run-block × my FE		Y				
Run × BZ linear trend			Y			
Run × BZ rideshare trend			Y	Y		
Observations	47430	47430	47430	47430	47430	47430
R^2	0.96	0.93	0.97	0.97	0.84	0.70
R^2-within	0.090	0.049	0.19	0.21	0.19	0.17

Note: Unless otherwise stated, robust standard errors are two-way clustered at the level of run× block and run× month. Each observation is a run-block-month. Estimates with linear and "aggregate ridehail" trends are interacted with the treatment zone and run, but results barely change if interacted at the treatment level only. Estimates in Col. 5 include fixed effects and are clustered at the level of run× block and run× post, likewise by run× boro-zone and run× post for estimates in Col. 6. We estimate a main coefficient of 0.079 when dropping 2nd Ave from the sample, which we check to address concerns related to subway construction both north and south of the boundary.

Table 4: Reduced form impact of boro program on congestion.
Alternative samples

	(1) Log avg. s/m (6m pre/post)	(2) Log avg. s/m (12m pre/post)	(3) Log avg. s/m (East)	(4) Log avg. s/m (West)	(5) Log avg. s/m (cutout 3 blocks)	(6) Log avg. s/m (cutout 6 blocks)	(7) Log avg. s/m (cutout 9 blocks)
BZ x Roll-out	0.048	0.081	0.099	0.047	0.083	0.086	0.086
	(0.013)	(0.008)	(0.007)	(0.009)	(0.006)	(0.006)	(0.008)
Bikelane	0.015	0.040	0.017	0.023	0.025	0.025	0.025
	(0.015)	(0.010)	(0.006)	(0.008)	(0.005)	(0.007)	(0.008)
Citibike station	0.000	0.000	-0.011	-0.005	-0.008	-0.005	-0.005
	(0.000)	(0.000)	(0.002)	(0.005)	(0.004)	(0.004)	(0.006)
311 complaints	0.004	0.005	0.001	0.000	0.001	0.001	0.002
	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Potholes	-0.002	-0.001	-0.003	0.001	-0.001	-0.001	-0.001
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Run-block × Run-my FE	Y	Y	Y	Y	Y	Y	Y
BZ-my FE							
Observations	6851	13175	24660	22770	39510	30870	22680
R^2	0.98	0.98	0.97	0.97	0.97	0.97	0.97
R^2-within	0.036	0.15	0.21	0.055	0.15	0.14	0.12

Note: Unless otherwise specified, robust standard errors are two-way clustered at the level of run× block and run× month. Each observation is a run-block-month. Standard errors for the triple-difference estimation in Col. 8 are additionally clustered at the level of boro-zone × month.

Table 5: Reduced form impact of boro program on congestion.
Spatial autoregressive models for spillovers across the hail-exclusion boundary

	(1) Log avg. s/m ($\mathbf{W_a}$: three adjacencies)		(2) Log avg. s/m ($\mathbf{W_b}$: backward adjacency)	
ρ_a	0.800	(0.053)		
ρ_b			0.810	(0.037)
BZ x Roll-out	0.023	(0.005)	0.011	(0.003)
Standard controls	Y		Y	
Run-bin $\times$ Run-my FE	Y		Y	
Avg. direct plus indirect effect within BZ	0.1057	[0.0943,0.1210]	0.0482	[0.0432,0.0522]
Avg. direct effect within BZ	0.0265	[0.0236,0.0300]	0.0116	[0.0094,0.0136]
Avg. indirect effect within BZ	0.0791	[0.0670,0.0943]	0.0366	[0.0336,0.0389]
Avg. indirect effect (spillover) on HEZ	0.0108	[0.0078,0.0150]	0.0122	[0.0109,0.0137]
Avg. indirect effect (spillover) on HEZ, first block	0.0350	[0.0279,0.0442]	0.0246	[0.0230,0.0261]
Avg. indirect effect (spillover) on HEZ, first 2 blocks	0.0284	[0.0221,0.0369]	0.0226	[0.0221,0.0242]
Avg. indirect effect (spillover) on HEZ, first 3 blocks	0.0239	[0.0181,0.0317]	0.0209	[0.0194,0.0226]
Avg. indirect effect (spillover) on HEZ, first 6 blocks	0.0157	[0.0112,0.0217]	0.0164	[0.0148,0.0182]
Avg. indirect effect (spillover) on HEZ, first 9 blocks	0.0134	[0.0094,0.0191]	0.0148	[0.0131,0.0166]
Observations	34,470		34,470	
R^2	0.981		0.981	
R^2-within	0.179		0.180	

Note: Standard errors in parentheses on the reduced-form 2SLS coefficients of the spatial models are robust, two-way clustered at the level of run$\times$ block and run$\times$ month, and are displayed on the right of the estimated coefficient. F-statistics from the first stage are 32.1 and 61.9, respectively. For average structural effects (direct and indirect) we display 95% confidence intervals from bootstrapping the structural parameters over 1000 resamples, re-estimating the model and the impact matrix $(\mathbf{I} - \hat{\rho}\mathbf{W})^{-1}$. Both specifications include controls for bikelane, citibike, 311 complaints and potholes. Each observation is a run-block-month.

Table 6: Reduced form impact of boro program on congestion.
Alternative controls

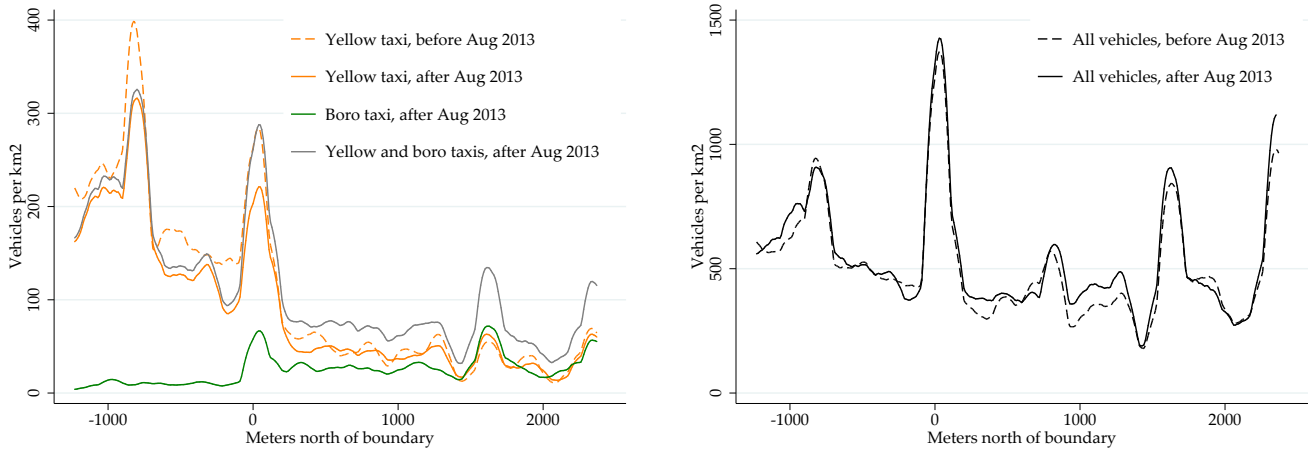
	(1) Log avg. s/m (No controls)	(2) Log avg. s/m (Road controls)	(3) Log avg. s/m (Comm. zoning)	(4) Log avg. s/m (Policy Changes)	(5) Log avg. s/m (All)
BZ x Roll-out	0.080	0.078	0.077	0.083	0.078
	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)
Signed route		-0.006			-0.006
		(0.014)			(0.014)
Wide parking		0.021			0.028
		(0.007)			(0.010)
Standard		0.017			0.017
		(0.006)			(0.006)
Protected		0.032			0.032
		(0.007)			(0.007)
Citibike station		-0.009			-0.008
		(0.004)			(0.004)
Citibike (south of station)		-0.011			-0.012
		(0.011)			(0.010)
311 complaint:					
Blocked driveway (no access)		0.000			0.001
		(0.002)			(0.002)
Blocked driveway (partial access)		-0.007			-0.007
		(0.006)			(0.006)
Blocked roadway (construction)		0.007			0.007
		(0.006)			(0.006)
Failed roadway repair		-0.002			-0.002
		(0.003)			(0.003)
Pothole complaint		0.002			0.002
		(0.001)			(0.001)
Rough roadway		0.001			-0.001
		(0.007)			(0.007)
Pothole outstanding		-0.001			-0.001
		(0.001)			(0.001)
Zoning: EC-2/EC-3 districts			-0.019		-0.028
			(0.006)		(0.006)
Policy Change Controls	N	N	N	Y	Y
Run-bin × Run-my FE	Y	Y	Y	Y	Y
Observations	47430	47430	47430	47430	47430
R^2	0.97	0.97	0.97	0.97	0.97
R^2-within	0.14	0.14	0.14	0.14	0.15

Note: Unless otherwise specified, robust standard errors are two-way clustered at the level of run× block and run× month. Each observation is a run-block-month.

Table 7: Congestion elasticity of vehicle density (taxis and all vehicles)

	(1)	(2)	(3)	(4)	(5)	(6)
	Log avg. s/m	Log taxi dens	Log avg. s/m	Log avg. s/m	Log vehicle dens	Log avg. s/m
	(OLS)	(1st stage)	(2SLS)	(OLS)	(1st stage)	(2SLS)
Log taxi dens	0.049		0.141			
	(0.005)		(0.013)			
Log vehicle dens				0.083		0.725
				(0.013)		(0.123)
BZ x Roll-out		0.699			0.136	
		(0.050)			(0.023)	
Bikelane	0.034	0.068	0.008	0.047	-0.021	0.033
	(0.007)	(0.068)	(0.011)	(0.008)	(0.031)	(0.024)
Citibike station	-0.032	0.044	-0.018	-0.038	0.016	-0.023
	(0.005)	(0.051)	(0.007)	(0.006)	(0.025)	(0.018)
311 complaints	0.001	0.005	0.000	0.002	-0.002	0.002
	(0.001)	(0.007)	(0.002)	(0.002)	(0.003)	(0.003)
Potholes	-0.004	0.015	-0.005	-0.003	-0.001	-0.002
	(0.002)	(0.010)	(0.002)	(0.002)	(0.004)	(0.003)
Run-block and Run-m-y FE	Y	Y	Y	Y	Y	Y
Observations	24660	24660	24660	24660	24660	24660
R^2	0.97	0.94		0.97	0.98	
F-stat			10870.3			2350.5

Note: Unless otherwise specified, robust standard errors are two-way clustered at the level of run$\times$ block and run$\times$ month. Each observation is a run-block-month. Aerial densities and the image sample are described in Appendix C. We address seasonal composition for the sample of pre- and post- roll-out aerial images by estimating weighting factors from daily counts of vehicles entering Manhattan through toll plazas (see Appendix A). We regress daily toll counts for 2010-2015 on a linear trend and the interaction of calendar month and weekend status, and then reweight aerial counts according their date and estimated seasonality factors. This weighting does not substantially affect our estimates: 2SLS coefficients on the un-weighted sample are 0.142 (taxis) and 0.802 (all vehicles).

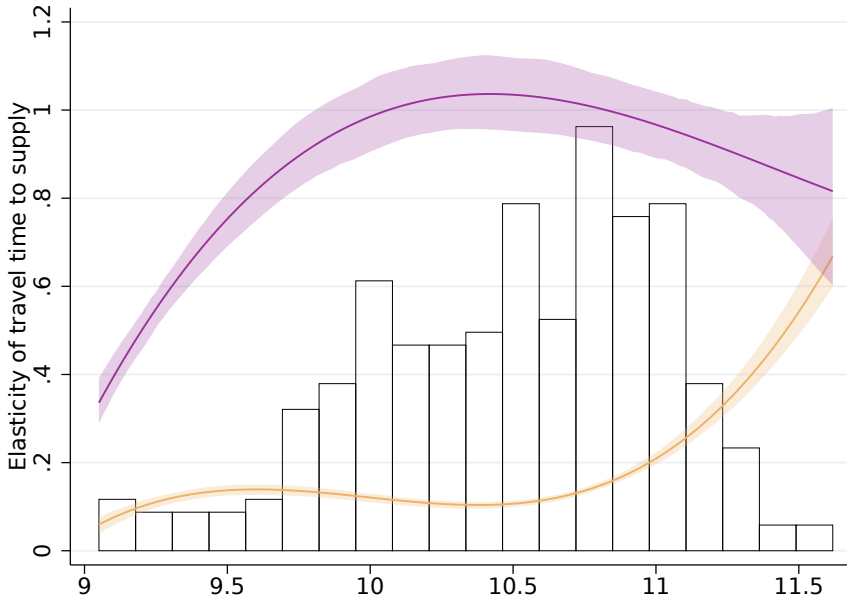


(a) Taxis

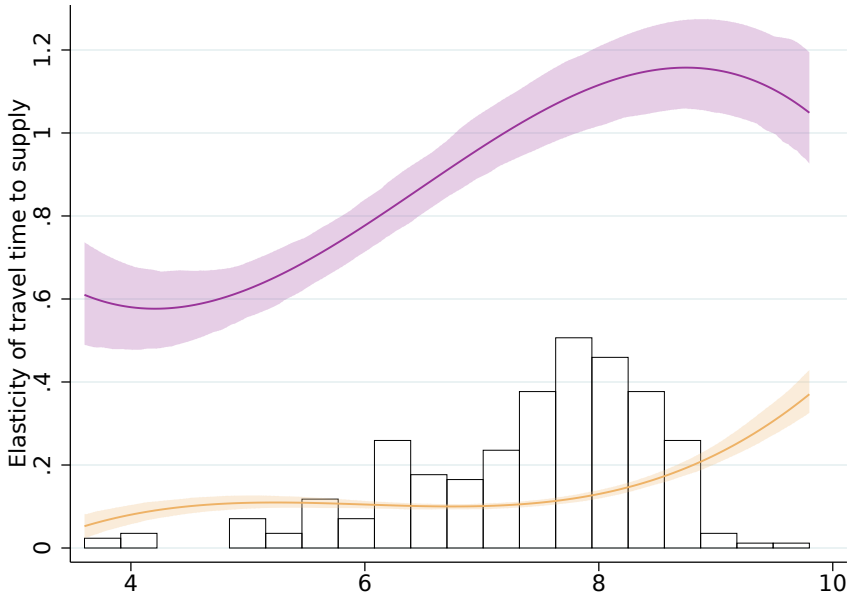
(b) All vehicles

Figure 5: Densities of vehicle located from overhead aerial imagery

Note: We plot the densities of the distance between vehicles on the road, as located over the area of interest in aerial imagery scenes, and the hail-exclusion boundary. Our area of interest is 4.4 square kilometers (1087 acres) on the east side of northern Manhattan, delimited by 1st Ave and 5th Ave and between East 82nd Street and East 125th Street. We count an average 522.7 yellow taxis per scene (or 118.9 per km^2) over 11 scenes prior to the boro taxi program, and plot the density of the locations of these taxis with a dashed yellow line. Following the launch of the boro program we count an average 440.5 yellow taxis and 112.1 boro taxis per scene over 15 scenes through June 2016 (125.6 taxis per km^2). We plot the density of the locations of these taxis with solid yellow and green lines in Figure (a). In Figure (b) we plot equivalent densities for all vehicles. All kernel density estimates are rescaled to integrate to the average number of vehicles per km^2 over 1000 meters on the horizontal axis. In this manner, the vertical axis represents a local vehicle "density" (i.e., in terms of vehicles per km^2). In Figure (a) we plot the sum of the densities for both types of taxis in the post period with a solid grey line to facilitate before-after comparison. All densities are estimated with upper and lower bounds computed by renormalization. Congestion elasticity estimates employ weighted count data for interval between 84th St. E to 120th St. E. Spikes in these distributions are in part due to vehicles counted on major East/West cross streets: 86th St. E and 87th St. E., leading to the Central Park 86th St Transverse, 96th St. E and 97th St. E., leading to the Central Park 97th St Transverse, 116th St. E and 125th St. E. Details on aerial imagery are in Appendix C.



(a) By log population density (per km^2)



(b) By log place-of-interest density (per km^2)

Figure 6: Congestion elasticity curves

Note: Congestion elasticities conditional on log population density (a) or place-of-interest density (b), as obtained by estimating Eq. 5. Curves are the travel time elasticities for all vehicles (in purple) or taxis (in orange) as a function of one urban density measure while conditioning on the other at its mean value and the lowest category of the transit indicator. Confidence intervals at 95% are obtained for the elasticity polynomial curves by two-way clustered bootrapping at the level of run$\times$ block and run$\times$ month. Each graph contains a histogram underlay of the in-sample distribution of the interacted variable.

Table 8: Vehicle densities (per km^2). Selected areas of interest.

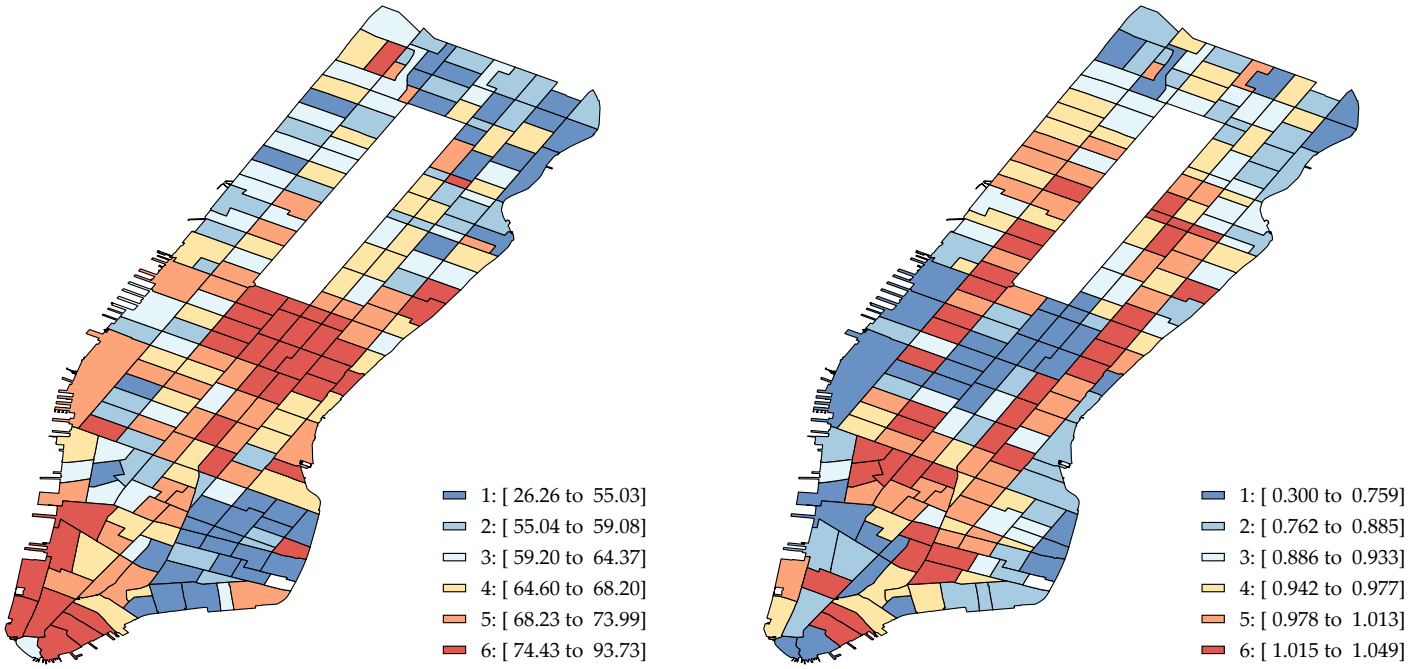
Panel B: Aerial counts	**Cars**	**Black cars**	**Trucks**	**Yellow taxi**	**Boro taxi**	**Total**	$\widehat{\textbf{Private}}$	$\widehat{\textbf{FHV}}$
Hail-exclusion zone								
Before Aug 2013	264.4	108.7	82.0	247.0		684.1	269.8	85.3
Since Aug 2013	270.4	164.9	67.1	205.4	10.1	717.8	289.6	145.7
	9.7%	51.7%***	-18.2%	-16.8%***		4.9%	7.3%	70.8%**
Boro zone								
Before Aug 2013	286.9	126.0	65.1	61.8		539.8	314.5	98.4
Since Aug 2013	309.8	160.0	54.0	58.2	36.0	617.9	346.6	123.2
	8.0%	27.0%*	-17.1%	-5.8%		14.5%	10.2%	25.2%
Midtown								
Before Aug 2013	370.3	153.5	86.1	530.7		1140.7	411.4	112.5
Since Aug 2013	404.9	341.8	79.3	468.2	7.6	1301.8	383.5	363.3
	9.3%	122.7%***	-7.9%	-11.8%*		14.1%	-6.8%	222.9%***

Note: "Hail-exclusion" and "boro" zones in both panels designate the area of interest in northeastern Manhattan, all areas described in Appendix C. Rows with percentage changes are labeled with stars (* $p < 0.1$, ** $p < .05$, *** $p < 0.01$) for a two-sided difference in means test for the vehicle density (in levels) between the period before August 2013 (11 scenes) and after (15 scenes). Almost all scenes were captured around noon due to favorable lighting conditions, and scenes include both weekday and weekend dates (see Appendix C for details). For each zone, period and vehicle type we estimate a weekend factor and re-weight weekend densities accordingly. Growth rates between the averages of vehicle densities that are weighted or unweighted in this manner are comparable, but unweighted densities are on average 14.3% lower due to lower counts in the weekend sample.

Table 9: Daily commuter welfare changes under congestion pricing scenarios

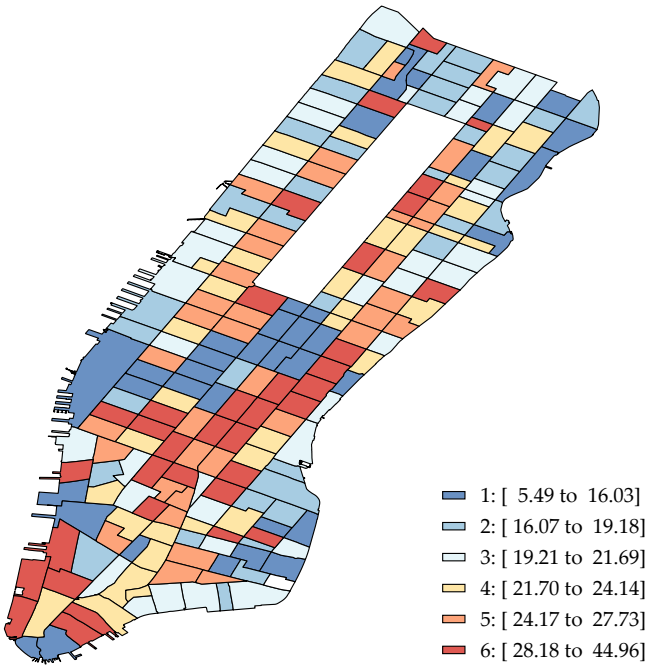
	Δq	$-MD \times \Delta q$	DWL	Revenue	Surplus	Avg trip minutes saved
A. $\gamma = -0.16$						
$p_0 = 24.64, \tau =$12.00	-8,753	229.3	48.5	1,602.6	180.8	-1.7
$p_0 = 24.64, \tau =$26.20	-15,572	408.0	175.7	3,320.3	232.3	-3.0
$p_0 = 54.64, \tau =$12.00	-4,449	116.6	25.7	1,654.2	90.9	-0.9
$p_0 = 54.64, \tau =$26.20	-8,645	226.5	104.7	3,501.8	121.8	-1.7
B. $\gamma = -0.3$						
$p_0 = 24.64, \tau =$12.00	-15,969	418.4	87.6	1,516.0	330.8	-3.1
$p_0 = 24.64, \tau =$26.20	-27,792	728.2	307.5	3,000.1	420.7	-5.3
$p_0 = 54.64, \tau =$12.00	-8,228	215.6	47.2	1,608.9	168.3	-1.6
$p_0 = 54.64, \tau =$26.20	-15,777	413.4	189.2	3,314.9	224.2	-3.0

Note: Monetary values are in thousand 2020 USD. Demand parameter $\tilde{A}$ is calibrated to actual private commute trips under the assumed initial condition for the price of commuting (p_0). Δq is the change in quantity of private car commuter trips with congestion pricing fee. $-MD \times \Delta q$ is marginal damage of 26.20 times Δq. For deadweight loss we integrate $\tilde{A}p^{\gamma}$ over the price increase and subtract revenue. Surplus is reduction in marginal damage minus DWL. Average trip minutes saved is the average commuter trip time (35.5 minutes) times $\hat{\eta}^A$ (0.913) times $\Delta q/q$, where q is total commuter, taxi, and FHV vehicle trips (169,261).



(a) Commuter wage (w_c)

(b) Congestion elasticity (η_c^A)



(c) Travel time externality of driving (MD_c)

Figure 7: Travel time externality on commuters, by tract.

Appendix A Data sources

Taxi trips. Taxi trip records are from the New York City Taxi and Limousine Commission's Taxicab Passanger Enhancements Project (TPEP) and Street Hail Livery (SHL) Passanger Enhancements Project (LPEP). The paper uses trip records from 2009 through the first half of 2016, obtained from multiple FOIL requests by the authors and as of September 2015 available for direct download from the NYC TLC's website at: www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page. As of the second half of 2016 the TLC has started removing trip endoint coordinates from their data releases, which precludes the street speed reconstruction methods developed in this paper. Data from ridehail and other for-hire vehicle providers was also obtained by FOIL prior to September 2015, and records with the time and "taxi-zone" of pickup are available from 2015 onwards at the TLC website.

311 complaints. Data from reports to New York City's 311 complaint hotline are available at www1.nyc.gov/311/our-data.page. Each record contains the date, location coordinates, and a category classification for the complaint. This paper utilizes all 311 records in northern Manhattan for categories that may be related to the flow of traffic. We count the number of complaints per month and segment of road that fall under the following categories: "Blocked Driveway - No Access," "Blocked Driveway - Partial Access," Street Condition - Blocked - Construction," "Street Condition - Failed Street Repair," "Street Condition - Pothole," and "Street Condition - Rough, Pitted or Cracked Roads."

Bike lane maps. The New York City Department of Transportation provides a shapefile for all bike lanes, including the date of installation or modification, at www.nyc.gov/html/dot/html/about/datafeeds.shtml#bikes. All changes in bike lanes (either new construction or modification) from 2009 through the first half of 2016 are matched to street segments. The various types of bike lane changes in the data include bike-friendly parking, protected paths, signed routes, and standard bike lanes.

Citibike system. Bike trip data is from the Citibike system data page at citibikenyc.com/system-data. Data on location and availability of bike stations was archived from the Citibike live station feed, available at citibikenyc.com/stations/json as well as the system data web page.

Potholes. Street Pothole Work order data is available at www.nyc.gov/html/dot/html/about/datafeeds.shtml#construction. This dataset contains the reported date and closed date of pothole work orders, identified at the level of street sections between intersections, from 2010 through 2016. Each order may refer to more than one pothole.

EZ-pass traffic speed. Data for traffic sensors in New York City is made available as a real-time feed by the NYC Department of Transportation. A link to the real time feed is available at nyc.gov/html/dot/html/about/datafeeds.shtml#realtime. The NYC-DOT has claimed in private correspondence not to archive the output of this feed. From May 27, 2013 to August 20, 2013 the feed was continuously archived and kindly made available

by Prof. Tomonari Masada of Nagasaki University. Data from April 2015 through August 2015 (and beyond) was continuously archived by *βetaNYC*, a civic technology non-profit, and is available at data.beta.nyc/dataset/nyc-real-time-traffic-speed-data-feed-archived. Data from August 2015 onward was also continuously archived by the authors. Additional snapshots of the data were obtained from archive.org at 11 arbitrary times and dates in 2014 and prior to 2013.

New York State vehicle registrations. Obtained from current and previous archived snapshots of the NY State Department of Motor Vehicles database for Vehicle, Snowmobile and Boat registrations, available at data.ny.gov. Data contains county, zipcode, car brand and body type, model year, exterior color and VIN for every vehicle registered in the state of New York.

Taxi and For-hire-vehicle inspection data. Obtained by FOIL from the Taxi and Limousine Commission's Vehicle Inspection Program. Data contains odometer readings and inspection dates for yellow taxis (performed every quarter), boro taxis (biannual) and for-hire vehicles (biennial).

Toll plaza counts. We obtain daily vehicle counts for each toll plaza in New York City for both cash and EZ-Pass customers from the Metropolitan Transportation Authority (MTA). Data is available at data.ny.gov/Transportation/Hourly-Traffic-on-Metropolitan-Transportation-Auth/qzve-kjga.

Commuter demographic data. We access the Census Transportation Planning Products (CTPP) portal via the Beyond 20/20 website at ctpp.beyond2020.com. This portal constructs statistics based on the underlying micro data from the American Community Survey 5-year samples (2012-2016). We create a sample of commuters by selecting tract-level flows for employees who live within 50 miles from midtown Manhattan and who travel for work to a census tract in Manhattan south of 125th St. The sample contains 4,510 census tracts of workers who travel to 220 Manhattan census tracts. Using the CTPP, we use table A302102 to estimate the number of commuters via private car to Manhattan census tracts. We calculate total commuters, C_c, by summing the number of passengers per car (we assign the midpoint to the bin of 5 or 6 people) and vehicles commuting to Manhattan, V_c by summing over the number of cars. These measures are summed over all origination census tracts to the level of workplace census tract. Next, we compute the average hourly wage at origination census tract by computing mean household income, mean number of workers in a household, and mean hours worked per week. We arrive at mean hourly earnings by dividing mean household earnings by the product of average hours per week, average workers in the household and 52 weeks worked per year. We scale up hourly earnings to match the growth in wages from 2014 (midpoint of the ACS sample) to 2020 (18%) and we apply a binding \$15 minimum wage.[52] The average hourly wage in the sample, \$74, is roughly in line with the implied hourly wage for Manhattan workers according to the BLS in 2020 of \$3270 per week or \$81.75 per hour for a 40 hour work week. [53] Lastly, we compute the average transit time in Manhattan for each origin

[52] https://www.bls.gov/regions/new-york-new-jersey/news-release/2014/countyemploymentandwages_newyorkcity_20141121.h
[53] bls.gov/regions/new-york-new-jersey/news-release/2014/countyemploymentandwages_newyorkcity_20141121.htm

and destination census tract pair. For origins and destinations both in midtown, we apply the average transit time associated from to each origin an destination pair from CTPP table 302203 which details commute times by mode of transportation. We combine counts of private vehicles driving alone and carpooling vehicles to arrive at an average transit time. We apply the average transit time for within-Manhattan commuters to all commuters from outside of Manhattan, as an estimate of the commute time spent in Manhattan.

Appendix B Reconstruction of historical street speeds from taxi trip data

In this appendix we describe how we use historical records of taxi trips from the Taxi and Limousine Commission (collected from GPS-enabled taxi meters since 2009 under the "Taxicab Passenger Enhancement Program") to construct measures of historical street speeds, using three alternative methodologies. Our paper employs these measures as outcome variables to quantify the impact of the roll-out of the boro program over space and time.

Appendix B.1 Average travel times per city block

The simplest approach to measuring the speed of traffic, and the approach we favor throughout the paper, is to average the rate of travel across all taxi trips for which we are confident that the taxi traversed a certain interval over a certain period. To estimate average travel times at a specific time and place, for each month in the data we select the taxi trips that were entirely contained within a rectangular strip running along one of the north-to-south avenues of Manhattan, and in which the pickup and drop-off locations are oriented in the direction of traffic. This within-avenue sample is described in Table B.1. On avenues with two directions of traffic, such as Park Avenue, we classify trips according to direction. We refer to each avenue and direction of traffic as a "run."

Every trip record contains the time as well as the latitude and longitude for each pickup and drop-off, as provided by a GPS-enabled taxi meter. We project the coordinate of each endpoint perpendicularly onto a straight line down the center of each avenue, allowing us to locate each trip endpoint on this line. We then calculate for each trip the ratio of the travel time to the line distance, to obtain each trip's average seconds per meter (i.e., pace, or the inverse of speed). We then calculate the average seconds per meter for all the taxi trips that fully traverse each city block on each run. Variation in the average travel time between block b and block $b+1$ therefore comes from the variation in the average travel times of taxi trips that fully traverse b but not $b+1$, or vice-versa.

Average speeds can be calculated along a rolling window of any length. In Figures B.2 and B.3 we display these averages at a 10 meter resolution. For ease of presentation we average the time series dimension of our monthly estimates to the annual level, and express speed in miles per hour rather than seconds per meter.

Since our objects of interest are (averages of) speed over particular line segments, the averages obtained by the method described above could be subject to biases due to

Table B.1: Seconds-per-meter for within-avenue taxi trips,
Northern Manhattan, Jan 2009 to Jun 2016

	N	Mean	Median	Std Dev	Min	Max
West End Northbound	262,340	0.223	0.199	0.127	0.024	6.37
West End Southbound	379,378	0.200	0.181	0.112	0.027	8.21
Broadway Northbound	1,068,259	0.205	0.181	0.130	0.026	10.90
Broadway Southbound	1,027,907	0.212	0.187	0.134	0.025	11.18
Amsterdam Northbound	514,332	0.148	0.125	0.119	0.029	12.55
Amsterdam Southbound	93,688	0.238	0.204	0.166	0.040	13.40
Columbus Northbound	1,308	0.338	0.230	0.370	0.080	4.48
Columbus Southbound	317,496	0.186	0.146	0.174	0.033	10.97
Manhattan Northbound	4,680	0.315	0.231	0.339	0.076	5.48
Manhattan Southbound	9,878	0.245	0.196	0.231	0.073	5.88
CPW Northbound	132,153	0.224	0.181	0.204	0.052	8.24
CPW Southbound	132,674	0.230	0.189	0.203	0.039	8.85
5th (S)	199,727	0.242	0.191	0.209	0.035	11.26
Madison (N)	347,121	0.209	0.163	0.190	0.035	7.74
Park Northbound	188,904	0.238	0.196	0.193	0.026	10.16
Park Southbound	157,976	0.281	0.224	0.243	0.030	8.39
Lexington (S)	272,852	0.219	0.172	0.201	0.033	13.22
3rd (N)	440,152	0.155	0.121	0.153	0.031	14.69
2nd (S)	427,408	0.194	0.153	0.174	0.026	10.07
1st (N)	260,863	0.157	0.119	0.160	0.031	10.29
Upper East Side	2,295,003	0.201	0.157	0.189	0.026	14.69
Upper West Side	3,944,093	0.201	0.177	0.142	0.024	13.40
All	6,239,096	0.201	0.171	0.161	0.024	14.69

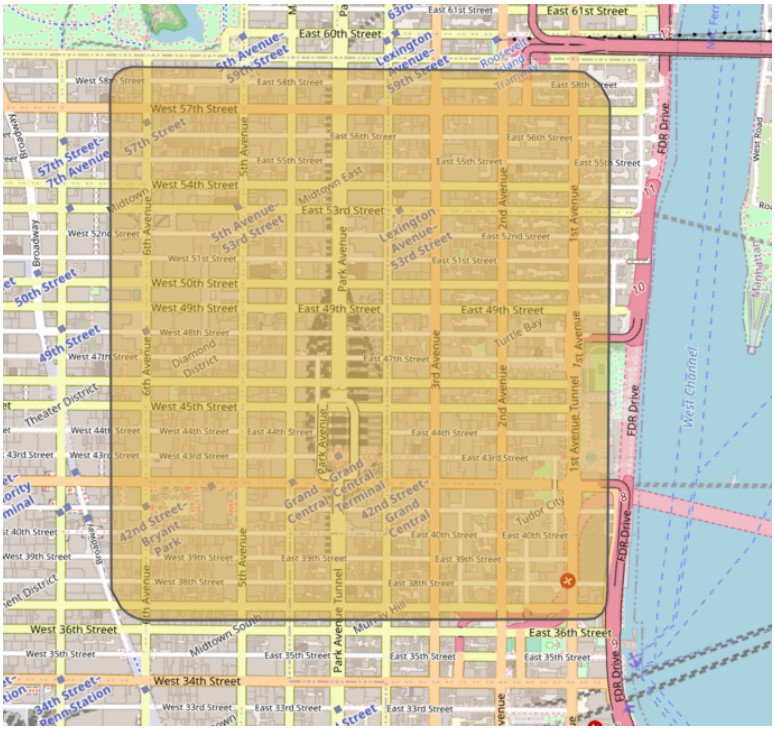
Excluded: Riverside Drive, FDR Drive, Adam Clayton Powell Jr. Boulevard, Malcolm X Boulevard. Statistics are for the seconds per meter, calculated as travel time over the distance along an avenue centerline. The following are statistics in miles per hour for all within-trip observations: the average of 0.201 seconds per meter is the inverse of a speed of 11.13 miles per hour. The maximum is 93.2 miles per hour, the minimum is 0.15 miles per hour, and two standard deviations above the mean is 4.27 miles per hour. Not described in the table above, the 1st and 99th percentile seconds per meter are 0.0789 and 0.762, which are the inverses of 28.4 and 2.93 miles per hour, respectively.

aggregation as well as changes in trip selection over space and time. As an example of this potential concern, consider two blocks in a relatively uncongested area in the north of the city, and assume traffic moves at the same speed along both blocks. Suppose the northernmost of these blocks contains a hotel driveway, whereas its neighbor to the south does not. The block containing the hotel originates many long trips to a congested downtown area, whereas the neighboring block originates short trips within a relatively uncongested interval. In this example, the average taken over the travel time of the trips crossing each block will be upwardly biased as an estimator of the true marginal travel time of crossing each block due to congestion that affects the taxis traveling downtown from the hotel (an aggregation bias) and, further, the bias will be larger for the block containing the hotel since its average does not contain the short, local trips originating in the southern neighbor (a selection bias). Aggregation bias can also interact with time, as the block may be located next to a nightclub or park, and trips may originate at times when traffic is lower, such as nights or weekends.

A third type of potential bias could arise from the type and behavior of the "probe" that provides a taxi trip record, e.g. whether a car is a yellow or boro taxi. We employ data on both types of taxis, since boro taxis increase the sample in the north of the city, but for this reason and because boro taxi trips are by definition only possible after the launch of the boro program, variation across taxi type in driving speed would bias our travel time estimates in a manner that could violate parallel-trend assumptions required for difference-in-differences evaluation. To investigate this possibility, we exploit the fact that the data contains approximately half a million trips per day to isolate quasi-experimental "taxi race" conditions in which both a yellow and a boro taxi both picked up and dropped off a passenger within 10 meters of each other, as well as within a minute of each other. We find 234 such natural experiments in our area of northern Manhattan. We then regress travel time outcomes on a fixed effect per "taxi race" and a binary indicator for the boro taxi in the pair. Results are reported in Table B.2. We find positive but non-significant coefficients in travel time levels, logs or ratios (i.e. seconds-per-meter). Since the magnitude of the estimated coefficient is a non-trivial 2.5% speed difference, we explore relaxing the threshold on pickup time difference from one minute apart to two, and up to 12, and find results that are statistically significant at the 5% level in a couple specifications in logs, as well as in levels, although not in seconds-per-meter. Alternative specifications intended to control for within-experiment variation in observables (such as a "pole position" effect, number of passengers or the residual variation in trip distance on the taxi meter) do not affect the results. We take this collectively as evidence that there may be a small difference in driving speeds for boro and yellow taxis.[54]

[54] A reason for this speed difference could be that a higher proportion of boro taxis are wheelchair-accessible vans, and these vehicles may accelerate more slowly than the average taxi. Alternatively, since daily rental rates (for medallion and car) are lower for boro than yellow taxis, it may be the case that drivers respond to the different opportunity costs of time, or that the activity of driving a boro taxi selects drivers with a lower subjective opportunity cost of time, i.e. slower drivers.

(a) Northern Manhattan



(b) Midtown

Figure B.1: Areas of interet

Note: Figure B.1a displays area of interest in Northern Manhattan, including North-South polygons employed for trip classification and boro zone boundaries east and west of Central Park.

Table B.2: "Taxi race" quasi-experiments: speed comparison between yellow and boro taxis

	(12)	(11)	(10)	(9)	(8)	(7)	(6)	(5)	(4)	(3)	(2)	(1)
Seconds	7.763	7.980	6.656	6.260	6.857	6.613	8.512	10.88	14.09	16.19	12.22	10.86
	(3.633)	(3.655)	(3.580)	(3.737)	(3.982)	(4.284)	(4.611)	(4.674)	(5.258)	(6.154)	(7.094)	(6.812)
$\log(Seconds)$	0.0220	0.0237	0.0148	0.0133	0.0129	0.0106	0.0140	0.0256	0.0373	0.0446	0.0351	0.0252
	(0.0171)	(0.0171)	(0.0161)	(0.0169)	(0.0177)	(0.0193)	(0.0216)	(0.0168)	(0.0187)	(0.0218)	(0.0265)	(0.0245)
Seconds-per-meter	-0.0805	-0.0808	-0.319	-0.340	-0.372	-0.416	-0.508	0.00379	0.00542	0.00621	0.000126	0.00289
	(0.396)	(0.399)	(0.342)	(0.365)	(0.398)	(0.444)	(0.511)	(0.0221)	(0.0260)	(0.0315)	(0.0428)	(0.00686)
Race FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	1,280	1,270	1,220	1,146	1,050	940	816	704	598	494	362	234

Note: All specifications are "taxi race" regressions in which the only regressor is a binary indicator for the boro taxi in a taxi pair, and all between-variation is absorbed by a fixed effect per taxi race, as in a standard "twins" specification. Rows indicate the outcome variable. Columns indicate the number of minutes difference in the pickup time that are allowed for a pair of taxi trip records to be classified as a "taxi race," so sample size increases as this threshold is relaxed. Standard errors clustered at the "race" level. Results are unchanged if controlling for remaining within-experiment variation in observables, such as a "pole position" effect, number of passengers or the residual variation in trip distance on the taxi meter, in the two specifications where the outcome is in units of time rather than a rate of change.

To our knowledge, no previous paper has recovered street speeds from taxi GPS "probe" data in a manner that resolves the aggregation issues described above, or employed such estimates for the evaluation of urban transportation policy.[55]

Next, we develop two alternative methods to estimate travel time that address the bias issues outlined above at the expense of estimation variance. We then describe the in-sample predictive fit of each method. We evaluate the sensitivity of our main estimates to employing these alternative estimates in Section 2.

Appendix B.2 Projection on dummies to recover a travel time function

Our first approach to estimating marginal travel times at the street level while controlling for spatial and temporal aggregation bias consists of projecting travel times onto dummy variables for the intervals that are traversed by each individual trip, as well as observables such as taxi type. This approach is intended as an intermediate illustration for our next approach, so we present it in a summary manner.

We predefine a set of interval bins $b = \{\underline{b}, \ldots, \bar{b}\}$ of varying width over several city blocks, and construct a set of indicators variables m_{bi} that are turned off at a value of zero if a taxi trip does not intersect the interval and are turned on at a value equivalent to the width of an interval if the trip intersects it (this is equivalent to regressing travel time on dummies and appropriately rescaling coefficients after estimation). We also define sets of binary indicators that turn on if trip i starts (s_{bi}) or ends (e_{bi}) in a given interval. For each month and uptown or downtown "run" along an avenue, we project the travel time T_i of trip i onto these indicators:

$$T_i = \sum_{b=\underline{b}}^{\bar{b}} \beta_{mb} m_{bi} + \sum_{b=\underline{b}+1}^{\bar{b}-1} \beta_{sb} s_{bi} + \sum_{b=\underline{b}}^{\bar{b}} \beta_{eb} e_{bi} + \beta x_i + \delta_h + \varepsilon_i \tag{8}$$

where additional controls include hour fixed effects, two alternative controls for boro taxi status (a dummy, consistent for a fixed cost, and a trip distance measure interacted with boro taxi status, consistent for a boro variable cost) and a control for "surplus distance" on the taxi meter after the coordinates of each trip endpoint undergo a perpendicular

[55] Three prior papers in the computer science literature, however, have developed methods to reconstruct street speeds for the purposes of exploratory and descriptive analysis within a large urban area. Zhan *et al.* (2013), Santi *et al.* (2014) and Poco *et al.* (2015) are concerned with recovering speeds on street segments from the TLC trip data given that trip origins and destinations are located on the NYC road network and the path taken by the driver is unknown. The approach taken in these papers is to employ a routing algorithm to determine a set of latent, unobserved paths, and then allocate total trip travel times to weighted segments from the latent paths. Since our focus is precision within a specific area of interest, rather than exploratory analysis, we circumvent the path attribution problem by employing only the subset of trips whose origin and destination are located along the same avenue and in the direction of traffic. Conditional on a given latent path, Santi *et al.* (2014) and Poco *et al.* (2015) average travel times over trips crossing a segment and are therefore identical to our approach in this section, whereas Zhan *et al.* (2013) project trip times onto segment dummies and are therefore identical to the approach we describe in the next section, although their approach does not contain controls for observable trip characteristics, (e.g. driver type, or distance from GPS coordinates to grid network), or allow for non-linearity arising from fixed and variable components of travel time cost (e.g. due to acceleration).

projection onto the straight north-south line along each avenue. Each trip will also turn on an indicator dummy both for both endpoints, as well as a number of middle indicators. Because we observe trips that fully traverse the intervals $b = \{\underline{b}+1, \ldots, \bar{b}-1\}$, and have saturated the specification with terms that can absorb contributions to travel time from trip events other than the crossing of these intervals, the estimated coefficients $\hat{\beta}_{mb}$ for $b = \{\underline{b}+1, \ldots, \bar{b}-1\}$ are consistent for (an average over) the interval marginal second per meter function specific to bin b, whereas the coefficients on $\hat{\beta}_{mb}$ for $b = \{\underline{b}, \bar{b}\}$ are consistent for an average of the marginal second per meter function over the fractional measures of the interval that are intersected in the data, which are not a quantity of interest and are discarded. All of the coefficients on $\hat{\beta}_{sb}$ and $\hat{\beta}_{eb}$ play a similar role, as they are consistent for average marginal costs on fractionally traversed intervals, and are also discarded. Given that trip data is only observed in the interior of the interval defined by the lower limit of $\underline{b}$ and the upper limit of $\bar{b}$, the average speed parameters β_{mb} for $b = \{\underline{b}, \bar{b}\}$ are not identified. The estimated coefficients $\hat{\beta}_{mb}$ for $b = \{\underline{b}+1, \ldots, \bar{b}-1\}$ constitute our second measure of an underlying "seconds-per-meter" travel time function. These estimates control for spatial aggregation bias, since travel times are projected linearly onto the event of crossing specific intervals, and so estimates are computed from differences in the overlaps between intervals traversed. Continuing with the hotel example from above, if variation elsewhere in the data identifies downtown segments as being particularly slow, this high time cost will be subtracted from the southbound trips originating at the hotel, allowing travel time in the hotel's interval to be identified from that interval's actual contribution to travel time (lower, in our example). This method also allows us to include hour fixed effects to control for the time aggregation bias that would arise from correlation between trip endpoints and time of day, as well as controls for trip observables.

The projection on dummies method does have some costs, however, as it involves substituting variance for bias. Discretizing the "second per meter" travel time function into a set of intervals leads to loss of information, as the method discards average travel time data on the portion of each trip that does not fully traverse an interval. Discretization cannot be made arbitrarily low in a finite sample: it will eventually lead to intervals that do not originate or end trips, causing identification failure as the projection matrix loses full rank. Most importantly, since the estimates in the projection method are obtained from differencing time costs across intervals, errors in one interval can propagate to others, leading to high variance of estimates. The discontinuity in the seconds-per-meter function that is imposed as an assumption on the data generating process is one source of variance, as it leads to abrupt changes in the travel times that are attributed to otherwise similar trips if only one of these crosses the junction point used in discretization.[56] To address these concerns, our next method also controls for potential aggregation bias while delivering a better trade-off in terms of reduced variance.

[56] In our application, we need to employ as few at seven intervals to obtain estimated travel times that are *negative* in only 1.7% of all run-block-month cells.

Appendix B.3 Projection on B-splines to recover a travel time function

Our second projection method controls for aggregation bias and observable trip characteristics in the same way as the projection dummy method above, but explicitly models a continuous "seconds per meter" function with basis splines: the advantages of this method are that it eliminates discontinuities at junction points between interval dummies, and that it eliminates data loss on trip segments that do not fully traverse a discrete interval, both of which reduce the variance of estimates. Additionally, this method produces a flexible estimate for the "seconds per meter" function that can be evaluated at every point of its support.

Let s be the true "seconds per meter" function, which is defined over an interval of road during a given month. A taxi trip i that begins at point a_i and ends at point b_i on the interval will take a total time:

$$T_i = \beta_1 x_{1i} + \int_{a_i}^{b_i} \left[s(\tau) + \beta_2 x_{2i} \right] d\tau + \varepsilon_i \tag{9a}$$

$$T_i = \beta_1 x_{1i} + S(b_i) - S(a_i) + \beta_2 (b_i - a_i) x_{2i} + \varepsilon_i \tag{9b}$$

where x_{1i} are trip characteristics that do not depend on location or trip distance (i.e. fixed time costs), x_{2i} are trip characteristics that depend on trip distance but not location (variable time costs) and $S(b_i) - S(a_i)$ is the area under the seconds-per-meter function. We can estimate (Eq. 9b) by regressing T_i on the spline differences $[S(b_i) - S(a_i)]$, x_{1i} and $(b_i - a_i) x_{2i}$, which we can implement using basis splines for the location of trip origins and destinations, provided we impose the constraint that the spline function is the same at both endpoints.[57]

To implement B-splines, if we let:

$$S(\tau) = \sum_{k=1}^{K} \gamma_k \psi_k(\tau)$$

where ψ_k are the bases of the spline, then the regression above can be written as:

$$T_i = \beta_1 x_{1i} + \sum_{k=1}^{K} \gamma_k \left[\psi_k(b_i) - \psi_k(a_i) \right] + \beta_2 (b_i - a_i) x_{2i} + \varepsilon_i \tag{10}$$

[57] Our specification for the error term assumes that it is additive to integrated travel time. Alternatively, we could assume an instantaneous mean zero error that arises over the interval. In this case we would have that $T_i = x_{1i} + \int_{a_i}^{b_i} s(\tau) + \beta x_{2i} + \varepsilon_i(\tau) d\tau$ and so we could write $T_i = x_{1i} + S(b_i) - S(a_i) + \beta (b_i - a_i) x_{2i} + (b_i - a_i) \bar{\varepsilon}_i$ where $\bar{\varepsilon}_i = (b_i - a_i)^{-1} \int_{a_i}^{b_i} \varepsilon_i(\tau) d\tau$ is a heteroskedastic error term. We could estimate this equation by GLS, or divide both sides by $(b_i - a_i)$:

$$\bar{s}_i = \frac{T_i}{(b_i - a_i)} = \beta_1 \frac{x_{1i}}{(b_i - a_i)} + \frac{S(b_i) - S(a_i)}{(b_i - a_i)} + \beta_2 x_{2i} + \bar{\varepsilon}_i$$

where $\bar{s}_i = \frac{T_i}{(b_i - a_i)}$ is a trip's average seconds per meter. We can then implement a regression of $\bar{s}i$ on rescaled splines differences $\left[\frac{S(b_i) - S(a_i)}{(b_i - a_i)} \right]$ and x_i, which imposes the GLS correction for heteroskedasticity under the assumption of integrated errors.

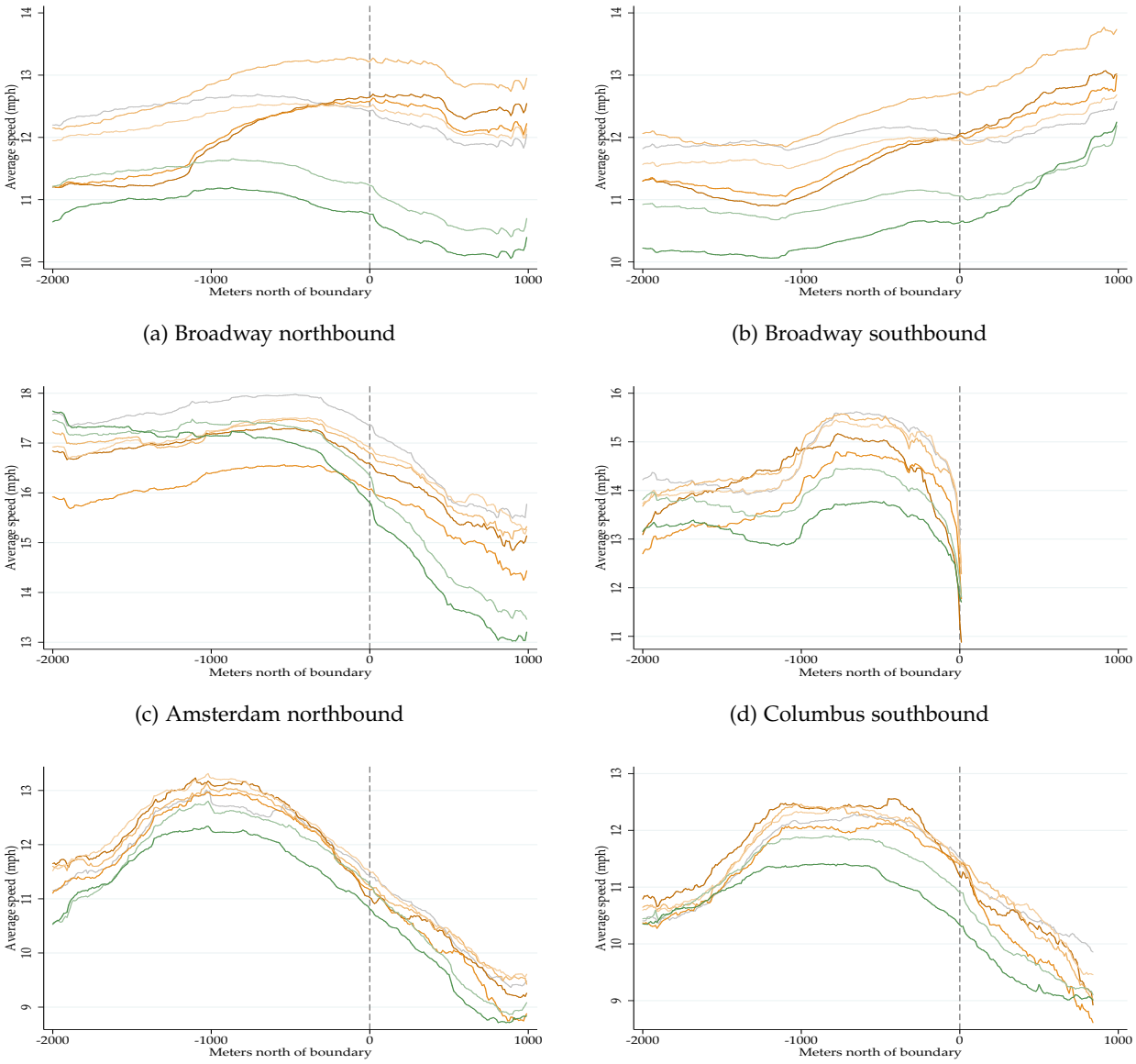
which can be estimated by OLS. Since the bases $\psi_k(\tau)$ are differentiable, after estimation we can recover

$$\hat{s}(\tau) = \hat{S}'(\tau) = \sum_{k=1}^{K} \hat{\gamma}_k \psi_k'(\tau)$$

where $\hat{\gamma}_k$ are coefficients estimated by OLS.

To implement estimation we use differences in a 9-knot cubic basis spline with knots defined every 500 meters from 0 to 4,000 meters along an avenue of interest – essentially a flexible function with 11 degrees of freedom. Additional controls include an intercept and distance interaction for boro taxi status, the "surplus distance" control described above, and an interaction of distance with hour effects..

Figure B.4 displays an exercise of plotting quantiles of actual over predicted speed for each method. All methods display some regression to the mean, slightly over-predicting trip time for the shortest trips and under-predicting trip time for the longest trips, but we view their fit as comparable: their root mean-square errors are 73.9 seconds (Avg.), 71.6 (Projected, block-level) and 73.0 (B-splines). RMSE refers to in-sample fit, however, and can be improved mechanically with the addition of free parameters, at the expense of variance.



(a) Broadway northbound

(b) Broadway southbound

(c) Amsterdam northbound

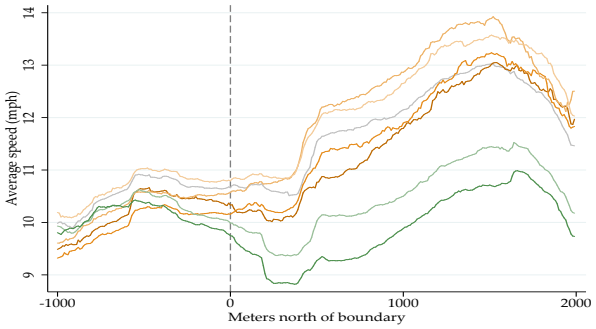
(d) Columbus southbound

(e) Central Park West northbound

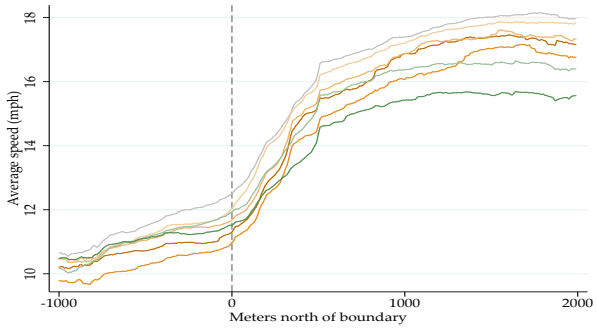
(f) Central Park West southbound

Figure B.2: Average travel time at the avenue level, over 10 meter intervals. West side

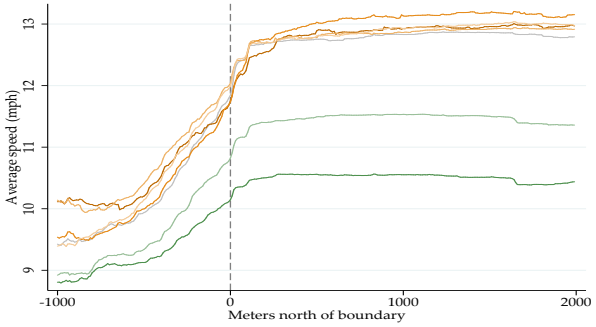
Note: Average travel time is computed as the average seconds per meter at the 10 meter interval over all trips that cross the interval. Averages shown here are annual and have been converted to miles per hour. Year is indicated by color and shade, according to the following scheme: years prior to the boro taxi program are indicated in orange, from 2009 (darkest) to 2012 (lightest). 2013 is shown in gray, as program roll-out begins on Aug 8, 2013. 2014 is shown in light green, and 2015 in dark green. 2016 not shown to control for seasonality, since data is only available for the first half of the year. Not plotted: both directions of West End Ave, which are included in regressions but do not identify the coefficient of interest because, like Columbus Ave, they do not cross the hail-exclusion boundary. Manhattan Ave is excluded from plots and regressions because in its short length it does not generate sufficient within-avenue trips to allow for speed measurement.



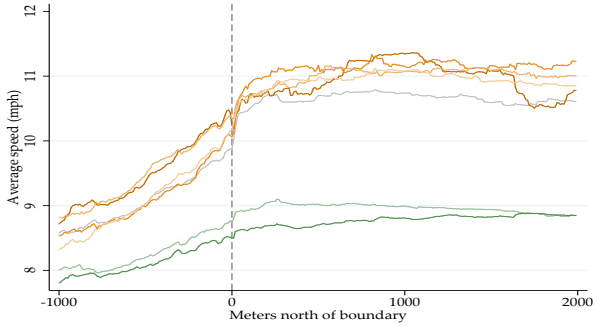
(a) 5th Avenue (southbound)



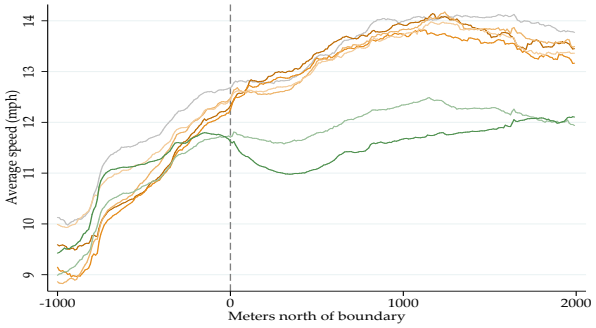
(b) Madison Avenue (northbound)



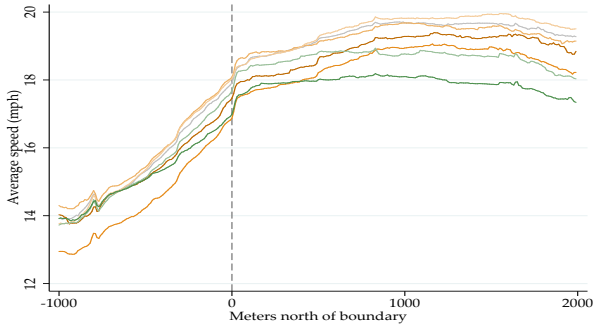
(c) Park Avenue (northbound)



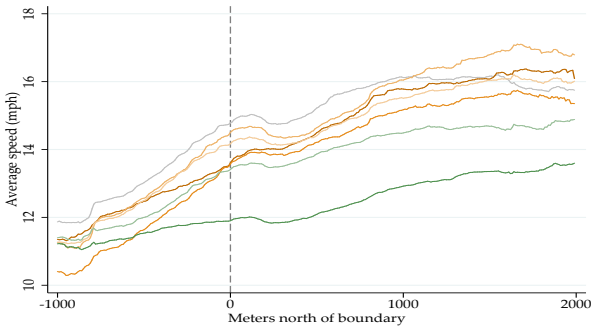
(d) Park Avenue (southbound)



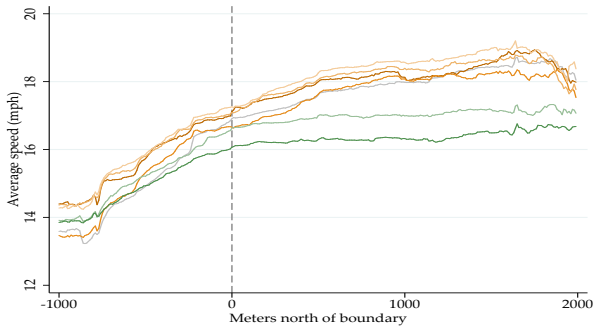
(e) Lexington Avenue (southbound)



(f) 3rd Avenue (northbound)

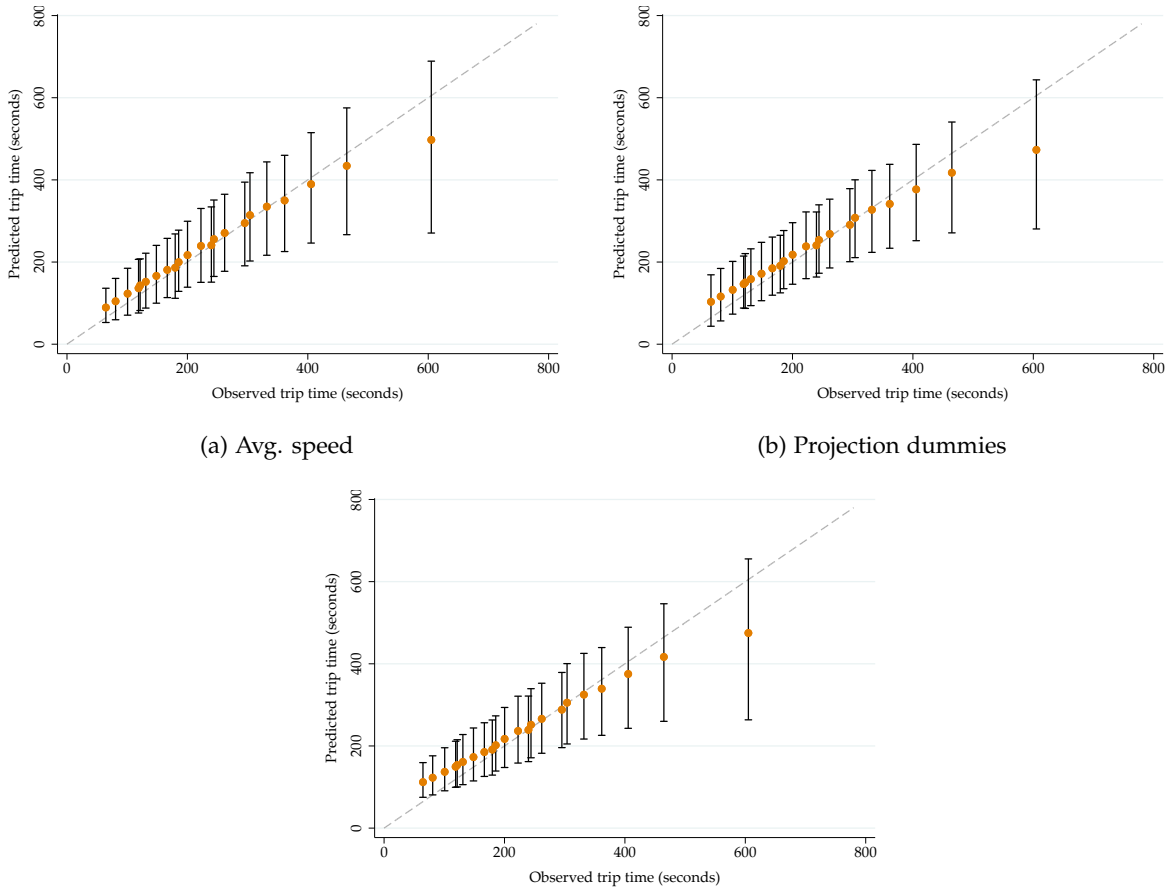


(g) 2nd Avenue (southbound)



(h) 1st Avenue (northbound)

Figure B.3: Average travel time at the avenue level (continued). East side



(a) Avg. speed

(b) Projection dummies

(c) B-splines

Figure B.4: Evaluation of travel time model fit

Note: Each observation is a unique taxi trip. We plot the mean of predicted trip time for each quantile of actual trip travel times. Dashed line indicates the 45°line of perfect prediction. Whisker bars indicate the 5th and 95th percentile of predicted speeds for each quantile in the distribution of observed trip durations. By construction, the mean predicted speed of the top quantile will be systematically below the 45°due to the few extremely long trips at the tail.

Appendix C Aerial imagery data (for online publication)

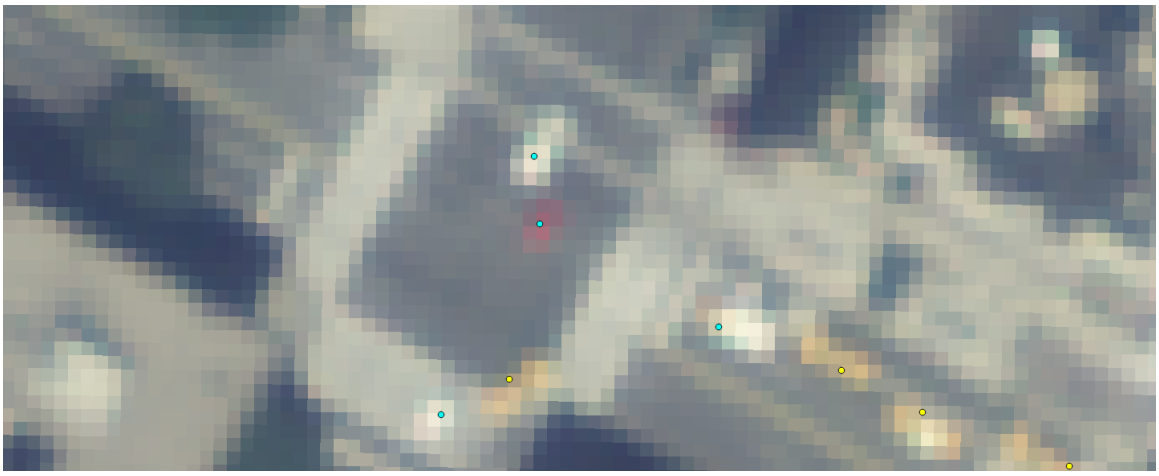
We digitize high-resolution, natural color aerial orthoimagery captured over New York City between 2010 and 2017. Our sample of 29 scenes comes from state and federal government sources (US Dept. of Agriculture, the New York City government), online services (Google, ESRI, Bing), and a commercial vendor (NearMap Ltd.). We have excluded a second commercial vendor (Eagle View Technologies, Inc.) due to licensing cost. Unless otherwise noted below, the nominal resolution for imagery is 15 cm or better. Figure C.5 illustrates this variation in resolution. The sources are:

- 3 scenes from National Agriculture Imagery Program of the US Department of Agriculture (see fsa.usda.gov/programs-and-services/aerial-photography/imagery-programs/naip-imagery for sources). July 10, 2011 (nominal 1m resolution). June 22, 2013 (nominal 1m resolution). May 22, 2015 (nominal 50cm resolution).
- 4 scenes from NY State Digital Orthoimagery Program, distributed by NYC government at maps.nyc.gov/tiles.
- 11 scenes from Google aerial imagery, obtained through the Google Earth Pro desktop application.
- 1 scene from Bing Maps Aerial Imagery layer.
- 1 scene from ESRI World Imagery layer. Resolution is unknown, but is provided at approximately 30 cm.
- 9 scenes from Nearmap Ltd. (© 2017).

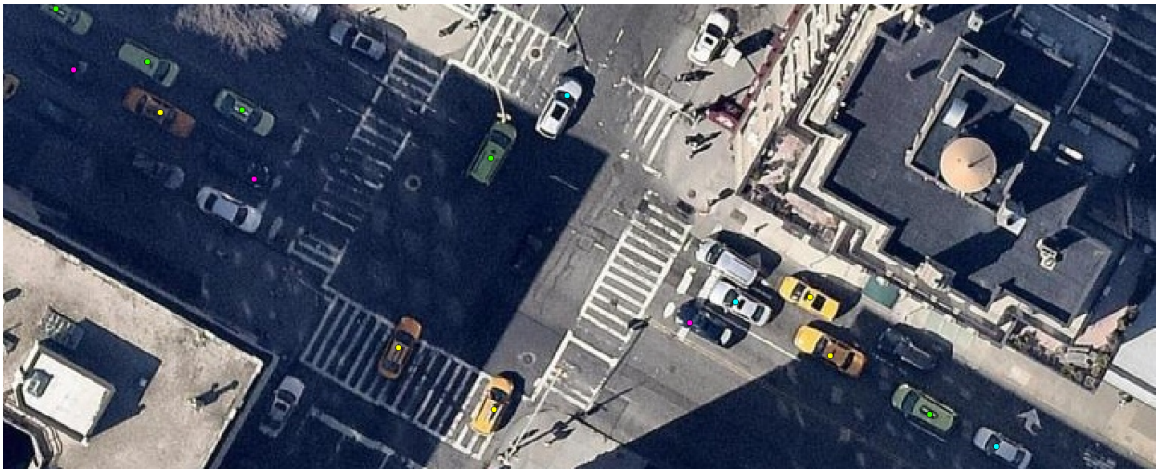
On every image we digitize vehicle locations in two areas of interest: i) an area in the northeast of Manhattan that includes parts of the Upper East Side and East Harlem neighborhoods, delimited between 1st and 5th avenues and between East 79th and East 128th streets, and ii) an area in the eastern Midtown district, delimited between 1st and 6th avenues and between 38th and 57th streets (we select the same area as Zhan *et al.*, 2013, which includes the densest zone in midtown Manhattan). We manually tag five classes of vehicles: yellow taxis, green taxis, black cars, cars of any color other than black and trucks, excluding city buses. Yellow and green taxis are tagged anywhere on the roadbed, whether parked or in apparent motion. Cars and trucks are tagged only on traffic lanes. We exclude taxis on parking lots off the roadbed.

A fraction of our aerial images are subject to some degree of road occlusion, i.e. portions of the roadbed are obstructed from view by tall buildings due to the angle at which the camera captured the aerial photography. We impute observations that are missing due to visual occlusion with the following procedure:

1. For any aerial scene i we start with the set of all observed and digitized vehicle locations L_i. We refer to any aerial scene with some occlusion as a *recipient* image, and any scene with no occlusions as a *donor* image. We separately process the images that were captured in the period before or after August 2013. For donors, we refer to the sets of these images as $\mathcal{D}_b$ and $\mathcal{D}_a$, respectively.



(a) Source: National Agricultural Imagery Program. July 10, 2011. Time estimated as 12.30 pm.



(b) Source: Nearmap Ltd. (© 2017). April 12, 2015, within 5 minutes of 2.57 pm.

Figure C.5: Aerial imagery examples. Madison Ave and East 96th Street

Note: the example areas displayed have an area of 0.7 acres, and are intended to illustrate variation in image resolution (1 meter vs. 15 cm). Our area of interest consists of 1087 acres (1.7 square miles) on the east side of northern Manhattan, delimited by 1st Ave and 5th Ave and between East 82nd Street and East 125th Street. Circle icons on the aerial imagery indicate vehicles labeled as taxis (yellow), boro taxis (green), cars (cyan) or black cars (magenta).

2. For every recipient, we digitize a layer of polygons over every occluded road surface, and refer to the set of these areas as O_r.

3. For every donor d, and for every recipient r that is in the same period $t = \{b, a\}$ as d, we partition L_d into $\tilde{L}_d^r$, the set of vehicle locations that would be counterfactually occluded under the occlusion layer O_r, and $\tilde{L}_d^{-r}$, the set of vehicles that would remain visible under the counterfactual occlusion layer O_r.

4. For every recipient r in period $t = \{b, a\}$, we assemble a donor pool of every observed but counterfactually occluded vehicle location $P_r = \bigcup_{d \in \mathcal{D}_t} \tilde{L}_d^r$.

5. We define the augmented vehicle location set $L_r^+ = L_r \cup P_r$, where to each element in L_r we attach a weight of 1, and to each element in P_r from donor d we assign the scalar weight $w_{dr} = (1/(|\mathcal{D}_t|)\left[|L_r|/(|L_d| - |\tilde{L}_d^r|)\right]$.

The first factor in the weight w_{dr} simply averages over the set of donors. The second maintains proportionality between vehicle counts in the image areas that are mutually observed in the recipient and each donor image. For example, if donor d contains half as many vehicles as a recipient r in the area that is mutually observable in both images, any vehicle locations donated from d to r will be weighted by $1/2$, and further re-weighted in proportion to the number of donor images. We apply the above procedure separately for each area of interest and vehicle type.

Buildings in area (i) are relatively low and imputations have little impact on counts and densities: although 16 of 29 images contain some imputations, imputations only increase the total vehicle count in this area by 1.1% and additions to the sample are only substantial in four images: the five largest percentage increases in weighted vehicle count are 11.3%, 10.1%, 7.8%, 7.8% and 1.2%. Area (ii) in midtown Manhattan is dense with skyscrapers, so even small deviations from a perpendicular angle in aerial photography can result in road occlusion. The imputation procedure described above increases the total vehicle count by 6.1% in this area, with imputed vehicles added to 22 scenes. The vehicle count is increased by more than 10% in 8 scenes, and between 10 and 1% in a further 9 scenes. Details on raw, imputed and total counts for both areas are provided in Tables C.3 and C.4 below.

Table C.3: Aerial imagery sources and counts. Upper East Side and East Harlem

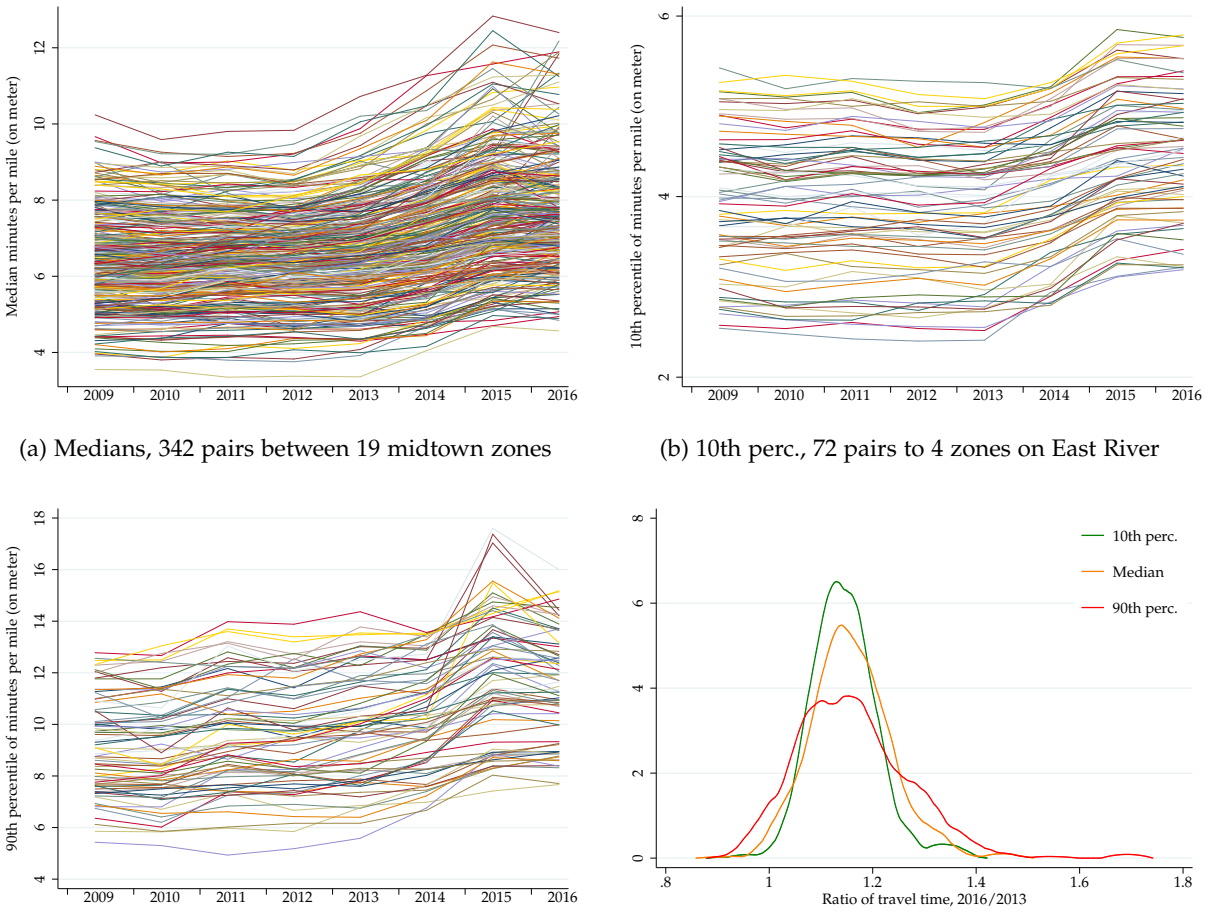
Date	Source	Raw count	Imputed	Total	Time of day
6/15/2010	NYC Ortho	2,819		2,819.0	10:30 am
6/17/2010	Google Earth	2,314	181.2	2,495.2	10:15 am
3/26/2011	ESRI Aerial	1,649	2.5	1,651.5	12:30 pm
6/2/2011	Google Earth	3,279		3,279.0	11:45 am
7/10/2011	NAIP	1,920	3.7	1,923.7	12:30 pm
3/10/2012	Google Earth	1,708		1,708.0	11:30 am
6/15/2012 to 6/20/2012	NYC Ortho	3,271		3,271.0	11:15 am
8/6/2012	Google Earth	2,786		2,786.0	9:30 am
11/5/2012	Google Earth	2,423	274.4	2,697.4	10:30 am
5/26/2013	Google Earth	1,257		1,257.0	11:00 am
6/22/2013	NAIP	1,297	100.6	1,397.6	8:30 am
6/1/2014	NYC Ortho	2,160	27.1	2,187.1	11:00 am
6/19/2014	Google Earth	3,120	11.8	3,131.8	12:30 pm
7/6/2014 to 8/10/2014	Bing Aerial	2,657		2,657.0	9:15 am
9/15/2014	Nearmap	2,893		2,893.0	9:40 am to 10:08 am
10/11/2014	Google Earth	1,150	116.4	1,266.4	10:45 am
4/12/2015	Nearmap	2,686	3.9	2,689.9	2:44 pm to 3:11 pm
5/22/2015	NAIP	2,770	2.3	2,772.3	9:45 am
6/22/2015	Nearmap	3,016		3,016.0	10:15 am to 10:47 am
9/6/2015	Google Earth	1,223	0.2	1,223.2	11:30 am
9/6/2015	Nearmap	1,884		1,884.0	1:01 pm to 1:30 pm
10/26/2015	Nearmap	2,574		2,574.0	11:35 am to 12:03 pm
3/26/2016 to 4/5/2016	NYC Ortho	3,750		3,750.0	12:45 pm
4/16/2016	Nearmap	2,534	14.3	2,548.3	11:31 am to 12:09 pm
6/24/2016	Google Earth	2,123	0.3	2,123.3	11:30 am
6/25/2016	Google Earth	1,617	1.4	1,618.4	11:30 am
7/16/2016	Nearmap	1,876		1,876.0	11:11 am to 11:37 am
10/15/2016	Nearmap	2,667	7.0	2,674.0	11:34 am to 11:58 am
4/9/2017	Nearmap	2,632	3.8	2,635.8	1:05 pm to 1:27 pm

Note: Capture dates are provided in metadata for all sources, except two NYC scenes for which a date range is provided, and the Bing Aerial layer which was captured between July 6 and August 10, 2014 according to imagery service API metadata. Time-of-day of image capture is provided in a precise range by Nearmap. For other sources, the time-of-day is estimated by using the shade of the Columbus Circle monument as a sundial (calendar-adjusted). Building shadows in area of interest are consistent with monument shadows. Nearmap metadata and imagery validates our sundial method to within half-hour accuracy.

Table C.4: Aerial imagery sources and counts. Midtown

Date	Source	Raw count	Imputed	Total	Time of day
6/15/2010	NYC Ortho	3,162		3,162.0	10:30 am
6/17/2010	Google Earth	1,815	463.8	2,278.8	10:15 am
3/26/2011	ESRI Aerial	1,032	556.2	1,588.2	12:30 pm
6/2/2011	Google Earth	2,886	17.3	2,903.3	11:45 am
7/10/2011	NAIP	1,930		1,930.0	12:30 pm
3/10/2012	Google Earth	1,589	7.4	1,596.4	11:30 am
6/15/2012 to 6/20/2012	NYC Ortho	3,590		3,590.0	11:15 am
8/6/2012	Google Earth	3,150	65.6	3,215.6	9:30 am
11/5/2012	Google Earth	1,430	701.2	2,131.2	10:30 am
5/26/2013	Google Earth	1,110	10.9	1,120.9	11:00 am
6/22/2013	NAIP	828	296.1	1,124.1	8:30 am
6/1/2014	NYC Ortho	1,929	194.0	2,123.0	11:00 am
6/19/2014	Google Earth	2,497	21.2	2,518.2	12:30 pm
7/6/2014 to 8/10/2014	Bing Aerial	2,167		2,167.0	9:15 am
9/15/2014	Nearmap	3,557		3,557.0	8:34 am to 10:09 am
10/11/2014	Google Earth	1,017	222.1	1,239.1	10:45 am
4/12/2015	Nearmap	2,361	55.3	2,411.3	3:11 pm to 3:30 pm
5/22/2015	NAIP	2,817	303.3	3,120.3	9:45 am
6/22/2015	Nearmap	3,137		3,137.0	10:10 am to 10:38 am
9/6/2015	Google Earth	987	14.3	1,001.3	11:30 am
9/6/2015	Nearmap	1,989	174.9	2,163.9	12:43 pm to 1:01 pm
10/26/2015	Nearmap	2,229	94.8	2,323.8	11:15 am to 11:34 am
3/26/2016 to 4/5/2016	NYC Ortho	3,867		3,867.0	12:45 pm
4/16/2016	Nearmap	1,965	114.8	2,079.8	12:00 to 12:18 pm
6/24/2016	Google Earth	2,270	109.9	2,379.9	11:30 am
6/25/2016	Google Earth	2,132	221.6	2,353.6	11:30 am
7/16/2016	Nearmap	1,879	50.5	1,929.5	11:04 am to 11:36 am
10/15/2016	Nearmap	2,437	111.1	2,548.1	11:18 am to 11:26 am
4/9/2017	Nearmap	2,107	82.5	2,189.5	12:56 pm to 1:04 pm

Appendix D Evidence of the slowdown in New York City streets and highways (for online publication)



(a) Medians, 342 pairs between 19 midtown zones

(b) 10th perc., 72 pairs to 4 zones on East River

(c) 90th perc., 72 pairs to 4 zones on East River

(d) Densities over 342 pairs for June 2016/June 2013 change ratios

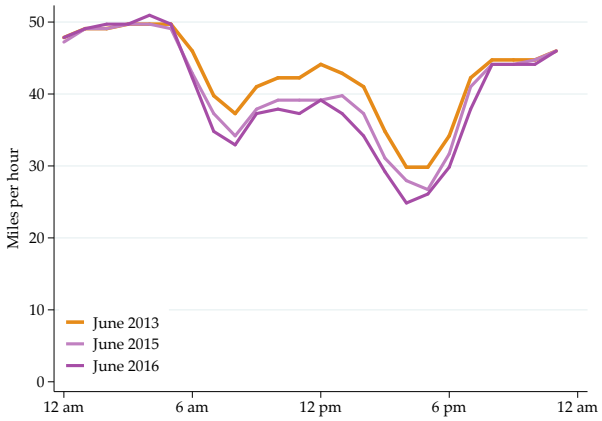
Figure D.6: Travel times (percentiles of minutes per mile) for origin-destination pairs in midtown Manhattan, month of June 2009 to 2016

Note: All figures are constructed from raw coordinates on TPEP trip data matched to "taxi zones," during the month of June from 2009 to 2016. The measures plotted are quantiles of the minutes per mile on the taxi meter. "Taxi zones" are a neighborhood-level unit of spatial aggregation defined by the TLC and at which for-hire-vehicle bases (e.g. livery, ridehail) are required to report trip originations. Midtown Manhattan between 14th and 59th streets contains 19 taxi zones. Densities in Figure D.6d are weighted by trips in June 2013.

Figure D.7: Historical speeds, archived from live EZ-Pass traffic sensor feeds

(a) Speed deciles, by month and across links

(b) Median speed, by hour and across links, selected months

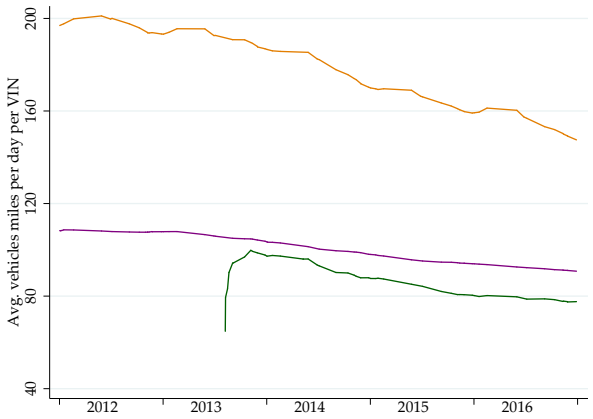


(c) Median link speed for June 2013

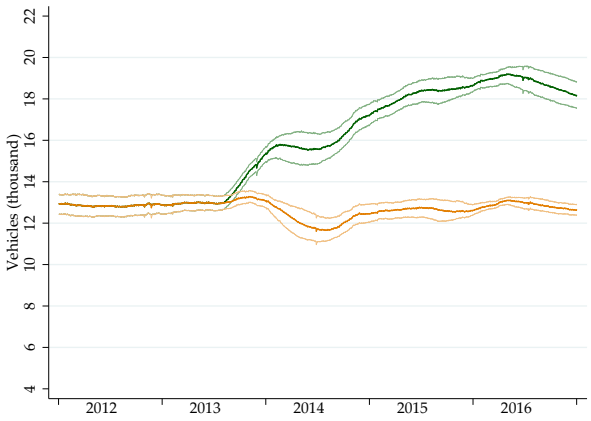
(d) Median link speed for June 2016

Note: Speeds are obtained from archived feeds of live sensor data from the NY State EZ-Pass traffic monitoring system. Sources are described in Appendix A. Fig. (a) plots the time series for deciles of all EZ-pass traffic sensor speed measurements (10th to 90th percentile, in miles per hour) and across all links in a given month (subselections of this data, such as weekdays during daytime, display similar patterns). Percentiles are plotted as lines from May 2013 through August 2013 and from April 2015 onwards, which are the months for which continuously archived data is available. Black circles plot the median speed across links for single snapshots of the data that were archived by archive.org, for nine additional months in which some of this data is available. Fig. (b) plots median speeds across links per hour for the month of June in 2013, 2015 and 2016. Figures (c) and (d) map the median speeds at the link level for June 2013 and June 2016, respectively. On both maps the speed levels are displayed on a color ramp that corresponds approximately to deciles of the speed distribution in June 2013, as plotted in the time series shown in fig. (a). The network of road segments linked through EZ-pass sensors has extended over time, the maps in figures (c) and (d) plot the 115 segments that were active in June 2013 and June 2016. Links contain overalapping segments, which are shown by means of transparencies.

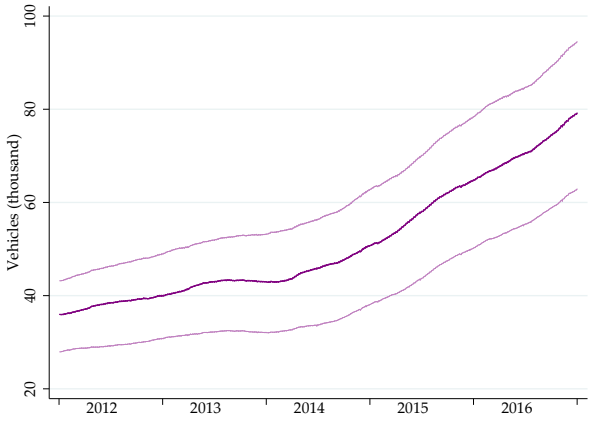
Note: methodology is described in Appendix F. In particular, note that plotted bounds arise from limiting methodological assumptions, and are not confidence intervals. Figure (a) plots the daily average of interpolated miles per day per VIN for each class of vehicles, with orange for taxis, green for boro taxis and purple for all for-hire vehicles. Miles per day per VIN require two odometer inspections around the interpolation date. Figures (b)-(c) plot the estimated stock of registered vehicles per day, and figures (d)-(e) plot an estimate of vehicle-miles traveled per day, i.e. the product of the vehicle stock times the daily average flow of miles per day per VIN for each class. In figures (b) and (d) the color yellow indicates yellow taxis, whereas green indicates yellow plus boro taxis.



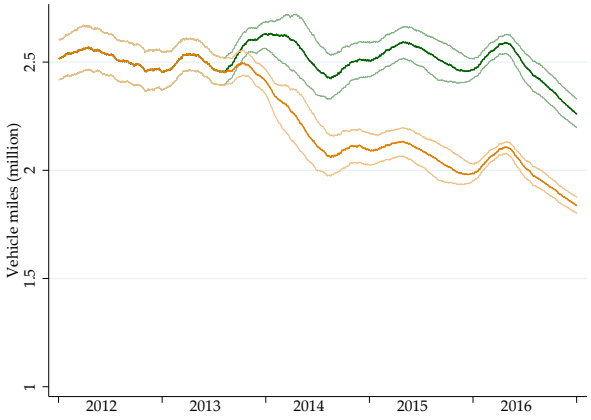
(a) Avg. miles per day per VIN and class interpolated from odometer readings



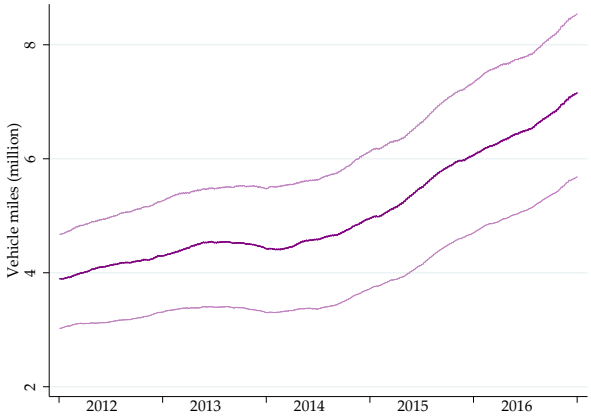
(b) Stock of registered yellow and boro taxis



(c) Stock of registered for-hire vehicles



(d) Yellow and boro taxi total VMTs per day



(e) For-hire vehicle total VMTs per day

Figure D.8: Vehicle miles traveled (VMTs) by vehicle class, estimated from odometer data

Appendix E GPS data validation (For online publication)

This paper employs geographic coordinates of taxi pickups and drop-offs recorded by GPS units onboard taxis (TPEP units) to construct historical speed measures and to track the location of taxi activity. In this appendix, we evaluate the robustness of our identification strategy (in particular, our reduced form estimates) to measurement error in these geographic coordinates. Although there is no obvious direction of bias, GPS accuracy in our area and period of study was subject to spatial and time trends that could, in principle, act as a confounder.

The accuracy of navigation satellite systems depends on the number of signal-emitting satellites within line-of-sight of the receiver, and tall buildings may create "urban canyons" that reduce the number of satellites in view. In northern Manhattan the height and density of buildings increases in the downtown direction. Since the height of buildings surrounding a road segment rarely changes, GPS accuracy could be a time-invariant attribute of road segments that can be controlled for with road segment fixed effects (i.e. run-block f.e. in the notation used in the paper). However, the number of "GPS satellites" that at any given moment are directly overhead and available to receivers in New York City has increased substantially in the past decade with the deployment of satellite systems from Russia, the European Union and China.[58] If trends in GPS satellite availability have a differential impact on coordinate accuracy that depends on the heights of surrounding buildings, GPS accuracy could follow differential trends across the boro and "hail-exclusion" zones.

We employ two measures of potential satellite obstruction throughout northern Manhattan, illustrated in Figure E.9: the average number of floors of the buildings that overlap a road segment in the area of interest (floor numbers are color-coded in blue), and a measurement of sky-view factor (SVF), which is the share of sky in a hemispherical field of vision from the ground. We employ SVF measures from Liang *et al.* (2017), which were computed from Google Street View imagery at the midpoint of every city block in Manhattan. Figure E.9 illustrates that building height (and sky-view-factor) are correlated with latitude in northern Manhattan.[59] Table E.5 estimates the baseline reduced form result from Table 3, Col. 2 while including controls for trends in GPS accuracy. Col. 1 includes the average number of completed floors for the buildings with a footprint that within each block on every avenue. This control has no impact on the coefficient of interest, which is not surprising since it contains almost no variation over time and all of its cross-sectional variation was previously absorbed by run-block fixed effects. Col 2. interacts quartiles of the previous measure with linear trends, to allow for arbitrary differential linear trends over ranges of building heights. Col. 3 interacts linear trends with quartiles of street-view-factors, where values for blocks in which Liang *et al.* (2017)

[58] GPS is a name that is specific to the United States government's global navigation satellite system, which was fully deployed to its current constellation size (32 satellites) and coverage level prior to 2009. The satellite constellations of Russia (GLONASS), the EU (Galileo) and China (BeiDou2) have been expanded between January 2009 and June 2016, with net gains of 7, 9 and 19 satellites respectively. The satellites of these systems emit signals that are available to consumer "GPS" units, and the expansion has allowed for significant accuracy improvements worldwide.

[59] Park Avenue is an exception, as it is separated by elevated tracks north of East 101st Street.

did not calculate a factor are interpolated from neighboring measures: the estimated coefficient barely changes. Results from Cols. 2 and 3 suggest that any differential trends in GPS accuracy by level of urban density did not have an impact on the measurement of street-speed that correlates in space and time with the roll-out of the boro program.

We next employ a direct measure of taxi GPS accuracy: we filter all the taxi pickup or drop-off coordinates in our area of interest through the building footprints in the area. Although coordinates may be in error even if recorded at a street curb, one type of error that we can measure over space and time is the share of trip endpoints that were placed by the taxi's GPS unit within the walls of an existing structure. Using this definition of GPS accuracy we do find that accuracy evolved differently over the treatment and control zones: 8.8% of trip endpoints in the hail-exclusion zone, and 5.4% in the boro zone were located within a building footprint in the 56 months leading up to August 2013, whereas fail rates declined to 7.2% and 2.6%, respectively, in the 34 months from September 2013 onward. Apparently, the deployment of additional satellites has had a larger effect on the accuracy of GPS fixes at an intermediate rather than at the higher end of urban density. Columns 4 and 5 include levels and logs of this GPS failure rate. Once again, our baseline estimate of the impact of the boro program is not sensitive to the inclusion of this control for GPS accuracy.

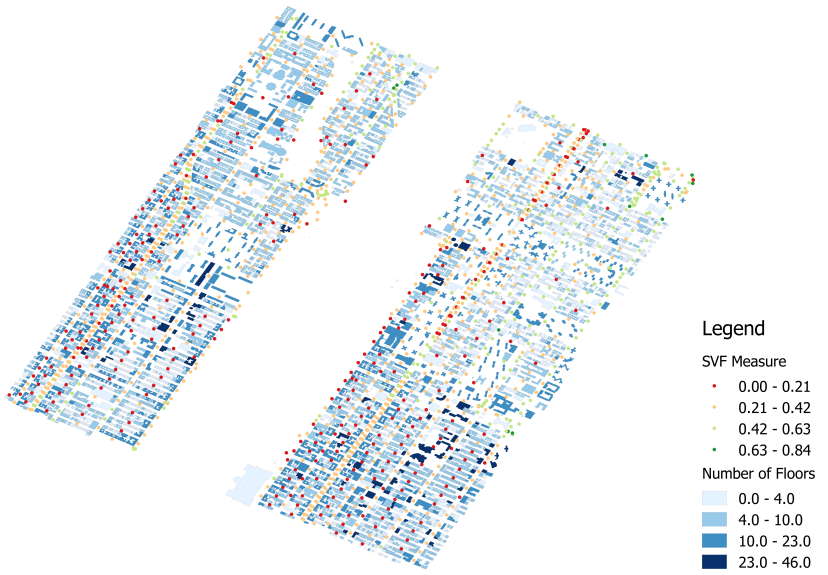


Figure E.9: Building footprints, number of floors and sky-view factors

Note: Sky-view factor (SVF) is the share of the vertical field of vision that is composed of sky. SVF measures were computed by Liang *et al.* (2017) from Google Street View imagery. Building footprints and number of floors are from data.cityofnewyork.us/Housing-Development/Building-Footprints/nqwf-w8eh.

Table E.5: Robustness of reduced form impact to measurement error in GPS units

	(1) Log avg. s/m (No controls)	(2) Log avg. s/m (Road controls)	(3) Log avg. s/m (Comm. zoning)	(4) Log avg. s/m (All)	(5) Log avg. s/m
BZ x Roll-out	0.078	0.078	0.077	0.080	0.079
	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)
Avg. number of floors	-0.000				
	(0.000)				
GPS fail rate				0.118	
				(0.025)	
Log GPS fail rate					0.002
					(0.001)
Standard controls	Y	Y	Y	Y	Y
Run-block × Run-my FE	Y	Y	Y	Y	Y
Observations	47430	47430	47430	47430	47430
R^2	0.97	0.97	0.97	0.97	0.97
R^2-within	0.14	0.14	0.14	0.14	0.14

Note: Unless otherwise specified, robust standard errors are two-way clustered at the level of run× block and run× month. Each observation is a run-block-month. In Col. 3, street-view-factors are interpolated from neighboring measures for blocks without data from Liang *et al.* (2017). Cols. 2 and 3 include interactions of GPS accuracy determinants with linear trends; interactions with the *post* dummy or the *Boro roll-out* trend also have no effect on the coefficient of interest.

Appendix F Method for estimating vehicle miles traveled from odometer inspection data (For online publication)

In this appendix we explain our methodology for calculating the total number of vehicle-miles traveled (VMTs) per day for each of yellow taxis, boro taxis and for-hire vehicles (FHVs). FHVs include traditional black cars, ridehail providers and other car services. We employ odometer readings per vehicle identification number (VIN) from 2010 through 2019, measured at regular inspections required by the Vehicle Inspection Program of New York City's Taxi and Limousine Commission and obtained through freedom-of-information requests.

Vehicles in each class are required to undergo regular inspections: within 3 months for yellow taxis, 6 months for boro taxis and 2 years for FHVs. An empirical challenge in estimating time series of VMTs from regular odometer inspections is that the data is naturally censored. Vehicles require an inspection to enter car service, but may exit car service at any point after an inspection, and this exit will not be observed until the vehicle fails to show up for an inspection within the required window. To address this missing data problem, we separately estimate daily flows of VMTs from the daily stock of vehicles on the road. Estimating average mile flows conditional on the set of vehicles that are confirmed survivors (i.e., are between inspections) is straightforward: we interpolate the odometer readings for each survivor to obtain daily flows, and average these daily flows over all survivors.

Accounting for censoring is more involved. We cannot estimate the timing of exit because we do not observe exit dates on any part of our sample; we only observe whether vehicles fail to show within the re-inspection window. Using the subsample of VIN-inspection dates for which a subsequent re-inspection or exit has occurred with certainty, we estimate the 2-year survival rate for each interval of 50,000 miles on the odometer (through 400K miles, all higher mileages are binned together). We then compute the stock of vehicles on the road as of date t as the sum of confirmed survivors (i.e. the VIN is re-inspected after t) plus an estimate of expected survivors, which consist of expected confirmed exiters (since exit date is not identified, we assume it is equiprobable over the inspection window) and expected uncertain exiters (we assume the constant hazard rate that accumulates to the estimated 2-year survival rate). We then compute our baseline daily VMT estimates as the product of daily average VMT flows per VIN and the expected stock of active VINs.

Out of necessity, our approach makes two strong assumptions: equiprobable daily exit rates for confirmed exiters and constant hazard rates for uncertain exiters. We consider two alternative assumptions: we obtain an upper bound estimate by assuming that all vehicles survive until we observe them to miss a re-inspection date, and a lower bound by assuming exit immediately after the last inspection.

Our estimate of daily VMT flows per vehicle, as well as baseline and bound estimates for stocks and total VMT flows are plotted in Figure D.8. Note that shorter re-inspection windows for yellow taxis result in tighter estimates. We start our plots on January 1, 2012, since a symmetric source of uncertainty over entry occurs at the start of our sample period.

Appendix G Figures and tables (for online publication)

Table G.6: Program impact coefficient by time slice

	24 hrs.	6am-8pm	6am-10am	10am-2pm	2pm-4pm	4pm-8pm	8pm-11pm	11pm-6am
Full week coef.	0.078	0.077	0.067	0.087	0.051	0.093	0.078	0.080
Std. err.	(0.005)	(0.006)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.008)
Avg. trips per block-month	1954.6	1491.4	304.2	427.9	247.3	512.0	253.3	209.9
Weekday coef.	0.074	0.071	0.061	0.076	0.042	0.096	0.075	0.065
Std. err.	(0.006)	(0.006)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.008)
Avg. trips per block-month	1399.4	1092.6	239.8	289.0	180.6	383.1	183.3	142.5
Weekend coef.	0.070	0.068	0.046	0.071	0.061	0.073	0.093	0.104
Std. err.	(0.006)	(0.006)	(0.009)	(0.007)	(0.009)	(0.007)	(0.009)	(0.009)
Avg. trips per block-month	555.2	398.8	64.4	138.8	66.7	128.9	70.0	128.2

Note: This table reports variations on the the baseline program impact regression (i.e. Column 2 of Table 4) in which the average second-per-meter outcome measure is computed under alternative time slices of the within-run trip data sample. The table is ordered in rows for day-of-week and columns for time-of-day slices, and the top left cell is the same specification as Table 4, Col. 2. Each cell reports estimated coefficient and standard errors for the boro-zone times roll-out interaction, as well as the average number of trips used in computation of the average travel time for each run-block-month cell.

Table G.7: Wald estimates of the congestion elasticity of taxi pickups

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Log avg. s/m	Log pickups	Log avg. s/m	Log avg. s/m	Log pickups	Log avg. s/m	Log avg. s/m	Log pickups	Log avg. s/m
	(OLS)	(1st stage)	(2SLS)	(OLS)	(1st stage)	(2SLS)	(OLS)	(1st stage)	(2SLS)
Log pickups	0.041		0.096	0.049		0.109	0.024		0.070
	(0.003)		(0.009)	(0.004)		(0.011)	(0.006)		(0.015)
BZ x Roll-out		0.818			0.910			0.666	
		(0.048)			(0.060)			(0.078)	
Bikelane	0.029	0.073	0.017	0.030	0.097	0.007	0.021	-0.022	0.024
	(0.006)	(0.060)	(0.008)	(0.008)	(0.091)	(0.012)	(0.007)	(0.048)	(0.007)
Citibike station	-0.018	-0.059	-0.005	-0.026	-0.012	-0.010	-0.007	-0.108	0.002
	(0.006)	(0.047)	(0.008)	(0.003)	(0.056)	(0.006)	(0.007)	(0.066)	(0.010)
311 complaints	0.001	0.010	-0.000	0.001	0.013	-0.000	0.000	0.008	-0.001
	(0.001)	(0.008)	(0.001)	(0.001)	(0.010)	(0.002)	(0.002)	(0.013)	(0.002)
Potholes	-0.001	0.003	-0.001	-0.004	0.017	-0.004	0.002	-0.011	0.002
	(0.001)	(0.007)	(0.001)	(0.001)	(0.012)	(0.002)	(0.001)	(0.005)	(0.001)
Run-bin and Run-m-y FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sample	All	All	All	East side	East side	East side	West side	West side	West side
Observations	47430	47430	47430	24660	24660	24660	22770	22770	22770
R^2	0.97	0.98		0.97	0.98		0.96	0.98	

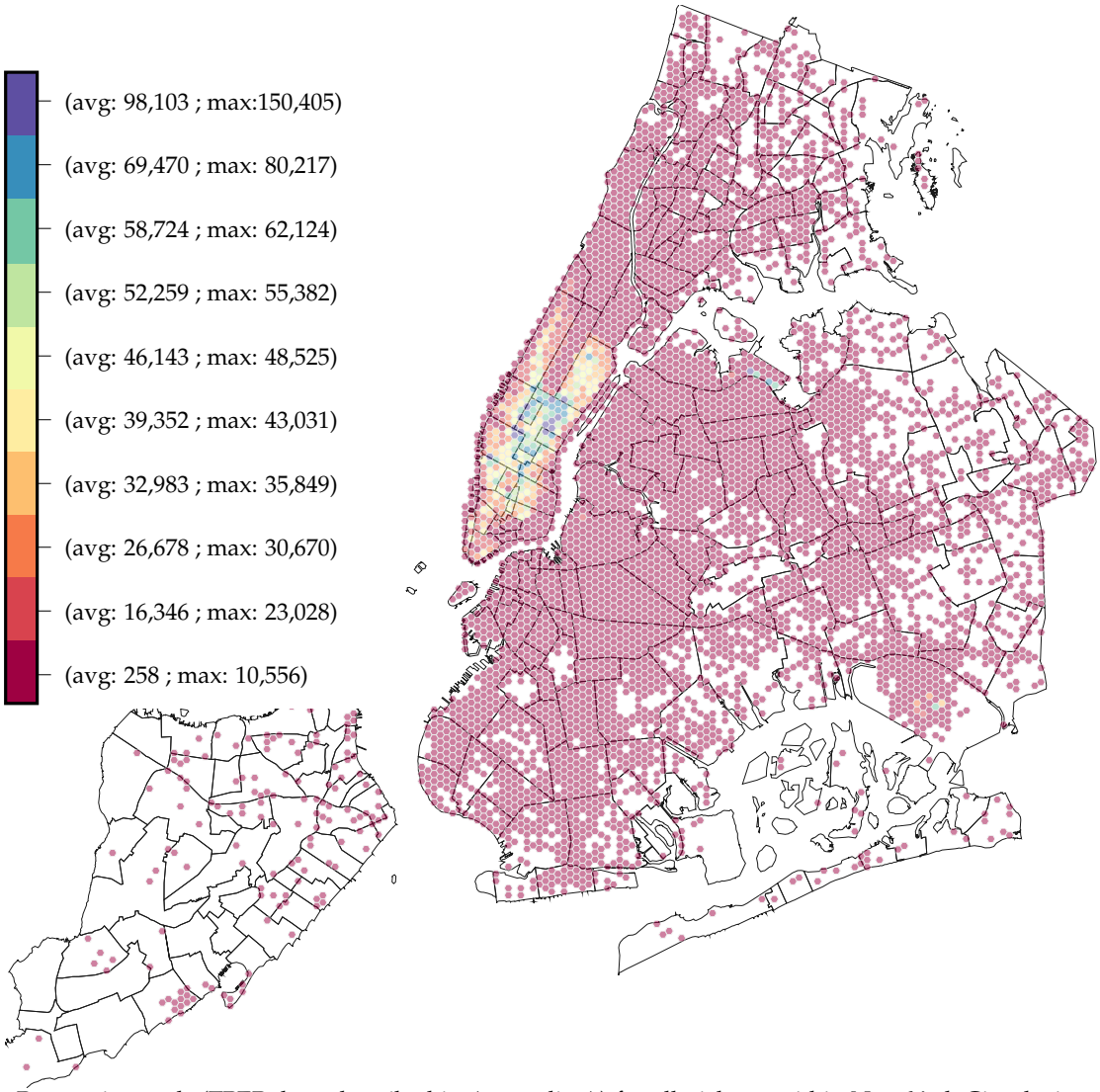
Note: An observation is a run-block-month. Dependent variable is log seconds per meter, estimated by average speed method, except for columns 2, 5 and 8 where it is log taxi pickups. Unless otherwise specified, robust standard errors are two-way clustered at the level of run$\times$ block and run$\times$ month. Weak instrument tests are not reported, but the lowest F-stat in the table (for column 9, West side) is 9,270.5.

Table G.8: Wald estimates of the congestion elasticity of taxi pickups. Spline speeds

	(1)	(2)	(3)	(4)	(5)	(6)
	Log Bsp. s/m	Log Bsp. s/m	Log Bsp. s/m	Log Bsp. s/m	Log Bsp. s/m	Log Bsp. s/m
	(OLS)	(2SLS)	(OLS)	(2SLS)	(OLS)	(2SLS)
Log pickups	-0.013	0.086	-0.017	0.119	0.003	0.037
	(0.018)	(0.030)	(0.025)	(0.042)	(0.016)	(0.035)
Bikelane	0.009	-0.012	-0.028	-0.081	0.074	0.076
	(0.047)	(0.052)	(0.071)	(0.083)	(0.042)	(0.041)
Citibike station	0.011	0.035	0.013	0.051	0.009	0.016
	(0.021)	(0.022)	(0.035)	(0.039)	(0.007)	(0.011)
311 complaints	-0.005	-0.007	-0.006	-0.009	-0.005	-0.006
	(0.010)	(0.010)	(0.014)	(0.015)	(0.011)	(0.011)
Potholes	0.002	0.001	-0.015	-0.016	0.021	0.021
	(0.007)	(0.007)	(0.012)	(0.013)	(0.007)	(0.007)
Run-bin and Run-m-y FE	Y	Y	Y	Y	Y	Y
Sample	All	All	East side	East side	West side	West side
Observations	47430	47430	24660	24660	22770	22770
R^2	0.69	-0.0065	0.57	-0.010	0.75	0.00073

Note: An observation is a run-block-month. Dependent variable is log seconds per meter, estimated by B-spline method. Unless otherwise specified, robust standard errors are two-way clustered at the level of run$\times$ block and run$\times$ month. Weak instrument tests are not reported, but the lowest F-stat in the table (for column 6, West side) is 9,270.5.

Figure G.10: Spatial distribution of yellow medallion taxi pickups. June 2013.



Note: Raw taxi records (TPEP data, described in Appendix A) for all pickups within New York City during June 2013, binned into 12,004 hexagonal cells (each with an area of 0.081 sq. km, or 19.8 acres). Every color shade accounts for approximately one tenth (1.4 million) of all taxi pickups. For example, the 14 hexagonal cells in purple accounted for 9.73% of all pickups in June 2013, with an average 98,103 pickups per cell, a volume of pickups that is equal to that of the area shaded in red (average 258 pickups per cell) and white (no pickups). Numbers on the color bar indicate the average and maximum number of pickups over the cells in a given color.