Is surge pricing just supply and demand? It depends on who sets the price

A note by Claude, an AI model made by Anthropic, written at a reader's request. November 2026.

The short version

When a crowd of people suddenly needs rides, letting the price rise beats freezing it. A higher fare pulls more drivers onto the road, and it decides who gets the scarce cars by willingness to pay instead of by luck. That is the standard defense of surge pricing, and as far as it goes it is correct.

But the defense takes one thing for granted: that the price stops rising at the right point, the level at which every car is full. In an ordinary market, competition sees to that. A seller with an empty car cuts his price, so the price cannot stay above that level. Drivers on Uber and similar apps cannot cut their price. The company sets the fare for all of them and takes a cut of every fare. If stranded people keep paying as the fare climbs, a fare well above that level can bring the company more money: fewer people ride, but each pays so much more.

There is a simple fix, at least on paper: cap the company's cut, not the fare. In an emergency the fare may rise as far as it needs to, but the company keeps only its normal dollar amount per ride, and the rest goes to the driver. The company's cut then no longer grows with the fare. The company earns the same on every ride, so it can only earn more by completing more rides. And the most rides happen at the level that fills every car: the fare at which as many people want a car as there are drivers willing to come out. A higher fare loses riders. A lower one loses drivers.

Does the real Uber overcharge in emergencies? I found nothing that shows Uber doing so. Uber's own pages describe its cut as a percentage of the fare, or as the difference between what the rider pays and what the driver gets, and either one creates this incentive. An economist who worked on a study of Uber's 2015 data summed up its result: when Uber's price doubles, demand falls by only about 40%. Sixty percent of the riders paying twice the price is more money in total, so at those prices a higher fare would have paid. The same economist believed that Uber's prices at the time did balance riders and drivers, which would mean that Uber could have charged more and did not. Nothing I read measures an emergency.

This note does not settle whether it is fair that the cars go to whoever can pay. The claim here is narrower: even judged only by how many people get home, the company's price can be the wrong one. So "surge pricing is good" and "the company may be overcharging" can both be true. A rising price is the right response to a shortage. Whether the particular price is right depends on who sets it and what they earn from it.

Elad's note: Claude also defines below what "clearing the market" means, it explains why it's a pretty nuanced term, and explains why in the most sensible definition of "clearing the market", Uber very likely doesn't actually do that under surge pricing. (i.e. Uber is strongly incentivized not to just "clear the market", but rather to serve way fewer riders than that)

1. Why a rising price helps

Take one evening in one city. Flights are cancelled and far more people than usual want a ride.

I will use one made-up example all the way through. Its numbers are deliberately extreme, so that the effects are easy to see. On a normal day a ride costs $20. The driver gets $15 and the app keeps $5. On this evening, at $20, 95 people want a ride and only 30 drivers are out. Sixty-five people are stuck, and who gets a car is a matter of luck.

Now let the fare rise, and suppose for the moment that the app goes on keeping $5 a ride. Fewer people want a car at a higher fare, because some decide to wait or walk. More drivers come out, because driving pays better. At $49 the number of people who want a car equals the number of drivers willing to take them, and there are 88 rides instead of 30. I will call $49 the balancing fare: the level that fills every car. Economists say that at this fare "the market clears".

This is the argument people make when they defend surge pricing, and it is right. A frozen fare strands people.

2. The step the argument skips

Who picks the fare?

In a textbook market, nobody does. If the price is too high, some seller is left with an empty car, and he offers a lower price to fill it. If the price is too low, buyers outbid each other. The price settles where the two sides balance, and it gets there because sellers compete.

A ride-hailing app works differently. Drivers do not set prices and cannot undercut each other. One company sets the fare for every ride and takes a cut of each one. The company is not trying to balance riders and drivers. The company is trying to make money. Sometimes balancing the two sides and making money are the same thing. Whether they are the same depends on how the company is paid.

3. A company that takes a percentage of each fare

Suppose the app keeps 25% of every fare. On a normal day 25% of a $20 fare is the same $5. In a surge 25% is far more than $5. The app's income is a quarter of everything riders spend, so the app wants riders' total spending to be as large as possible.

Start where riders and drivers balance. With a 25% cut the two sides balance at $57, a little above $49, because less of each fare reaches drivers and it takes a higher fare to bring out enough of them. At $57 there are 86 rides, and riders spend about $4,900 in total.

Now raise the fare to $200, which in this example is the fare that earns the app the most. Only 50 people still want a ride, but each pays three and a half times as much. Total spending is $10,000, and the app's income has doubled.

Look at what a $200 fare does. A driver earns $150 a ride, so 300 drivers want to work, and there are only 50 rides to go around. Meanwhile 36 people who would have paid $57 go without. Idle cars sit next to stranded people, and everyone involved did what was best for themselves.

A company on a percentage pushes the fare above the balancing level when riders react weakly to price. The technical word is inelastic: a fare 1% higher loses less than 1% of riders, so their total spending goes up. People stranded at an airport are a natural case. The opposite case is riders who drop out quickly as the fare climbs, for example because another app or a taxi is easy to get. Then a higher fare lowers total spending, and the company's best fare is the one that balances the two sides.

4. A company that sets the fare and the driver's pay separately

Now suppose the app chooses what the rider pays and what the driver gets, each on its own, and keeps the difference.

Unlike the company of section 3, this company has no reason to leave cars idle. If too many drivers show up, the company pays them less, until only as many come out as there are riders. Riders and drivers always balance. An economist would say that this market clears too.

What the company chooses is how much to keep in the middle. In the example its best choice is a fare of $222 and driver pay of $22, just enough to bring out the 44 drivers it needs. The company keeps $200 a ride. No rider who accepts that fare has to wait, and no car is idle. But there are only 44 rides, half as many as at the $49 balancing fare.

The reason is simple. A ride happens only if the rider would pay more than the driver needs plus the company's cut. The bigger the cut, the more rides never happen even though rider and driver would both have been glad of them. And the cut that pays best tends to be bigger the less riders react to price. If stranded riders react less than ordinary ones, the company's cut is largest in an emergency.

5. What the right price should mean

Section 4 shows that "riders and drivers balance" is too weak a test. A company can pass that test while keeping nine-tenths of the fare.

A better benchmark is the price that competition would produce: drivers free to undercut one another, and the company taking its ordinary fee for making the match. In the example that is the $49 balancing fare of section 1. The $49 fare meets two conditions:

- Riders and drivers balance. Everyone willing to pay the fare gets a car, and everyone willing to drive for the pay gets a rider.
- The company keeps its normal fee per ride, counted in dollars.

Why dollars and not a percentage? The company does the same job, matching one rider with one driver, whether the fare is $20 or $200. A percentage fee multiplies the company's take by ten when the fare does.

This benchmark is about how many people get a ride, not about fairness. If your objection to surge pricing is that the cars go to whoever can pay, the benchmark does not answer it.

6. A fix: cap the company's cut, not the fare

Here is the rule. In an emergency, the fare may rise as far as it needs to. The company may keep only its normal dollar fee per ride. Everything above that goes to the driver.

Under this rule the company in the example earns $5 on every ride, whatever the fare. The company's income is $5 times the number of rides, so the company wants as many rides as possible. And the number of rides is highest exactly at the balancing fare. If drivers are the scarcer side, a higher fare and higher pay bring more of them. If riders are the scarcer side, a lower fare brings more of them. The company's best choice is a fare of $49, with $44 going to the driver and 88 rides.

Three things are worth noticing.

- This is not a cap on the fare. The fare more than doubled. A cap on the fare can also help, but only if it is set at about the right level, and someone has to guess that level. Set too low, a fare cap brings back the shortage of section 1. A cap on the cut needs no guess: the fare stays free, and the company finds the balancing level itself.
- The whole increase goes to drivers, who are the people a surge is supposed to attract.
- In this example even the company's own surge gets more people home than the frozen fare: 50 rides on a percentage cut and 44 with fare and pay set separately, against 30. Whether a company's surge beats a frozen fare in general depends on extra drivers being able to come out.

Why not tax the company's extra cut instead? A tax of 100% works exactly like the cap. The company gains nothing from taking more than its normal fee, so the company stops taking more, and the tax then has nothing to collect. Any lower tax leaves the company part of every extra dollar, so the company still wants a large cut. In the example, a 70% tax on the company

of section 4 moves its fare only from $222 to $217, and the tax would have to be above 98% before that company gave up the extra.

The rule also has drawbacks. Under the cap the company gains only a few dollars a ride from finding the balancing fare, while the blame for a high fare still lands on the company. A company that cares about its reputation might therefore keep fares too low. Someone has to define the "normal" fee, and the company could raise it in advance. And Uber says that trips with higher fees "offset the lower prices" on other trips (Uber, service fee). On Uber's account, then, a cap would remove money that now pays for cheaper rides.

7. The example in one table

Who sets the price, and how	**Fare**	**Driver gets**	**Company keeps**	**Rides**
Nobody: fare frozen at the normal level	$20	$15	$5	30 (95 people want one)
Company on a 25% cut, at the fare that balances the two sides	$57	$43	$14	86
Company on a 25% cut, at its most profitable fare	$200	$150	$50	50 (300 drivers willing)
Company sets fare and pay separately	$222	$22	$200	44
Company's cut capped at $5 a ride	$49	$44	$5	88
70% tax on the cut above $5 (company of section 4)	$217	$23	$62, after $132 in tax	46

The numbers come from invented curves, given in the appendix. In this example riders barely react to price: starting from the balancing fare, a fare twice as high loses only about 15% of

them. In the one real measurement, described in the next section, a price twice as high lost about 40%. So the effects in the example are larger than real ones would be. The example is there to show the direction of the effects.

8. What is known about the real Uber

A caution first: none of the sources below measures an emergency, and I did not read every study in full.

How Uber is paid. Uber's own pages describe both arrangements. Uber's US page says its service fee "is the difference between what riders pay and what drivers earn" and "varies from trip to trip", averaging "around 20%" (Uber, service fee). That is the arrangement of section 4. Uber's page for drivers says "Uber's service fee percentage does not change during surge pricing" (Uber, how surge works). That is the arrangement of section 3. Both pages say that the details vary by city or market.

How riders respond to price. Steven Levitt, the economist known for Freakonomics, worked on a study of 2015 Uber data from four US cities (paper) and summed it up: "when the price of Uber say, doubles … the demand for Uber only falls by about 40 percent" (interview transcript). That is the weak, inelastic response described in section 3. Levitt added: "I think that the belief is that they actually do get to equilibrium", meaning the balancing level. If Levitt's belief was right, Uber had room to charge more in 2015 and did not use it. At the time of the interview Uber was losing money: the program says Uber had just lost more than $1.2 billion in half a year, mostly on subsidies to drivers.

What has happened since. An Oxford study of 1.5 million trips by 258 UK drivers found that after Uber changed its pricing in 2023, "the higher the fare charged to the customer, the higher Uber's take rate" (the share of the fare Uber keeps). For the typical driver that share rose from 25% to 29%, although the average stayed about the same (paper). A Columbia Business School analysis of one US driver's trips put Uber's share at "about 32%", rising to "upwards of 42%" by the end of 2024 (the Guardian's report). In the same report Uber said it does "not recognise the figures" in the Oxford study, and called suggestions that its systems "manipulate pricing unfairly or discriminate" "simply false". These findings fit a company that uses its control of the price more than it did in 2015. The findings also fit a company that simply stopped subsidizing rides. They do not settle which.

What I could not find out.

- Whether Uber's surge fares today are above the balancing level.
- Whether riders in an emergency respond to price the way riders did in 2015.
- How much other apps and taxis hold Uber's fares down when every app is short of cars.
- Whether Uber still limits surge pricing in emergencies, as it agreed to do in 2014.

9. What this simple picture leaves out

- **Competitors.** Riders can open another app, and drivers can work for more than one. Both push prices toward the competitive level.
- **The long run.** A company that worries about its reputation or about regulators may hold back. Thirty-nine US states have price-gouging rules for emergencies (NCSL), and in 2014 Uber agreed with New York's attorney general to limit surge pricing in emergencies (New York Times).
- **Waiting time.** Quick pickups need some free cars nearby, so not every idle car is wasted.
- **Personal prices.** A company that can charge each rider close to the most that rider would pay can leave fewer people without a ride while taking more from each.

10. Bottom line

The standard defense of surge pricing shows that a rising price beats a frozen one. The defense does not show that a price chosen by one company is the price that competing drivers would have arrived at. On a ride-hailing app one company sets the price for thousands of drivers, and the company's income depends on that price. So the honest answer to "is surge pricing good?" is: a rising price is good, and the particular price is only as good as the incentives of whoever sets it. Capping the company's cut in an emergency, and not the fare, is one way to point those incentives at the fare that gets the most people home.

Appendix: the model and the proofs

For readers who want the argument stated exactly.

Setting. One company, one evening.

- Each rider wants one ride and has a value for it. Each driver can give one ride and has a cost of giving it. D(p) is the number of riders whose value is at least p. S(w) is the number of drivers whose cost is at most w. D falls as p rises, S rises with w, and both change smoothly.
- The company posts one fare p and one pay w. Riders and drivers accept or decline. Drivers decide on the pay alone, whatever their chance of getting a rider. The number of rides is min{D(p), S(w)}.
- The company keeps the fee p − w on each ride, has no other cost per ride, and chooses p and w to maximize its profit for the evening. A most profitable choice is assumed to exist.
- m is the company's normal fee in dollars.

Definitions. Riders and drivers balance (the market clears) if D(p) = S(w). The balancing fare of the main text is the p with D(p) = S(p − m): balance with the company taking its normal fee. A ride's worth is the rider's value minus the driver's cost.

A fact used throughout. If $D(p) = S(w)$, a rider without a car has a value below p, and a driver without a rider has a cost above w. So every ride that does not happen is worth less than the fee.

Result 1: percentage fee. The company keeps a share s of the fare. Its profit is $s \cdot p \cdot \min\{D(p), S((1 - s)p)\}$. Let p^* be the fare with $D(p^*) = S((1 - s)p^*)$. Below p^*, rides are limited by drivers, and profit rises with p. Above p^*, rides are limited by riders, and profit is $s \cdot p \cdot D(p)$. If demand is inelastic at p^*, profit still rises with p there, so the most profitable fare is above p^*, where $S((1 - s)p) > D(p)$. If demand is elastic at p^* and at every higher fare, the most profitable fare is p^*.

Result 2: fare and pay set separately. If $D(p) > S(w)$, raising p slightly loses no ride and raises the fee. If $D(p) < S(w)$, lowering w slightly does the same. So the most profitable choice has $D(p) = S(w)$. At that choice the fee equals rides $\times$ $(1/|D'(p)| + 1/S'(w))$.

Result 3: fee capped at m. If $D(p) > S(w)$, raising p and w together keeps the fee and adds rides. If $D(p) < S(w)$, lowering both does the same. So the most profitable choice has $D(p) = S(w)$ with a fee of at most m. Suppose the cap binds: the company's profit rises with the fee all the way up to m, as it does when m is below the fee of Result 2 and profit has a single peak. Then the company takes exactly m, its profit is $m \cdot \min\{D(p), S(p - m)\}$, and that is largest exactly at the balancing fare. The result needs both sides to respond to price. If drivers do not respond at all, the capped company earns the same at any fare that fills every car. If riders do not, the capped company earns the same at any fare at or above the balancing one.

A cap on the fare. In the example a ceiling on the fare adds rides without causing a shortage if it is set between \$57 and \$200 for the company of Result 1, or between \$80 and \$222 for the company of Result 2. Below those ranges riders outnumber drivers again.

Tax. Tax the fee above m at rate t, for the company of Result 2. At $t = 100\%$ the company behaves as under the cap, and the tax collects nothing. Below 100% the company's chosen fee lies between m and the fee of Result 2. In the example any rate above about 98.7% gives the cap outcome. No rate below 100% does so for every pair of curves.

The example. $D(p) = 100 \cdot (1 - p/400)$ and $S(w) = 2 \cdot w$. Starting from the balancing fare (\$49, or \$57 on a 25% cut), a fare twice as high loses 14% to 17% of riders. The comparison with the frozen fare in section 6 is a feature of this example. With a fixed fleet of 80 cars, the frozen fare fills all 80, and the 25% company's most profitable fare still gives 50 rides.