Effective and Equitable Congestion Pricing: New York City and Beyond

Michael Ostrovsky
Stanford University and NBER
ostrovsky@stanford.edu

Frank Yang*
University of Chicago
franky@uchicago.edu

November 12, 2024

Abstract

In this paper, we argue that the New York City congestion pricing plan whose implementation was paused in the summer of 2024 had a major shortcoming: as designed, it would have had a much more severe impact on the drivers of personal vehicles than on the passengers of taxis and ride-hailing vehicles, as well as on the clients of delivery services. In addition to being inequitable, this shortcoming would have likely made the congestion pricing scheme ineffective at solving the traffic congestion problem, due to the fact that the drivers of personal vehicles constitute a minority of traffic in the congestion pricing zone. We propose a simple modification to the scheme that addresses this shortcoming.

*We thank the National Science Foundation (grant #1824317) and the Business, Government & Society Initiative at the Stanford Graduate School of Business for financial support. We are grateful to seminar participants at Stanford, NYU, the San Francisco Bay Area IOFest, and the NBER Workshop on the Economics of Transportation in the 21st Century for helpful comments and suggestions.

1 Introduction

On June 30, 2024, the City of New York was scheduled to launch its congestion pricing system, following a series of earlier failed initiatives, decades of political and legal debates, and hundreds of millions of dollars of investment in supporting infrastructure. Just weeks before that anticipated historic event, on June 5, 2024, New York State governor Kathy Hochul announced an “indefinite pause” to the plan[1], to the dismay of its supporters who viewed this as a major setback not just for New York City itself, but also for numerous other regions suffering from traffic congestion. In the words of a former executive of New York’s Metropolitan Transportation Authority (MTA), “Kathy Hochul didn’t just kill congestion pricing in New York, she killed it across the country for a long time.”[2]

In this paper, we argue that this pause to the implementation of congestion pricing in New York City may have been a blessing in disguise, because the delay makes it possible to address a major shortcoming of the plan’s design. Moreover, to the extent that the congestion pricing plan implemented in New York City would serve as an example and a blueprint for future congestion pricing plans in other cities in the United States and the rest of the world, it is especially important to “get it right.”

The core parts of New York’s paused plan were as follows. Cars entering the Central Business District (CBD, the part of Manhattan south of the 60th street) during peak hours (5am–9pm on weekdays and 9am–9pm on weekends) would be charged $15 for crossing into the congestion pricing zone, once per day. Larger vehicles (excluding transit or commuter services) would be charged $24–$36, depending on their sizes. Toll rates during overnight off-peak hours would be reduced by 75%. Passengers of taxis and app-based for-hire vehicles (FHVs) would be exempted from the $15 charge, and would instead pay an additional surcharge of $1.25 for taxis and $2.50 for FHVs, 24 hours a day, on any trip to, from, or within the CBD. As explained in the recommendation by the Traffic Mobility Review Board (2023), on which New York’s congestion pricing plan was based, “The recommended tolls of $1.25 for yellow taxis, green cabs and black cars, and $2.50 for app-based FHVs, were determined by dividing the daily base auto toll rate ($15) by the average number of trips that taxis and FHVs make in the CBD today, which are 12 and 6, respectively” (p. 23).

As we argue in this paper, the key shortcoming of this plan is that as designed, it charges different types of travelers dramatically different amounts per “unit of congestion” that they generate. Commuters who drive into the CBD in the morning, work during the day, and drive home in the evening in effect pay $7.50 per trip ($15 daily fee spread over two trips). Taxi passengers pay only $1.25 per trip—one sixth of that amount—even if one ignores the fraction of the time that taxis create congestion when driving empty. And an Amazon, FreshDirect, UPS, or FedEx delivery vehicle that might enter the CBD before 5am to avoid paying the full peak-hour toll would pay the off-peak rate of $6 (25% of the $24 regular rate for package delivery trucks) that would cover a full day of driving around CBD delivering packages, inducing vastly

[1] https://www.governor.ny.gov/news/what-they-are-saying-governor-hochul-announces-pause-congestion-pricing-address-rising-cost

[2] https://www.nytimes.com/2024/06/16/nyregion/congestion-pricing-delay-effects.html

more congestion than one commuter or taxi trip.[3]

These dramatic differences between "per-unit-of-congestion" prices for different types of travelers are undesirable for two reasons. First, they are patently inequitable. It is not clear why a nurse from Queens driving to work in Manhattan should pay six times as much as a business traveler taking a taxi there from La Guardia, or one or two orders of magnitude more than a customer having a truck drive to his or her house to have an Amazon package delivered. The basic economic logic of Pigouvian taxation says that these prices should be the same (at least if they cover similar distances over similarly congested areas in Manhattan), and while in principle there may be other reasons for price discrimination or redistribution based on various group characteristics (e.g., income levels), it is not at all clear that they would align directionally with the price differences above, and even with such considerations, the magnitudes of suitable "per-unit" price differences are unlikely to be so dramatic.

Second, and perhaps more critically, such a toll scheme is unlikely to be effective at solving the traffic congestion problem. The key observation behind this second claim is that the private vehicles that would be substantially affected by the scheme constitute only around one third of all vehicles traveling in the CBD. Traffic Mobility Review Board (2023, p. 11) reports that private autos and motorcycles are responsible for only 35% of traffic in the CBD, while taxis and FHVs are responsible for 52%, with the rest being split approximately equally among buses, trucks, and commercial vans. Similarly, a recent study conducted by the former Commissioner of NYC's Department of Transportation Lucius Riccio estimates that of the 2,000 vehicles recorded in randomly selected locations in Midtown New York City, only 32.7% were personal cars, while 50.5% were taxis and FHVs.[4] So even if the scheme is effective at reducing the number of private vehicles in Manhattan, it will barely affect the remaining two thirds of the traffic, and is thus unlikely to meaningfully relieve traffic congestion (in fact, with less congestion from private vehicles, the number of miles driven by taxis, FHVs, and delivery vehicles may increase). This observation is consistent with the experience of London's congestion pricing scheme, which exempts taxis from the toll and charges FHVs only once per day, thus making a "per unit of congestion" toll for such vehicles minimal and ineffective.[5] Despite raising the daily fee on private vehicles from £5 at the scheme's introduction in 2003 to its current rate of £15, the city still suffers from heavy traffic congestion. E.g., a 2024 study by the navigation company TomTom has awarded London the dubious honor of being the most congested city in the world (out of 387 for which TomTom has collected traffic data).[6] Of course, it is possible that without London's congestion pricing scheme, traffic there would have been even worse, but it is nevertheless safe to say that the scheme has failed to solve London's traffic congestion woes.

So is there a solution? To a properly trained economist, the answer appears immediate and long-

[3] A typical delivery driver makes 100–200 stops per day (`https://www.aboutamazon.com/news/transportation/photos-day-in-the-life-amazon-delivery-driver`, `https://www.businessinsider.com/im-a-ups-delivery-driver-what-my-job-is-like-2022-2`, `https://www.quora.com/How-many-stops-does-a-UPS-DHL-FedEx-delivery-truck-make-daily-in-NYC-With-how-many-packages-avg`).

[4] `https://www.gothamgazette.com/city/11828-congestion-pricing-for-hire-vehicles-midtown-traffic`

[5] `https://tfl.gov.uk/modes/driving/congestion-charge/discounts-and-exemptions`

[6] `https://www.tomtom.com/newsroom/explainers-and-insights/tomtom-traffic-index-2024-london-is-slowest/`

established: Pigouvian taxation, going back to Pigou (1920). The regulator should simply charge each vehicle for the externality it imposes on others, and the problem is solved. However, a moment's reflection makes it clear that such a scheme, taken literally, is too complex and impractical. First, the externality varies by the specific location in the CBD, the specific time and day of the week, weather conditions, traffic conditions on the neighboring roads, traffic accidents and special events, vehicle size, and so on. This leads to a massive, highly multidimensional space of prices, which is both very hard to compute and hard to communicate to the drivers. Second, imposing such a pricing scheme would require tracking all cars in real time, which raises potential privacy issues and imposes additional technological implementation costs. While not insurmountable on its own, this issue adds to the political costs of congestion pricing that are already very substantial.

In this paper, we propose a congestion pricing scheme that is equitable (different vehicle types pay tolls commensurate with the externalities that they impose) and practical (technologically, it is only a minor deviation from New York's "paused" congestion pricing plan, and it can be launched quickly). The scheme can be viewed as an approximation of the idealized Pigouvian pricing discussed in the previous paragraph, and is based on two observations. The first observation, which we have already mentioned, is that personal cars are responsible for only around a third of traffic in the CBD, while taxis, FHVs, delivery trucks, and other commercial vehicles are responsible for the remaining two thirds. The second observation is that driving patterns are very different between typical personal and commercial vehicles. A typical commuter in a personal vehicle drives to the CBD in the morning, spends the day at his or her workplace, and then drives back home in the evening. So the amounts of congestion caused by different commuters are comparable: "All happy families are alike." By contrast, different commercial vehicles have dramatically different driving patterns inside the CBD, with some circling throughout the CBD for much of the day (e.g., delivery vehicles) and others driving through it several times per day (e.g., Uber drivers) on trips of varying lengths: "Each unhappy family is unhappy in its own way."

These observations lead us to the following proposed congesting pricing scheme. For personal vehicles, the scheme is the same as the "paused" plan: such vehicles will pay a certain daily amount for entering the CBD (as discussed below, given the other modifications in the plan, this amount can be lower than in the "paused" plan while achieving similar revenue and traffic reduction goals). Given their minority share in the overall traffic volume and the relative similarity of their driving patterns, this is a reasonable approximation of the "idealized" Pigouvian taxation, and has the additional practical benefits of being very simple to communicate and not requiring tracking the vehicles throughout the CBD. For commercial vehicles, however, the scheme is different. Given the dramatic differences in their driving patterns, the scheme charges these vehicles on a per-trip or per-mile basis, during peak times. In Section 3, we discuss specific calculations for the per-day tolls for personal vehicles and the per-trip or per-mile tolls for commercial vehicles, based on the principle that on average, the different types of vehicles should pay comparable amounts per unit of congestion that they create. Note that introducing per-trip or per-mile tolls is much easier for commercial vehicles than for personal ones. FHVs and taxis already track trips and mileage on passenger trips, so no additional infrastructure is required. (We discuss the issue of "empty miles," when these vehicles travel without passengers, in Section 2.) For delivery vehicles (Amazon, FedEx, UPS, etc.), their movements are likewise tracked in real time by the companies (e.g., Amazon, FedEx, and UPS have customer apps that display real-time locations of their deliveries), and thus the information can be easily collected. The remainder of commercial vehicles constitute only a small fraction of the overall traffic, and can be required to either install a tracking device and pay per mile (possibly after some initial grace period, not necessarily immediately at the launch of the plan) or be folded under the "pay-per-day" case without substantially affecting the overall congestion levels.

The "paused" NYC congestion pricing scheme does impose fees on taxis and FHVs, but they are very different from the ones in the scheme that we propose. First, as already discussed above, they are much smaller in magnitude (roughly speaking, following the principle of "same congestion fee for the same amount of congestion created," an average taxi or FHV trip should pay a half of the daily fee paid by a commuter, i.e., $7.50 per trip rather than $1.25 or $2.50, plus, as we discuss below, an additional increment due to the "empty miles" that taxis and FHVs drive; see Section 2 for detailed calculations). Second, unlike the paused proposal, ours would charge these congestion fees only during peak travel times (Section 3.1). There is no need to charge a congestion fee to a passenger taking an Uber from a bar at 1am. By contrast, the paused proposal would charge the same congestion fees to FHVs and taxis throughout the day.[7] Finally, as we argue in Section 3.2, per-mile rather than per-trip tolls on taxis and FHVs are better aligned with the negative externalities they create: a congestion fee for a short trip of a few blocks should be substantially lower than that for a long trip that crosses the entire CBD and thus creates a lot more congestion.

The last part of our proposal does not directly affect economic incentives "on the margin," but instead highlights the point that parts of the additional revenues generated from the modified (and generally higher) tolls on taxis/FHVs and delivery vehicles can be deployed to alleviate their impact on some of the key constituents affected by those changes, in a way that (crucially!) does not "undo" their incentives that help relieve congestion. One such key constituency are the owners of taxi medallions, whose values are likely to be negatively affected by the increased tolls. This reduction in medallion values is an important concern, and a commonly voiced suggestion to address it is to exempt taxis from the congestion fee or charge them a substantially reduced one. Of course, a major problem with such an exemption is that since taxis are a substantial part of traffic in the CBD, it would severely lower the effectiveness of the overall congestion scheme (in particular, the number of taxi trips in the CBD would not just remain the same, but would likely rise, as passengers would substitute in higher numbers from FHVs to taxis; see, e.g., Leccese (2024) for evidence of this effect after such an "asymmetric" exemption in Chicago). Our proposal is to instead compensate medallion holders in a lump-sum fashion (e.g., a fixed per-medallion payment per year for a certain number of years, or a reduction in various annual fees that taxi operators and medallion owners are required to make), with the amount commensurate with the estimated negative impact on

[7] Note that taxis and FHVs already pay certain per-trip congestion fees for trips south of the 96th street. Such fees may also be in principle replaced by a time-varying fee. Our current analysis focuses only on the modifications to the most recent, paused proposal, but would be straightforward to extend to the other congestion fees.

the value of the medallion. Crucially, this approach allows regulators to make medallion holders "whole" without negatively impacting the overall effectiveness of the congestion pricing scheme. And to maintain the overall financial soundness of the congestion pricing plan, this compensation can be taken from just the overall congestion pricing fees collected from the taxi rides. We provide further details in Sections 2 and 3. Similarly, increased revenues generated from higher tolls on delivery vehicles can be deployed in a "lump-sum" fashion to expand programs like the NYC DOT's Off-Hour Deliveries Incentive Program pilot, which "provide[s] financial incentives for businesses to shift deliveries to off-peak hours" by "fund[ing] tools and strategies to make overnight deliveries feasible for businesses, [including] the installation of low-noise equipment for delivery vehicles (such as newer hand pallet trucks and backup alarms), building security retrofits to enable unattended deliveries, and safety equipment such as security cameras."[8]

The remainder of the paper is organized as follows. We begin with "back-of-the-envelope" calculations in Section 2, considering the most basic changes to the paused plan: keeping the overall structure of the plan unchanged, but changing the specific amounts of the per-day fee for personal cars and the per-trip fees for taxis and FHVs. In Section 3, we discuss two additional changes: imposing congestion tolls on taxis and FHVs only during peak traffic times and replacing per-trip tolls with per-mile ones. In Section 4, we discuss the tolling of delivery vehicles. Section 5 concludes.

2 Back-of-the-Envelope Calculations

We start out with the most basic, "back-of-the-envelope" calculations to illustrate the effects of our alternative plan and compare them to those of the baseline "paused" plan. Our model in this section is deliberately streamlined, to the very basics, to make our calculations as transparent as possible. In particular, we focus on just two parts of traffic, personal vehicles and taxis/FHVs, leaving aside commercial vans and trucks of various sizes (incorporating such vehicles would only strengthen our conclusions). We also consider a "per-trip" toll on taxi and FHV riders in the alternative plan, to make the comparison between the "paused" plan and the alternative one as direct and transparent as possible.

2.1 Data

We get the current number of daily taxi and FHV trips from the NYC Taxi & Limousine Commission's Trip Record database.[9] On average, there are approximately 235 thousand daily FHV trips and 85 thousand taxi trips that start or end (or both) in the CBD.

We assume that FHVs and taxis are occupied 50% of the time. This assumption is based on the estimates from Cramer and Krueger (2016), who report utilization rates of taxi and Uber drivers between 48.3% and 51.2% in New York City.[10]

[8]See `https://ohdnyc.com/` and `https://www.nyc.gov/html/dot/html/pr2024/reduce-truck-deliveries.shtml` for details.

[9]`https://www.nyc.gov/site/tlc/about/tlc-trip-record-data.page`

[10]Buchholz (2021) estimates that taxis in New York City are vacant 47% of the time. Castillo (2023) reports that in Houston,

To estimate the number of personal cars, we start with the New York Metropolitan Transportation Council's report (NYMTC, 2022, Tables 16 and 17), showing the numbers of vehicles (except buses) entering and exiting the CBD. Averaging the two numbers, we get 623 thousand vehicles entering/exiting the CBD per day. Next, according to a study by the NYC Department of Transportation, 56% of these vehicles are personal cars (NYCDOT, 2019, p. 16). Multiplying the two numbers, we get an estimate of approximately 350 thousand personal cars entering the CBD every day.

Thus, in the calculations below, we assume that there are approximately 700 thousand personal car trips and $(235 + 85) \cdot 2 = 640$ thousand FHV and taxi trips (including empty ones) per day in the CBD. These numbers are broadly consistent with the estimates cited in the Introduction (from the Traffic Mobility Review Board (2023) who report that private autos and motorcycles are responsible for 35% of traffic in the CBD while taxis and FHVs are responsible for 52%, and from Riccio who estimates that 32.7% of traffic is due to private autos and 50.5% is due to FHVs), but do have a relatively higher share of personal car trips than in those estimates. The discrepancy may be due to data issues, or possibly due to different average lengths of trips between personal cars and FHVs/taxis. It may also be due to the fact that cars with the special FHV license plates provide delivery services (Instacart, Amazon Flex, Gopuff, and so on) in addition to carrying passengers, and while both types of trips appear in the Traffic Mobility Review Board (2023) and Riccio data, only passenger trips are recorded in the NYC Taxi & Limousine Commission's Trip Record database. Note, however, that to the extent that our input numbers possibly overestimate the relative amount of personal car travel and underestimate the relative amount of FHV/taxi travel, our estimates about the impacts of switching from the paused plan to the alternative ones that we discuss are conservative: If the true relative amounts of personal vs. FHV/taxi traffic in the CBD are closer to the 2:3 ratio, our conclusions would become only stronger.

The next input in our calculations are the elasticities of demand for driving and hiring a taxi/FHV. For driving, Lehe and Devunuri (2022) report demand elasticities of -0.55, -0.67, and -0.53 from the introduction of cordon pricing in London, Stockholm, and Gothenburg, respectively. For FHVs and taxis, there are several recent studies estimating elasticities of demand. Castillo (2023) reports the elasticity of -0.633 from running a pricing experiment at Uber in five Latin American cities. Cohen et al. (2016) use variation in prices in Uber's four largest U.S. markets and estimate the elasticity of demand to fall between -0.4 and -0.6. Almagro et al. (2024) use price variation due to a surcharge on ride-hailing trips in Chicago and obtain the elasticity estimate of -1.42. Buchholz (2021) estimates price elasticities of demand for taxis in New York City to range between -1.074 and -2.220 (Table 7). For the calculations in Sections 2.2 and 2.3, we use the same elasticity of -0.6 for both taxis/FHVs and personal driving. In Section 2.4, we explore a range of parameter values, and show that the key conclusions are robust to the specific choice of elasticity parameters.

Uber drivers are available to be matched 45.2% of the time, are on the way to pick up a rider 17% of the time, and are actually taking a rider to the destination 37.8% of the time.

2.2 Paused Plan

We start out by calculating the estimates for the paused plan.

First, we discuss our calculations for personal cars. The nominal toll on these vehicles is $15, but due to various credits and reductions (e.g., discount for using a tolled bridge or tunnel; exemptions for some groups; lower toll during off-peak hours), we assume the effective average toll to be two thirds of the nominal one, $10. Another parameter we need for our calculations is the average base cost of driving into the city before the toll, which includes the costs of fuel, parking (which in Manhattan is substantial[11]), wear and tear of the car, and the existing tolls on some of the bridges and tunnels. We estimate this baseline number to be $30.

Assuming constant elasticity of demand, with the toll, we get the average predicted daily number of personal vehicles equal to

$$350,000\left(\frac{30+10}{30}\right)^{-0.6} = 294,513,$$

with the corresponding annual toll revenue from personal drivers equal to

$$294,513 \times 365 \times \$10 = \$1,074,973,274.$$

Next, there are 85,000 paid taxi trips per day that involve the CBD. The average price of a taxi trip is approximately $25[12]. With the per-trip toll of $1.25 and a similar calculation to the one above, we get the predicted post-toll daily number of taxi trips as

$$85,000\left(\frac{25+1.25}{25}\right)^{-0.6} = 82,548,$$

with the corresponding annual toll revenue from taxi passengers equal to

$$82,548 \times 365 \times \$1.25 = \$37,662,420.$$

Finally, there are 235,000 daily paid FHV trips that involve the CBD. The average price is $40. With the per-trip toll of $2.50, the predicted post-toll daily number of FHV trips is then

$$235,000\left(\frac{40+2.50}{40}\right)^{-0.6} = 226,606,$$

with the corresponding annual toll revenue from FHV passengers equal to

$$226,606 \times 365 \times \$2.50 = \$206,777,544.$$

[11] E.g., parking aggregator company SpotHero reports that the average parking rate for a Broadway show is $30 and the cost of a monthly parking spot is on average $19 per day (`https://spothero.com/artist/broadway`, `https://spothero.com/city/monthly/nyc-parking`).

[12] Based on the NYC Taxi & Limousine Commission's Trip Record Data for 2024.

2.3 Alternative Plans

As discussed in the Introduction, the paused plan disproportionately affects personal cars relative to the passengers of taxis and FHVs. The calculation in Section 2.2 makes this asymmetry clear: despite comparable aggregate numbers of trips, the drivers of personal cars would pay \$1.075 billion in tolls per year, while the passengers of taxis and FHVs would only pay \$244 million. The reduction in traffic is also disproportionate: the number of personal car commutes would drop by 16% (55 thousand trips per day), while the number of taxi and FHV trips would drop by only 3% (11 thousand trips). So how can the plan be made more equitable? As we discussed in the Introduction, our proposal for modifying the toll structure on taxis and FHVs involves three changes: the higher absolute level, the time-varying structure (high during peak times, zero during off-peak times), and the per-mile rather than per-trip computation of the fee. In this section, we focus just on the first of these three changes, for simplicity, transparency, and immediate comparability with the paused plan. We discuss the other two changes in Section 3.

What should the toll on taxis and FHVs be to have a comparable payment "per unit of congestion generated" to that of personal cars? Note that \$15 cordon fee pays for two trips: one trip to CBD in the morning, and one trip from CBD in the evening. So the per-trip fee is \$7.50. However, to be comparable on the "per unit of congestion" basis, the fee on taxis and FHVs needs to be higher than that, because in addition to the trips for which they are paid (and on which the congestion fee is imposed), taxis and FHVs also drive a fraction of the time empty—and that needs to be taken into account. Using the estimate of "empty miles" of 0.50 that we reported Section 2.1 (i.e., a typical taxi in the CBD is empty 50% of the time), to be commensurate with the personal car cordon fee, the per-paid-trip congestion fee on taxis and FHV should also be equal to $\frac{\$7.50}{.5} = \$15$. We call this congestion pricing scheme "Plan 1."

With this toll level, the predicted daily number of taxi trips drops to

$$85,000\left(\frac{25+15}{25}\right)^{-0.6} = 64,113,$$

while the corresponding annual toll revenue from taxi passengers increases to

$$64,113 \times 365 \times \$15 = \$351,019,352,$$

i.e., an almost tenfold increase relative to the paused plan! The number of FHV trips drops to

$$235,000\left(\frac{40+15}{40}\right)^{-0.6} = 194,127,$$

while the corresponding annual toll revenue from FHV passengers increases to

$$194,127 \times 365 \times \$15 = \$1,062,846,060,$$

i.e., a more than fivefold increase relative to the paused plan.

With the per-trip toll of \$15, both the overall toll revenue and the overall trip reduction for taxis and FHVs are commensurate with those of personal cars: both types of vehicles pay comparable tolls "per unit of congestion." However, this is only part of the story. The second observation is that, unsurprisingly, the overall toll revenue grows dramatically, by more than \$1 billion per year. This means that if the tolls on taxis and FHVs are set at the level commensurate with that of personal cars, the latter *can be substantially reduced* while the overall congestion pricing plan's revenue goals will continue to be met.[13] Our second alternative plan (which we call "Plan 2") illustrates this possibility.

Consider setting both the cordon toll for personal cars and the per-trip toll on taxis and FHVs at \$9 (instead of \$15 – a 40% reduction). Repeating the above calculations with this lower toll level,[14] we get

Number of personal cars	$350,000\left(\frac{30+6}{30}\right)^{-0.6}$	=	313,732,
Revenue from personal cars	$313,732\times 365\times \$6$	=	\$687,073,837,
Number of taxi trips	$85,000\left(\frac{25+9}{25}\right)^{-0.6}$	=	70,680,
Revenue from taxi passengers	$70,680\times 365\times \$9$	=	\$232,183,287,
Number of FHV trips	$235,000\left(\frac{40+9}{40}\right)^{-0.6}$	=	208,059,
Revenue from FHV trips	$208,059\times 365\times \$9$	=	\$683,473,346.

Note that compared to the paused plan, Plan 2 reduces the overall number of trips by a similar amount; charges a lot less to the commuters (\$9 vs. \$15); and raises substantially higher toll revenue (\$1.603 billion vs. \$1.319 billion). This increase in revenue makes it possible to meet the same congestion pricing revenue goals as what the paused plan would have achieved, *and* in addition to compensate the owners of taxicab medallions for the reduction in their value due to the introduction of congestion pricing (with 13,500 medallions, the extra revenue would be sufficient to compensate each one up to $\frac{\$1.603B-\$1.319B}{13500}=\$20,986$ per year, which would be much higher than the total current annualized value of a medallion, let alone the reduction in its value from the introduction of tolls). The regulators could also lower registration and other annual fees for taxicab drivers. Crucially, these would be fixed-type fees that would not affect the taxi drivers' marginal incentives to drive in CBD during congested times. There would also be sufficient extra revenue to increase the funding to such programs as the NYC DOT's Off-Hour Deliveries Incentive Program discussed in the introduction, whose budget is currently only \$6 million.[15]

[13] The plan's revenue goals are mandated by law; see §1704-A of the New York State Senate's Traffic Mobility Act, `https://www.nysenate.gov/legislation/laws/VAT/1704-A`.

[14] As in Section 2.2, we assume that due to various exemptions and discounts, the effective average toll on personal cars is two thirds of the nominal one, i.e., \$6.

[15] `https://www.nyc.gov/html/dot/html/pr2024/reduce-truck-deliveries.shtml`

Table 1: Alternative Elasticity Specifications

	(-0.6, -0.6)	(-0.3, -0.3)	(-1.2, -1.2)	(-0.3, -1.2)
Number of personal cars	314	331	281	331
Revenue from personal cars	\$687	\$726	\$616	\$726
Number of paid taxi trips	71	78	59	59
Revenue from taxi trips	\$232	\$255	\$193	\$193
Number of paid FHV trips	208	221	184	184
Revenue from FHV trips	\$683	\$726	\$605	\$605
Total number of trips	1,185	1,260	1,048	1,123
Total revenue	\$1,603	\$1,707	\$1,414	\$1,482
Δ trips vs. paused plan	-22	-11	-45	-90
Δ revenue vs. paused plan	\$283	\$286	\$274	\$116

Note: Numbers of cars and trips are in thousands. Revenues are in millions.

2.4 Robustness to Elasticity Assumptions

In this section, we explore the robustness of the above calculations to alternative assumptions on the values of demand elasticities. Specifically, we consider three alternative specifications: uniformly less elastic demand ($\nu_p = -0.3$, $\nu_t = -0.3$, where ν_p denotes the elasticity of demand from personal car drivers and ν_t denotes the elasticity of demand from taxi and FHV passengers), uniformly more elastic demand ($\nu_p = -1.2$ and $\nu_t = -1.2$), and less elastic demand from personal car drivers and more elastic demand from taxi and FHV passengers ($\nu_p = -0.3$ and $\nu_t = -1.2$).

Table 1 summarizes the results for Plan 2 (\$9 toll applied to both personal cars and taxi/FHV trips) for the baseline specification above with $\nu_p = -0.6$ and $\nu_t = -0.6$, along with the three alternative specifications of elasticity parameters. Note that the row "Total number of trips" adds up the number of personal cars, the number of paid taxi trips, and the number of paid FHV trips, and then multiplies the sum by 2, to account for the fact that each personal car takes two trips per day and for the empty trips by taxis and FHVs. The last two rows of the table show the differences in the total numbers of trips and the total revenues under Plan 2 vs. the paused plan.

Our main conclusions are robust across these alternative assumptions on elasticities of demand: the alternative toll plan that charges taxis and FHVs equitably in comparison to personal cars is both more effective in reducing traffic and generates higher toll revenue than the paused plan, despite charging a much lower cordon crossing fee of \$9 instead of \$15.

3 Time- and Distance-Based Tolls on Taxis and FHVs

Plans 1 and 2 discussed in Section 2 are most directly comparable to the paused plan, changing only the levels of various tolls but not their structure.

However, there are additional improvements and modifications that we can make.

First, as does the paused plan, Plans 1 and 2 charge the congestion pricing fee to taxis and FHVs 24 hours per day. When the level of those fees is low, like in the paused plan, this does not introduce a big distortion during off-peak hours. However, once the fees are set at the level of $9 or $15, the distortion (relative to the ideal Pigouvian taxes, which are close to zero when there is no congestion) becomes substantial. Moreover, it violates the "similar payment for similar amount of congestion" principle when compared with personal car drivers, who pay 75% less when they drive into the CBD during off-peak hours. Thus, one improvement to Plans 1 and 2 is to only charge the tolls on taxis and FHVs during peak hours.

Second, when the level of fees is low, they do not introduce a big distortion to demand for short vs. long trips. However, once the level of fees becomes substantial, the relative effects become very different. A $15 surcharge on a short trip that would otherwise by itself cost $15 has a very different effect vs. the same surcharge applied to a, say, $70 trip from the JFK airport to Manhattan. So this pricing structure not only violates the "similar payment for similar amount of congestion" principle, but also skews the overall passenger trip composition toward those who take long trips—and create more congestion. So another improvement to the plans is to replace the per-trip fee on taxis and FHVs with the commensurate (on average) per-mile fee, that more accurately reflects the amount of congestion caused by such trips.

In this section, we present calculations showing the effects of these changes. In Section 3.1, we keep the per-trip fee structure, but eliminate the toll on FHVs and taxis during off-peak hours. In Section 3.2, we further replace the per-trip fee structure with the per-distance one. In both sections, we focus on Plan 2 with the $9 toll.

3.1 Removing Taxi and FHV Tolls during Off-Peak Hours

In the NYC TLC Trip Record Data, approximately 20% of taxi trips and 25% of FHV trips take place during off-peak hours. Under Plan 3, we set the toll for those trips to zero, while keeping the rest the same as in Plan 2. Adjusting the calculations from Section 2.3,[16] we get

Number of taxi trips	$0.8 \cdot 85,000 \left(\frac{25+9}{25}\right)^{-0.6} + 0.2 \cdot 85,000$	=	73,544,
Number of peak-hour taxi trips	$0.8 \cdot 85,000 \left(\frac{25+9}{25}\right)^{-0.6}$	=	56,544,
Revenue from taxi passengers	$56,544 \times 365 \times \$9$	=	$185,746,629,
Number of FHV trips	$0.75 \cdot 235,000 \left(\frac{40+9}{40}\right)^{-0.6} + 0.25 \cdot 235,000$	=	214,794,
Number of peak-hour FHV trips	$0.75 \cdot 235,000 \left(\frac{40+9}{40}\right)^{-0.6}$	=	156,044,
Revenue from FHV trips	$156,044 \times 365 \times \$9$	=	$512,605,010.

Adding the $687 million revenue from personal car tolls, the total annual revenue from this modified program is $1.385 billion: lower (of course) than under full-day tolling, but still substantially higher than

[16] With time-varying tolls, these calculations implicitly assume away time substitution by FHV and taxi passengers. The magnitude of this effect is unlikely to meaningfully affect our conclusions, so we do not take it into account in our calculations.

the \$1.319 billion revenue from the paused plan. The total number of trips under this plan would be only slightly lower than under the paused plan: there would be $2 \cdot (314 + 74 + 215) = 1,204$ thousand trips under Plan 3, vs. 1,207 thousand trips under the paused plan. However, crucially, Plan 3 reduces disproportionately more traffic during peak hours, which is of course a primary goal of any congestion pricing program.

3.2 Distance-Based Tolls

The final dimension that we consider is distance-based tolling of taxis and FHVs. Setting the toll at the per-trip level disproportionately affects short-distance passengers, much more so than those who travel over a long distance, thus distorting the incentives and shifting the outcome away from the socially optimal one. In this section, we evaluate the viability of replacing the per-trip toll with a per-distance one. Specifically, in our Plan 3, the toll on taxis and FHVs was set at the \$9 level during peak hours (and zero during off-peak hours). With the average within-CBD taxi and FHV trip distance of approximately 2 miles, the corresponding distance-based toll is \$4.5 per mile (still only during peak hours). To evaluate the effect of this shift on the outcomes, we proceed as follows (Plan 4), using the NYC TLC Trip Record Data, which contains such details as the length of the trip and the fare paid.

Separately for taxis and for FHVs, we break all trips into 21 buckets of trip lengths: 0–0.5 miles, 0.5–1 mile, …, 9.5–10 miles, and more than 10 miles.[17] For each bucket, we compute the average toll and the average trip length (and thus the corresponding distance-based toll), and perform the computations similar to those above. We then aggregate the results across all buckets.

For example, there were 9,495 within-CBD taxi trips in the (0.5–1] bucket. The average length of such trips was 0.79 miles, and the average total fare was \$15.03. With the fixed \$9 toll, the number of such trips would drop to 7,166—a 25% reduction. By contrast, with the distance-based toll of 0.79·\$4.50=\$3.56, the number of such trips drops much less, to 8,365.

Conversely, there were 2,429 within-CBD taxi trips in the (2.5–3] bucket. The average length of such trips was 2.74 miles, and the average total fare was \$25.88. With the fixed \$9 toll, the number of such trips would drop to 2,031. With the distance-based toll of 2.74·\$4.50=\$12.33, the number drops to 1,922.[18]

[17] Perhaps surprisingly, there is a positive, though small, number of trips fully within the CBD that are more than 10 miles long. Less surprisingly, there are many long trips that either start or end in the CBD, but not both.

[18] One additional detail is that each bucket contains two types of trips: within-CBD trips, for which both the origin and the destination are within the CBD, and the remaining ones, for which either the origin or the destination, but not both, are within the CBD. For the former, we simply assume that the entire length of the trip is within the CBD and would be tolled at the \$4.50 per mile rate. For the latter, for trips longer than 5 miles, we assume that 2.5 miles of the trip was within the CBD (and thus would be tolled at that rate), and for trips shorter than 5 miles, we assume that half of the length of the trip was within the CBD and would be tolled. To get the cutoff value of 5 miles (or equivalently the cap of 2.5 miles in the CBD for trips with exactly one end in the CBD), we perform the following calculations. We know from the data that for a taxi or FHV trip fully within the CBD, the average distance is 2 miles. Intuitively, this should mean that for an average trip with exactly one end within the CBD, the distance that trip should cover within the CBD should be higher than 2 miles (because the distance between two randomly chosen points within the CBD is shorter than between a randomly chosen point within the CBD and a randomly chosen point on its boundary, where the trip enters or exists the CBD). To estimate the exact adjustment factor, we perform the following calculations. Take a disc or a square. For the distance, consider either the Euclidean distance or, quite fittingly, the Manhattan

Aggregating across all the buckets and both taxis and FHVs, we find that the overall toll revenue from taxi and FHV passengers would remain virtually unchanged if we switched from the fixed-toll Plan 3 to the distance-based Plan 4 ($679,734,078 vs. $680,219,877). The total number of paid daily trips increases from 206,921 to 211,573 (vs. 242,592 without any tolls). By contrast, the total number of *miles* on such trips *decreases*, from 1,034,986 to 1,023,165 (vs. 1,146,630 without any tolls). Thus, switching from the fixed to the distance-based toll plan dampens the reduction in the number of FHV and taxi trips by 13% while simultaneously boosting the reduction in the total distance of such trips by 11%, further easing the overall traffic congestion while continuing to meet the program's revenue goals.

4 Delivery Vehicles

Aside from personal vehicles and taxis/FHVs, another substantial source of traffic in Manhattan are commercial trucks, vans, and other vehicles, in particular those performing delivery services. The reason for singling out delivery services (over other commercial vehicles—e.g., a plumber traveling to a job site) is that by the nature of its service, a delivery vehicle generates a disproportionate amount of traffic congestion per day (while, to continue with the example, the plumber only uses the vehicle to get to the job site, and then spends a considerable amount of time actually performing the plumbing service while the vehicle is parked and is not creating any congestion).

For delivery trucks, the paused proposal would charge $24 or $36 daily toll (depending on their size), with a 75% discount if they enter the CBD before the beginning of peak tolling period (5am on the weekdays and 9am on the weekends). As discussed in the Introduction, the problem with this approach is that with the fee being split across many deliveries, it is inequitable (compared to the per-unit-of-congestion toll paid by personal car drivers) and is unlikely to be effective. Moreover, to the extent that it tries to give incentives to the delivery vehicles to shift deliveries to off-peak hours, it would only succeed in shifting the *initial* deliveries of a given truck to the pre-peak hours: a truck that enters the CBD at 4:50am on a weekday and performs the deliveries (and thus creates congestion) during much of the peak 5:00am–9:00pm time window would still pay the dramatically reduced toll.

While in principle, a delivery vehicle may create more congestion per mile than a car (due to its larger size and other related factors), at least in the initial implementation, keeping the toll for such vehicles at the same per-mile level as for taxis and FHVs (computed in Section 3.2) would keep the scheme simple, equitable, and transparent. The specific numbers can be iteratively fine-tuned by policymakers at a later date, perhaps for some subsets of the vehicles (e.g., tractor-trailers that are particularly detrimental to traffic).[19]

distance (`https://en.wikipedia.org/wiki/Taxicab_geometry`). For each of the four cases, suppose the size of the underlying geometric object is such that the distance between two randomly chosen points within it is 2 miles. What is then the distance between a randomly chosen point within the object and a randomly chosen point on its boundary? In all four cases, the answer is in the 2.47–2.5 miles range, so we use 2.5 miles in our estimates.

[19] The per-mile fee on taxis and FHVs computed in Section 3.2 takes into account "empty miles" driven by such vehicles when they are not carrying passengers. If delivery services accurately track every single mile driven by their vehicles, the per-mile fee

Two additional points deserve particular attention regarding congestion tolls on delivery trucks.

First, there is the issue of double- (and on occasion, triple-[20]) parking, which is of course hugely detrimental to the effective flow of traffic. While already prohibited, double-parking by delivery vehicles in New York City continues to be prevalent, with companies viewing the fines as a "cost of doing business."[21] So unlike most other vehicles, delivery trucks often create traffic congestion even when they are not actively moving, and per-mile (or per-trip, or per-day) tolls cannot address this issue. A logical way to address it is through increased levels and enforcement of parking fines. Note that the level (and the level of enforcement) of these fines may in principle vary by time and location to focus on the most congested areas during peak times: a delivery truck double-parked on a wide avenue between the hours of 2am and 3am does not create much if any traffic congestion, while the same truck double-parked on a narrow street between the hours of 6pm and 7pm is likely to create large negative externalities.

Second, unlike most other types of traffic, delivery vehicles can be much more flexible in terms of *when* deliveries happen: e.g., it is much easier for Amazon to ship packages between the hours of 9pm and 5am than it is for a typical office worker to shift his or her work hours to the 9pm–5am time window. This observation makes it easier and less disruptive to the lives of New Yorkers to impose high per-mile tolls and high and actively enforced parking fines on delivery vehicles during peak hours, while at the same time not imposing any congestion fees on deliveries that take place during off-peak hours. And as discussed in the Introduction, parts of the revenue from these tolls and fines can be used to help delivery drivers, fleet owners, and recipients of those deliveries to implement technologies that make these overnight deliveries more convenient and less costly.

In addition to trucks, some deliveries take place in personal cars via freelance delivery services like Instacart, Amazon Flex, Gopuff, etc.[22] Under the paused proposal, many of these deliveries would not face any tolls at all.[23] Charging these delivery trips the same per-mile toll during peak hours in the CBD as the per-mile taxi and FHV toll computed in Section 3.2 would keep the overall scheme equitable and effective, and similarly to the case of delivery trucks, would create an incentive for delivery services and their customers to shift deliveries to off-peak hours.

---

should be adjusted accordingly.

[20]`https://www.uppereastsite.com/fresh-direct-takes-over-upper-east-side-nyc-block-with-triple-parked-trucks-barrier-cones/`

[21]`https://www.freightwaves.com/news/todayspickup/ups-fedex-parking-fines`, `https://nyc.streetsblog.org/2019/05/02/the-dot-quandary-double-parking-isnt-illegal-except-when-it-actually-is`.

[22]Here we have in mind deliveries of relatively large orders that rely on cars. In the CBD, smaller deliveries, such as restaurant orders like those on DoorDash, Uber Eats, Grubhub, etc., are typically done on scooters or e-bikes, and are thus much less detrimental to the flow of traffic. Of course, if tolls on other forms of vehicles *are* imposed, and traffic in Manhattan starts flowing more freely, some of these restaurant delivery drivers may choose to switch to cars—in which case they would of course need to be tolled just like the other delivery drivers.

[23]FHV drivers may choose to deliver packages during the day, in addition to carrying passengers, and are only tolled for the latter. Likewise, if a CBD resident delivers an Instacart order within the CBD using his or her own car, there would be no toll under the paused plan.

5 Conclusion

Congestion pricing has the potential to improve traffic and generate revenue in New York City and many other congested cities worldwide. However, to be equitable and effective, it needs to follow the "similar tolls for similar amounts of congestion" principle. We discuss how to modify New York's paused congestion pricing plan to closer follow this principle, and show that such a change would make it possible to substantially lower the cordon fee paid by the drivers of personal cars while achieving the paused plan's revenue and traffic reduction goals, and generating additional revenue that can be used to alleviate the effects of congestion pricing on taxicab medallion owners and other business owners.

References

Almagro, M., F. Barbieri, J. C. Castillo, N. G. Hickok, and T. Salz (2024). Optimal urban transportation policy: Evidence from chicago. NBER Working Paper 32185, available at `https://www.nber.org/papers/w32185`.

Buchholz, N. (2021). Spatial equilibrium, search frictions, and dynamic efficiency in the taxi industry. *Review of Economic Studies 89*(2), 556–591.

Castillo, J. C. (2023). Who benefits from surge pricing? University of Pennsylvania Working Paper, available at `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3245533`.

Cohen, P., R. Hahn, J. Hall, S. Levitt, and R. Metcalfe (2016). Using big data to estimate consumer surplus: The case of Uber. NBER Working Paper 22627, available at `https://www.nber.org/papers/w22627`.

Cramer, J. and A. B. Krueger (2016). Disruptive change in the taxi business: The case of Uber. *American Economic Review 106*(5), 177–82.

Leccese, M. (2024). Asymmetric taxation, pass-through and market competition: Evidence from ride-sharing and taxis. University of Maryland Working Paper, available at `https://ssrn.com/abstract=3824453`.

Lehe, L. J. and S. Devunuri (2022). Large elasticity at introduction. *Research in Transportation Economics 95*, 101116.

NYCDOT (2019). New York City mobility report. Report of the New York City Department of Transportation, available at `https://www.nyc.gov/html/dot/downloads/pdf/mobility-report-singlepage-2019.pdf`.

NYMTC (2022). Hub bound travel data. Report of the New York Metropolitan Transportation Council, available at `https://www.nymtc.org/Portals/0/Pdf/Hub%20Bound/2022%20Hub%20Bound/May%202022/2022%20Hub%20Bound%20Report-%205.17.24-FINAL%20corrected.pdf`.

Pigou, A. C. (1920). *The Economics of Welfare*. Macmillan.

Traffic Mobility Review Board (2023). Congestion pricing in New York: A toll structure recommendation from the Traffic Mobility Review Board. Available at `https://new.mta.info/document/127761`.