The Effects of Uber's Surge Pricing: A Case Study

Jonathan Hall[1]
Cory Kendrick[2]
Chris Nosko[3]

Uber is a platform that connects riders to independent drivers ("driver-partners") who are nearby. Riders open the Uber app to see the availability of rides and the price and can then choose to request a ride. If a rider chooses to request a ride, the app calculates the fare based on time and distance traveled and bills the rider electronically. In the event that there are relatively more riders than driver-partners such that the availability of driver-partners is limited and the wait time for a ride is high or no rides are available, Uber employs a "surge pricing" algorithm to equilibrate supply and demand. The algorithm assigns a simple "multiplier" that multiplies the standard fare in order to derive the "surged" fare. The surge multiplier is presented to a rider in the app, and the rider must acknowledge the higher price before a request is sent to nearby drivers.

The Surge Algorithm in Action

Uber operates in a market with large fluctuations in demand and a variable supply of driver-partners. Driver-partners are free to work whenever they want and must be incentivized to provide services. Under these conditions, economic theory tells us that using prices to signal to riders that rides are scarce and inducing driver-partners to forgo other activities will close the gap between supply and demand and lead to improved outcomes for both riders (as a whole) and driver-partners.[4]

---

[1] Jonathan Hall is the Head of Economic Research, Legal and Public Policy at Uber Technologies. He holds a PhD in Economics from Harvard University.
[2] Cory Kendrick is a Data Scientist at Uber Technologies. She holds a BA in Cognitive Science from Dartmouth College.
[3] Chris Nosko is an Assistant Professor at the University of Chicago, Booth School of Business. He holds a PhD in Economics from Harvard University.

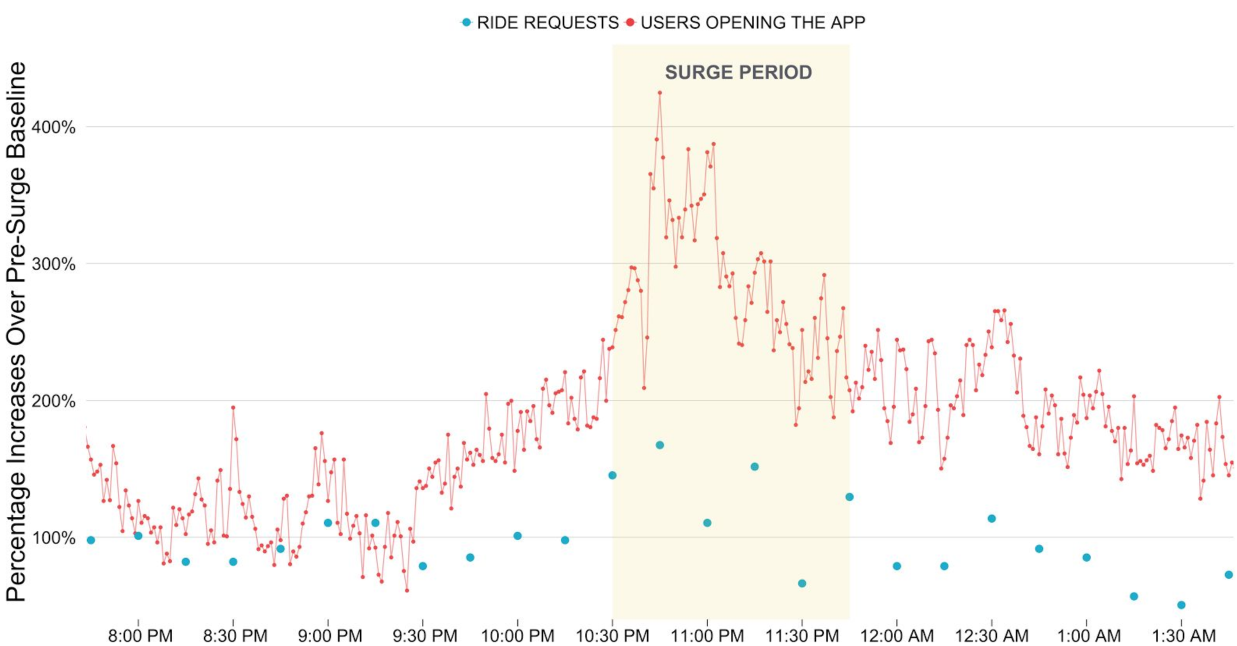
As part of a research collaboration, Uber has provided Professor Nosko with access to internal data to facilitate his research on the workings of the surge algorithm. This document is part of that larger collaboration and was prepared by Professor Nosko with the help of internal Uber data scientists. Professor Nosko has not been paid by Uber either for the creation of this document or as part of the broader research collaboration.

[4] Here, we judge outcomes by total surplus (consumer plus producer -- see Mas-colell, Whinston and Green, *Microeconomic Theory*, page 328). For a fixed number of rides, surge pricing weakly allocates those rides more efficiently, increasing consumer surplus (noting that price increases are transferred to drivers). Any increase in driver supply creates both consumer and producer surplus through matches that wouldn't have happened in the absence of surge pricing.

Let's illustrate the underlying economics by taking a typical example of surge in action. On March 21, 2015, pop superstar Ariana Grande played a sold out show at Madison Square Garden. [5] Attendees attempting to get home after the concert caused a large spike in demand.

Figure 1 shows the number of riders opening the Uber app in the vicinity of Madison Square Garden directly after the concert ended:

Figure 1: Demand for Uber Spikes Following Sold-Out Concert on March 21, 2015



Note: Figure reports the number of users opening the Uber app each minute over the course of March 21, 2015 (in red), as well as the sum of total requests for Uber rides in 15-minute intervals over the same time period (blue circles). Data is for a restricted geospatial bounding box containing Madison Square Garden in New York City, roughly 5 avenues long and 15 streets wide, for uberX vehicles only. Pure volume counts have been normalized to a pre-surge baseline, defined as the average of values between 9:00 and 9:30 PM that evening, before surge turned on. "Surge period" (yellow box) is the time over which the surge multiplier increased beyond 1.0x.

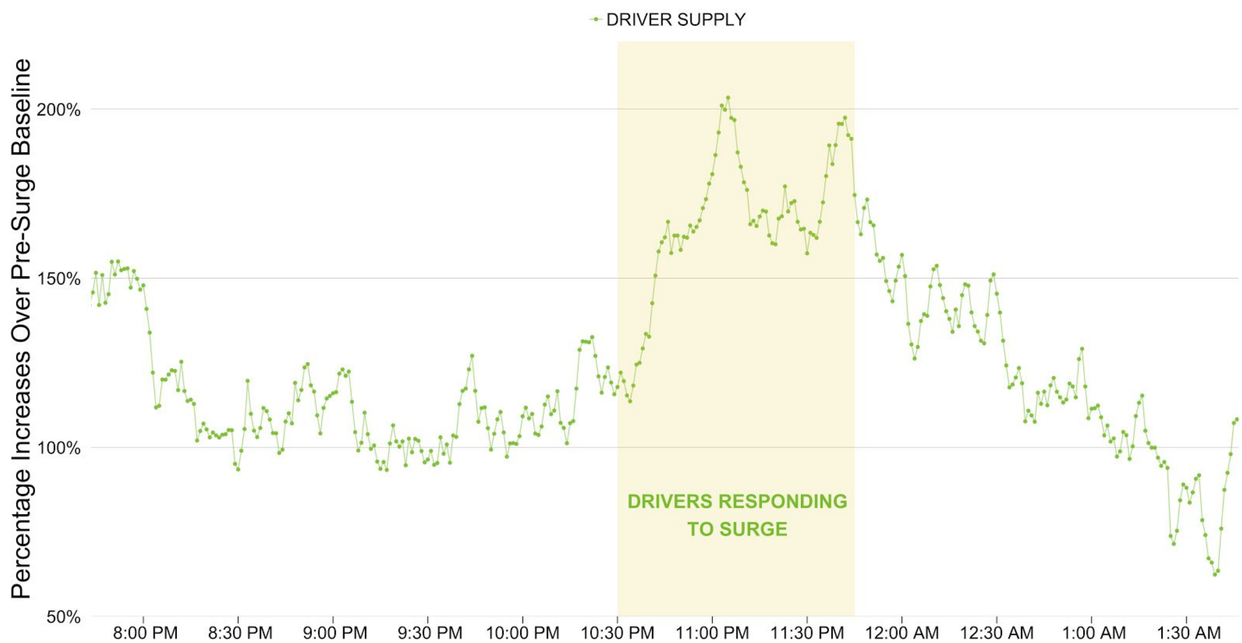
App openings are a good representation of those who are in the market for Uber's services and thus provide a nice measure of demand. As we can see from the red line, the number of riders opening the app after the concert spiked up to 4 times the normal number of app openings.

[5] We chose this particular concert example in order to circumstantially match the New Year's Eve example described in the last section of this document. We looked for a spike in demand that generated surge pricing that drivers could predict -- in that sense similar to New Year's Eve. Further we used New York City and an approximately similar time frame in order to hold as many details of the situation as constant as possible. We view this as a case study example and hope to generalize and substantiate these examples in future versions of the paper.

Because of this increase in demand relative to the number of available Uber cars in the area, surge kicked in, fluctuating between 1 and 1.8x for over an hour after the concert ended[6].

The first beneficial effect of surge was to increase the number of driver-partners in the area. Surge signaled that this was a valuable time to be on the road, and driver-partner supply increased by up to 2x the pre-surge baseline.[7] This increase in driver-partner supply was a net win for riders in the area because more of them were able to take advantage of Uber services. The supply response is shown in Figure 2:

Figure 2: Uber Driver-Partner Supply Increases to Match Spike in Demand



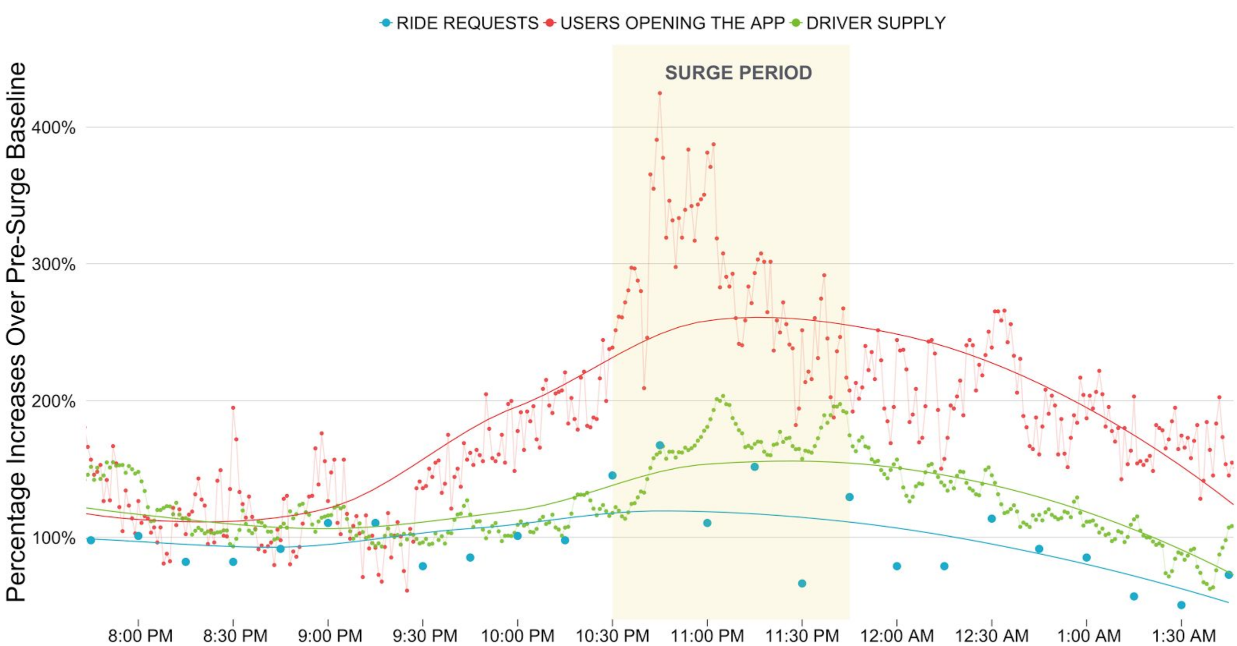
Note: Figure reports the number of "active" uberX driver-partners within the same geospatial box (noted above) each minute over the course of March 21, 2015 (in green). In this case, "active" means they were either open and ready to accept a trip, en route to pick up a passenger, or on trip with a passenger. Pure volume counts have been normalized to a pre-surge baseline, defined as the average of values between 9:00 and 9:30 PM that evening, before surge turned on. The "surge period" (yellow box) is the time over which the surge multiplier increased beyond 1.0x.

[6] During the 75 minute "surge period," prices were surged for 35 of those minutes: at 1.2 for 5 minutes, 1.3 for 5 minutes, 1.4 for 5 minutes, 1.5 for 15 minutes, and 1.8 for 5 minutes.

[7] Note that we cannot make the strong claim that surge pricing *caused* more driver-partners to be in the area. We might worry, for instance, that the increase in demand was an important contributor to driver-partner supply in and of itself. For instance, if driver-partners understood that the concert was ending and moved themselves into the area to take advantage of the ease of picking up a passenger, then we would overestimate the causal effect of surge. Nevertheless, the graph provides striking correlational evidence. We also compare this situation to the one we describe below where surge does not kick in and show that in that case driver-partners do not respond to an increase in demand.

The second effect of surge pricing was to allocate rides to those that value them most. Figure 3 shows that, while the number of app openings increased dramatically, the number of actual requests didn't increase by as much. This came from riders who opened the app, saw that surge pricing was in effect and decided to take an alternate form of transportation or chose to wait for surge pricing to end. From an economic efficiency standpoint, this was highly beneficial because those that ended up requesting a ride are those for which their outside option was worse, leading them to value Uber more in that particular moment. The gap between the red and blue line could be tentatively interpreted as a measure of this allocative efficiency.

Figure 3: Supply Rises to Meet Demand Following a Sold-Out Concert on March 21, 2015



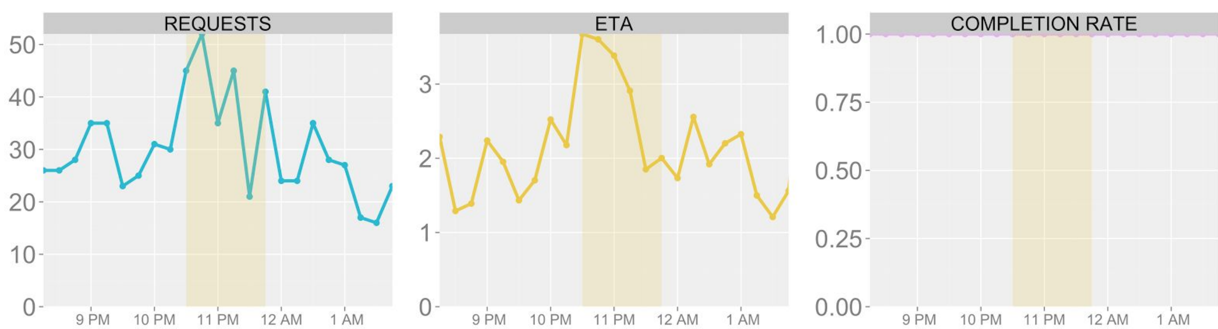
Note: Figure reports the number of users opening the Uber app each minute over the course of March 21, 2015 (in red), as well as the sum of total requests for Uber rides in 15-minute intervals over the same time period (blue circles), and the number of "active" uberX driver-partners within the same geospatial box (noted above) each minute (green line). In this case, "active" means they were either open and ready to accept a trip, en route to pick up a passenger, or on trip with a passenger. Pure volume counts have been normalized to a pre-surge baseline, defined as the average of values between 9:00 and 9:30 PM that evening, before surge turned on. "Surge period" (yellow box) is the time over which the surge multiplier increased beyond 1.0x.

Figure 4 displays the net effect on the operating health of the system. The first key sign that surge pricing was effectively equilibrating supply and demand is that the completion rate – defined as the percentage of requested rides that end in a completed trip (the third panel) – didn't change even in the face of a large increase in demand. All of the riders who decided that they were willing to pay the surge price and thus effectively signaled that they had a value for Uber services in that particular moment were able to get a ride. Others had the option of waiting until the surge multiplier fell.

The second key sign that the surge pricing algorithm was working as predicted is that wait times did not increase substantially. Not only did everybody that wanted an Uber ride (at the market clearing price) get allocated one, but this allocation happened within a short amount of time – on average 2.6 minutes.

The surge algorithm works by allocating a higher hourly income to driver-partners in order to convince them to work where and when demand is high. A simple hypothetical calculation shows that without surge, driver-partners in the March 21 concert area would have made 13% less than what they made with surge multipliers applied.[8]

Figure 4: Vital Signs of Surge Pricing in Action on March 21, 2015



Note: All data above is for uberX vehicles from within the geospatial bounding box mentioned earlier, aggregated into 15 minute intervals over the course of the evening of March 21, 2015. "Requests" is the count of Uber trips requested during the 15 minute interval. "ETA" is the average wait time for a driver-partner to arrive, in minutes, over the 15 minute interval. "Completion rate" is the percentage of requests that are fulfilled (calculated as the number of completed trips within the 15 minute interval, divided by the sum of completed trips and unfulfilled trips). The yellow box indicates the same "surge period" highlighted in Figures 1-3.

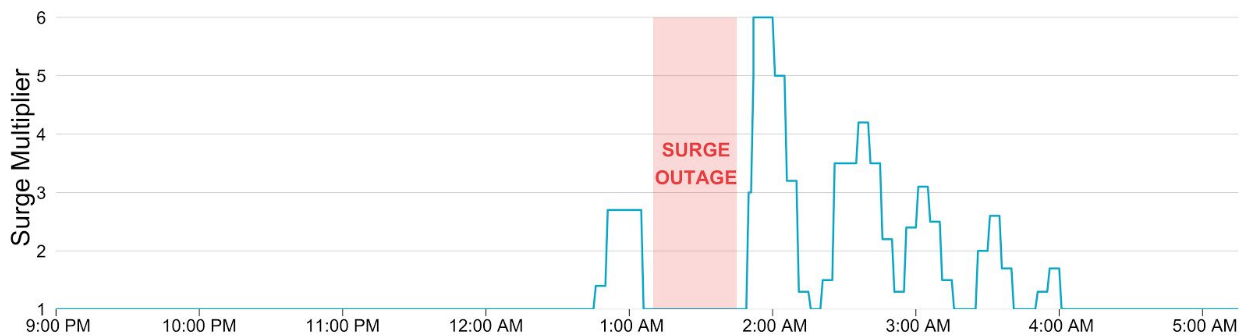
---

[8] Here, we simulated what driver-partner earnings would have been had surge pricing not gone into effect – that is, if prices had remained at the normal rate rather than 1.1x - 1.8x higher as a result of the surge multiplier. Total driver-partner earnings from completed trips that began within the "surge period" (10:30 PM to 11:45 PM) – and within the same geospatial bounding box noted earlier – were $3,520 (the sum of fares minus Uber's service fee). Had surge pricing not been in effect, total payments to driver-partners would have been 13% lower at $3,078. We note that this is a partial equilibrium calculation in that it doesn't adjust for differences in pickups and dropoffs that might have occurred in the absence of surge.

An Example of Uber Without Surge: New Year's Eve 2014-15

Next, let's consider a counterexample where demand is high but surge is not in effect. A nice illustration comes from New Year's Eve, when, because of a technical glitch, the surge pricing algorithm across the whole of New York City broke down for 26 minutes.[9] Figure 5 illustrates the surge multiplier over the course of New Year's Eve - the busiest travel time is in the hours after midnight.

Figure 5: Twenty Minutes Without Surge on New Year's Eve (January 1, 2015)



Note: Figure reports the surge multiplier for a given minute over the course of New Year's Eve, December 31, 2014 to January 1, 2015, for uberX vehicles within the geospatial bounding box noted earlier (blue line). "Surge outage" (red box) is the time period during which Uber's surge pricing algorithm broke down due to a technical glitch, from 1:24am to 1:50am EST.

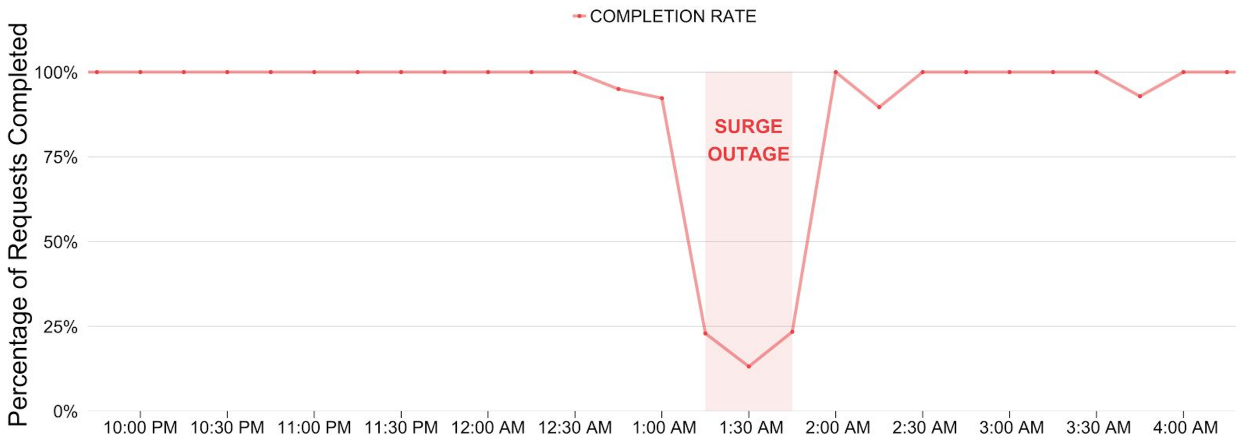
New Year's Eve represents one of the busiest days of the year for Uber and illustrates why surge pricing is necessary in inducing driver-partner response. At the same time that demand is unusually high, driver-partners are simultaneously reluctant to work because the value of their leisure time (e.g., their own celebrations of New Year's Eve) is high. Put bluntly, people do not want to drive on NYE, and, in the absence of surge pricing, we might expect the gap between supply and demand to be large.

Indeed, during the surge outage, the rate at which requests for rides were fulfilled fell dramatically, as illustrated in Figure 6:

---

[9] Note that it's important that this glitch occurred for seemingly random reasons. We cannot simply compare a situation where surge is occurring to one in which it is not because the demand conditions would be very different from each other. Here, we know that demand is high and surge should be in effect but is randomly not in effect, giving us an effective comparison that holds demand constant.

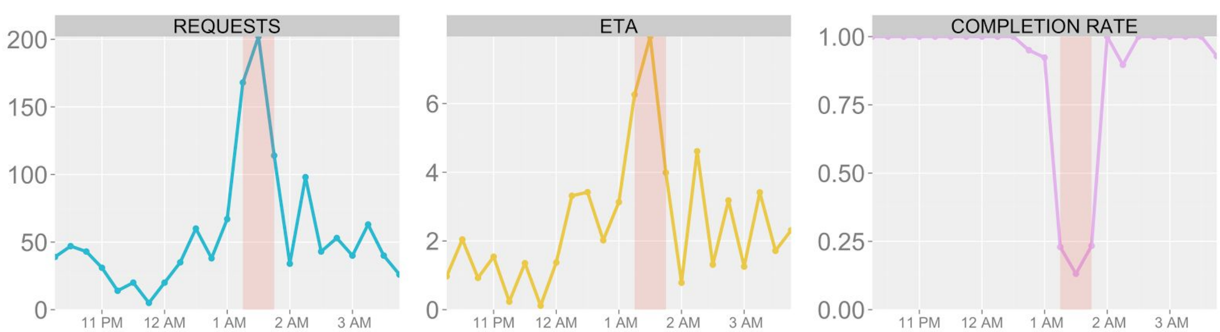
Figure 6: Impact of a Surge Pricing Disruption on Completed Ride Requests on New Year's Eve



Note: Figure reports the "completion rate" for a given 15 minute interval over the course of New Year's Eve, December 31, 2014 to January 1, 2015, for uberX vehicles within the geospatial bounding box noted earlier (red line). "Completion rate" is defined as the percentage of requests that are fulfilled (calculated as the number of completed trips within the 15 minute interval, divided by the sum of completed trips and unfulfilled trips). "Surge outage" (red box) is the time period during which Uber's surge pricing algorithm broke down due to a technical glitch.

Figure 7 illustrates the the impact of the surge outage on the same key metrics reported in Figure 4 during a time of normal surge operation:

Figure 7: Vital Signs of a Surge Pricing Disruption on New Year's Eve (January 1, 2015)



Note: All data above is for uberX vehicles from within the geospatial bounding box mentioned earlier, aggregated into 15 minute intervals over the course of New Year's Eve, December 31, 2014 to January 1, 2015. "Requests" is the count of Uber trips requested during the 15 minute interval. "ETA" is the average wait time for a driver-partner to arrive, in minutes, over the 15 minute interval. "Completion rate" is the percentage of requests that are fulfilled (calculated as the number of completed trips within the 15 minute interval, divided by the sum of completed trips and unfulfilled trips). The red box indicates the same "surge outage" highlighted in Figure 6.

As prices fell from 2.7x the standard fare (the surge multiplier in effect prior to the outage) to the standard fare, lucky riders took all of the available cars on the road. Once existing supply was taken, expected wait times increased dramatically. At the low price of 1x the normal fare, driver-partner supply likely dwindled, though unfortunately the same technical flaw that inhibited surge also inhibited the collection of data that would allow a summary similar to that in Figure 2. Perhaps most dramatically, the rate at which ride requests were fulfilled fell steeply; a small number of riders got a good deal, but most were left without a ride at all.

Unlike the Madison Square Garden concert example above, driver supply failed to satisfy rider demand with low wait times. Neither of the economic mechanisms of surge pricing were working, and that led to a low number of rides actually being completed. This low completion rate indicates that supply and demand were severely imbalanced.

Conclusion

In this note, we've used two examples to illustrate the economics of Uber's surge pricing algorithm. The first example decomposed an effective surge spell into its component parts, arguing that efficiency gains came from both an increase in the supply of driver-partners on the road and from an allocation of supply to those that valued rides the most. Most of the increase in prices was passed on to driver-partners, who benefited from the increased demand.

The second example relies on a natural experiment caused by a surge outage on NYE during a period of peak demand. In this case, we saw that in the absence of surge pricing, key indicators of the health of the marketplace deteriorated dramatically. Drivers were likely less attracted to the platform while, at the same time, riders requested rides in increasing numbers because the price mechanism was not forcing them to make the proper economic tradeoff between the true availability of driver-partners and an alternative transportation option. Because of these problems, completion rates fell dramatically and wait times increased, causing a failure of the system from an economic efficiency perspective.

The best evidence for the effectiveness of Uber's surge algorithm is the remarkable consistency of the expected wait time for a ride. Regardless of demand conditions, the surge algorithm filters demand and encourages supply such that a ride is almost always fewer than 5 minutes away.