Economics 301

Autumn 2015

A Guide to Preparing Price Theory Homework, with an application to the Economics of UBER

Table of Contents

I. The Questions

Below is an example problem set question. Parts (b) and (c) are conceptually difficult, even by Econ 301 standards, in order to illustrate how students should expect the quality of their answers to improve from one problem set to the next as they acquire additional price theory skills.[1]

UBER is a business that matches automobile drivers with paying passengers using an electronic mapping system. When a passenger enters UBER's market using his smartphone, he sees the price p and his approximate waiting time t for a vehicle/driver. For simplicity, assume that the price is quoted per minute of driving time. The revenue paid by each passenger is delivered to the driver, minus a 20 percent UBER handling fee.

a. Does it matter whether the UBER handling fee is paid by passengers or drivers?

b. If UBER ran an experiment of raising the posted price for a period of time, what would happen to passenger wait times and the number of rides? Would it matter whether UBER had announced (to potential drivers and passengers) this experiment in advance?

c. What can you say about UBER's optimal price p and percentage handling fee?

d. Could regulators use a price ceiling to encourage more efficient outcomes?

Before turning the page, take a few minutes to consider how you would answer these questions.

[1] This question was actually assigned, in a longer format, in a prior year toward the end of the course after students had seen much of problem set and lecture material.

II. A Textbook Answer

(The textbook audience is students, who are not assumed to already know price theory. A model student answer would be addressed to professor/TAs, and therefore shorter as shown below.)

<u>Setup the problem so that is no longer "open ended."</u>

Let t denote the time that passengers wait for pickup and $1-\theta$ denote the fraction of time that drivers are without a fare. As we increase the ratio of drivers to passengers, θ and t fall. Ie., drivers wait longer for a fare and passengers wait less.

θ and t are intrinsically connected: passengers wait less because there are more drivers waiting for each of them, and vice versa. Define a function $\theta(t)$ to represent this relationship, with $\theta'(t) > 0$, and hold fixed the *function* for many of the exercises below.

Among other things, the total scale of the market (e.g., the total number of drivers or riders, and not just their ratio) could shift the $\theta(t)$ function and thereby embellish the answers below that assume that waiting times are scale free.

Supply of drivers: Driving is a low-skill activity and lots of people are capable of doing it, so we model the supply of drivers as perfectly elastic with respect to minutely earnings, which is θ*0.8*p, where p is the price per minute paid by passengers and θ is the fraction of time that drivers have a passenger in the car. Let r denote driver's minutely reservation wage, which is constant by the infinite-elastic assumption. The driver supply condition is therefore

$$0.8\theta(t)p = r \tag{1}$$

Furthermore, note that p is set by UBER, and that as long as p is constant, the margins of adjustment that "clear the market" are the wait times.

Demand by riders: The full per-minute price u of being driven is $u = p + wt/m$, where m is the length of a trip in minutes, t is minutes waiting for pickup, and w is the waiting cost per minute. A passenger participates in the market if the full price is less than the price of the next best alternative (walking, using public transportation, driving own car, not traveling at all, etc.). Use the law of demand to conclude that the quantity of rides demanded $Q(u)$ falls with u. We assume that minutes per trip m is a constant.

This is an example of another general principle in economics: that the true price of consuming a good or service (by which we mean the amount of resources foregone at the margin) is often different than the cash outlay that a layman would refer to as "price." The "full price" terminology highlights the distinction, especially when customer time is part of the production process.[2]

The rider formulation above includes the wait time, but not the transit time itself. One reason to exclude the rider's transit time is that a demand curve refers to *alternatives*, and the rider's alternative transport modes takes time too. However, the more terse formulation above would not be appropriate for examining changes in the amount of transit time for UBER relative to alternatives. If the alternatives take a different amount of transit time than UBER does, then the more terse formulation would not be appropriate for examining changes in the wage rate w (even if the transit times were held constant).

Interpretation as a classic production problem. As an equivalent alternative to viewing the "product" as a trip that requires customer time, the product can be understood as an augmented product whose full price per minute traveled is p (i.e., no "customer time"), but is produced with driver time $1/\theta(t)$ and "waiter" time t/m. In other words, the rider is a single person serving two economic functions: as a consumer of trips and as one of the "employees" in producing trips. Production costs per minute traveled are wage-weighted sum of the time of the two employees:

$$c(t) = \frac{r}{\theta(t)} + \frac{w}{m}t \tag{2}$$

A "two hats" analysis is common in economics because laymen's terminology does not always conform with the economic concepts. What the laymen calls a "homeowner" is, for example, understood in economics as a person performing two economic functions: a consumer of housing and a landlord (i.e., owner of housing capital). Another example: a small business owner can be both a worker in the business as well as the owner of the business' capital, and his total income understood as the sum of wage and asset income.

Production costs per minute traveled are the sum of a term that falls with t and a term that is linear in t. They are independent of the number of trips taken. We therefore have a cost structure that has marginal costs for the industry that are independent of the number of trips but that the level of those costs depends on the equilibrium waiting time.

Figures 1a and 1b show the cost structure using some of the classic cost diagrams.[3] The vertical axes measures costs and prices in dollars per minute and the horizontal axes measures wait time. Figure 1a shows per-minute wait time only, t. Figure 1b repeats Figure 1a but then overlays market-aggregate wait time $t_{min}Q$ too. The black curve shows $c(t)$, assuming for Figures 1a and 1b only that $c''(t) > 0$.

---

[2] See Becker (1965).
[3] These diagrams were invented by Jacob Viner, who was teaching them to Chicago Econ301 students even before they were first published (see Viner (1932) for the published version and Viner (1930/2013, esp. the July 2 lecture) for the Price Theory lecture notes).

Figure 1b's blue horizontal line is drawn through the minimum feasible cost and represents the marginal costs of adding trips in the industry that are produced with the cost-minimizing wait time t_{min}. Figure 1b's red curve is the riders' demand curve scaled by t_{min}. The efficient quantity of rides is $Q(c(t_{min}))$: where the marginal value of the ride equals the minimum feasible marginal cost.

a. Does it matter whether the UBER handling fee is paid by passengers or drivers?

Note that UBER is an intermediary between drivers and riders. Per minute of riding, UBER receives *p* from the rider, pays 0.8*p* to the driver, and keeps the remaining 0.2*p* for itself. As the problem (and the real-world UBER) is set up, (a) the rider sees one payment that combines the revenues for UBER and driver and (b) the driver owes 0.2*p* to UBER for the privilege of being in UBER's network.

But consider an alternative setup where drivers receive *P* (note upper case) from customers and keep it all, while riders have to pay 0.25*P* as an additional fee to UBER in order to use UBER's network. The combined payment by riders is 1.25*P*, which means that drivers are still receiving 80 percent of what riders pay and UBER is receiving the other 20 percent. The difference between the *p*-setup and the *P*-setup is just the units in which prices are quoted. In the former system, prices are quoted in terms of the total rider payment, whereas the latter system quotes them in terms of driver receipt.

We assumed above that (a) driver behavior depends on driver receipts and wait time, and (b) rider behavior depends on rider payments and wait time. We did not assume that the *units* in which price are quoted matters *additionally*. Thus, the market outcomes – number of riders, number of drivers, wait times, expenditure by riders, and driver incomes – would be the same in a *p*-market with *p* = \$1 as it would be in *P*-market with *P* = \$0.80. Moreover, if *p* = \$1 is the equilibrium outcome in the *p*-market, then *P* = \$0.80 is the equilibrium in the *P*-market, and vice versa. This is not to say that UBER's intermediation fee is irrelevant, just that it could collect the fee from either riders or drivers and get essentially the same outcomes.

This is a major conclusion in the economic analysis of markets, not just rider-driver markets. Lots of evidence confirms it. The spot market for gold, for example, quotes prices inclusive of the market-maker fee, whereas the gold futures market does not (buyers in a futures market have to pay an exchange fee in addition to the quoted gold price). Nevertheless, the two gold markets react to essentially the same supply and demand fundamentals and their price paths are essentially parallel over time.

Countries around the world vary widely in terms of how they collect labor market taxes from employers versus employees. Employment and wages seem to be affected by the total tax, but not the breakdown of the tax between employer and employee.

Of course, the real-world UBER had to decide the units in which it would quote prices. For that purpose, the answer above would be unsatisfactory. But our purposes here are to understand wait times, quantities of riders, drivers' incomes, etc., for which the opportunity costs of drivers and riders, for example, are far more important than the units of price quotes. Economic theory by definition involves focusing on the important factors and abstracting from the unimportant ones.

b. If UBER ran an experiment of raising the posted price for a period of time, what would happen to passenger wait times and the number of rides? Would it matter whether UBER had announced (to potential drivers and passengers) this experiment in advance?

Assuming that drivers know about the price increase (it's their business to know such things), then drivers enter in response to the higher price. Driver entry reduces θ and t, and must reduce θ enough to stay on the supply curve (i.e., $\theta(t)*0.8*p = r$).

A general economic principle here is that people are more specialized as producers (drivers in this case) than they are as consumers (riders). Thus, if there are market participants lacking specialized knowledge, skills, or equipment, we expect them to be on the consumer side of the market. It is not particularly useful[4] to assume that drivers would be ignorant of the price of driving.

Entry by drivers enhances the "quality" of the product for consumers – more specifically, riders have to wait less for their ride and thereby demand more rides at each p. If passengers do not know about the price increase, then we have the result that there are more rides.

If passengers know about the price increase, then we need to look at the full price u of being driven. It depends on p both directly (because riders pay it) and indirectly through the wait time t:

$$u(p) = p + \frac{w}{m}\theta^{-1}\left(\frac{r}{0.8p}\right) \tag{3}$$

If we begin with an arbitrary p, then we cannot sign the effects of p on u and Q because the direct effect is to increase u and the indirect effect is to decrease it by decreasing wait times. But the p we observe without the experiment is not arbitrary: UBER presumably chose it to either maximize revenue or, as it is getting customers and riders to join the platform, to enhance Q so as to enhance future revenue. Either way, we rule out cases in which p is so low that increasing it increases both the number of rides and revenue per ride because UBER should have already implemented such price increases.

Because we are holding fixed UBER's share of the cash fare (20%), UBER's revenue is a fixed proportion of cash revenue. In this case, it helps to look at the relationship between cash revenue and the quantity of rides, as in Figure 2. Note that Figure 2 has the cash price p on the vertical axis, whereas Figures 1b has the full price u.

If p were too high, then rides would be essentially, if not literally, zero because potential riders would find the alternatives to be cheaper. But if p were less than r/0.8, then rides would also be zero because there would be no drivers and riders would find themselves waiting forever. Even a price p somewhat higher than r/0.8 would result in zero rides because, although finite, wait

[4] As always, "useful" means helping to explain behaviors in real-world markets.

times would be long enough to keep u above the choke point on the riders' demand curve.[5] This reasoning gives us two points on the "demand curve", shown as A and C in Figure 2.

More generally, Figure 2's demand curve is the inverse of the graph of $Q(u(p))$, where $u(p)$ is defined above. Now we see why the slope of u is not globally positive. Cash prices that are too low – below the value p_B shown in Figure 2 – actually result in high full prices, and thereby low quantities, because of the excessive wait time.

Figure 2 appears to violate the "law of demand" because the graph of p versus quantity slopes up over some range. The trick is that p is not the full price u that determines the demand for rides, and that p and u are inversely related over that range. A graph of u versus quantity slopes down throughout.

Points on Figure 2's demand curve between A and B have marginal cash revenue that is *above* the price, and thereby above zero. It wouldn't make sense for UBER to operate on this part of the demand curve because raising p would give it more rides and more revenue per ride. That is why we assume that UBER operates on some part of Figure 2's demand curve that is between B and C, where $u'(p) > 0$.

The answer is somewhat different if we assume that UBER adjusts its 20% share in order to minimize its losses from the experiment. This is addressed in part (c).

The amount that riders adjust to the higher u depends on whether UBER had announced the price change. A general economic principle is that market participants adjust more to price changes in the long run, when substitutable and complementary activities can be adjusted too, than in the short run. For example, forewarned riders could arrange their alternatives in order to make them most available during the time of the price-increase experiment.[6]

[5] The demand curve's (full price) choke point is defined as $Q(\text{choke point}) = 0$.

[6] A similar argument applies to driver adjustments (such as new drivers' joining the UBER network), but we have already assumed that the supply of drivers is infinitely elastic.

c. What can you say about UBER's optimal price p and percentage handling fee?

Let τ denote the handling fee (expressed as a share). Assuming that, as an intermediary, UBER has no marginal costs of rides beyond the aforementioned "production" costs incurred by drivers and riders (i.e., its electronic mapping system has plenty of capacity), it wants to maximize its revenue as intermediary τpQ. As noted above, t must be on the driver supply curve with

$$(1-\tau)\theta(t)p = r \tag{4}$$

Also recall that the passenger full price u is:

$$u = p + \frac{w}{m}t \tag{5}$$

Use the two equations above to substitute and look at UBER's revenue as a function of u and t,

$$\left[u - \frac{w}{m}t - \frac{r}{\theta(t)}\right]Q(u) = [u - c(t)]Q(u) \tag{6}$$

The term in square brackets is UBER revenue for every minute of customer travel, which (per minute of customer travel) is the customer's full payment minus the cost of customer wait time minus the cost of driver time. The two subtractions together are therefore the per-minute production cost $c(t)$. UBER's revenue maximization problem is therefore an example of the classic monopoly problem that is solved in two stages: first production techniques t are chosen to minimize cost $c(t)$ and then the price u is chosen to equate marginal revenue to the minimum feasible marginal cost.

The classic result is that monopolists charge above marginal cost but still minimize their costs, at least if the costs are not affecting the structure of demand.

Figure 3 illustrates the determination of quantity of rides under monopoly. It is similar to Figure 1b, but with the addition of a marginal (full) revenue curve and without scaling quantity by t_{min}.[7] The equilibrium quantity n_m is at the intersection of the marginal revenue curve and the marginal cost curve, which is horizontal at $c(t_{min})$. The equilibrium full price u_m is the full price that elicits demand in the amount n_m. The equilibrium cash price p_m and handling fee τ_m are:

[7] If n denotes the number of rides, then the marginal (full) revenue curve is the derivative of $nQ^{-1}(n)$ with respect to n. Note that $Q^{-1}(n)$ is a full price and not just the cash price p.

$$p_m = Q^{-1}(n_m) - \frac{w}{m} t_{min} \tag{7}$$

$$\tau_m = 1 - \frac{r}{p_m \theta(t_{min})} \tag{8}$$

If there were competition among intermediaries, then u would have to reflect the various costs, including $c(t)$. For example, if the costs of the electronic mapping system were negligible, then the competitive result would be u equal to the minimum feasible $c(t)$. The quantity of rides would be $Q(c(t_{min}))$. Note that competition results in more rides and a lesser cash price than monopoly does but the same (efficient) wait time. The competitive cash price is:

$$p_c = c(t_{min}) - \frac{w}{m} t_{min} < p_m \tag{9}$$

Price experiment revisited. Note that the full rider price u and the production cost $c(t)$ are different by UBER's revenue per minute of customer travel. As a function of wait time, the former is:

$$u(t) = c(t) + \frac{\tau}{1-\tau} \frac{r}{\theta(t)} \tag{10}$$

It follows that the full rider price u is not minimized at t_{min}. In the neighborhood of t_{min}, a reduction of t due to an increase in p – recall part (b) above – has no first-order effect on $c(t)$ or on UBER's aggregate revenue. But, holding τ fixed, equation (10) says that UBER's revenue per trip, and thereby the full rider price, increases. This result is useful in answering part (b), where an UBER "experiment" increases p without changing τ. A greater full rider price means fewer rides. In summary, an increase in the cash price p reduces the number of rides, each of which has a lesser wait time.

The appendix shows that a similar result holds if instead τ were to be adjusted in order to minimize UBER's revenue loss from the p-increase experiment.

d. Could regulators use a price ceiling to encourage more efficient outcomes?

Define efficiency to be UBER's surplus, plus riders' surplus, plus drivers' surplus, each measured in dollars.

Suppose for the moment that regulators were sure that UBER was charging $p_u > p_c$, and diligently enforced a prohibition of cash prices above a regulated ceiling that is set somewhere in the interval $[p_c, p_u)$. By definition, the regulation pushes the cash price closer to the competitive price. If the price reduction were small enough, we could apply the marginal $u'(p)$ result from parts (b) and (c) that more rides will result. Judging from price and quantity alone, the regulation appears to be a success. However, this both assumes that the regulation only nudges prices and ignores the quality effects.

On the first point, Figure 4 illustrates by drawing Figure 2 for various values of τ. It is possible that the quantity-maximizing cash price increases with τ (as it does with a linear θ function), as drawn in Figure 4. Moving from top to bottom along any one of the demand curves is associated with longer wait times. As discussed before, the curves are not monotonic because the extra wait time can more than offset the lower cash price.

Note that in the competitive case the cash price is the one that maximizes quantity on the $\tau = 0$ demand curve. If UBER operates unregulated at point *U*, forcing it to charge p_c would move the market to point *R*, which involves fewer rides and a higher full price for riders because of the extra waiting time. A cash price ceiling, even one above the competitive cash price, results in inefficient wait times (that is, they do not minimize $c(t)$) and can make riders worse off. As shown by point *D* in Figure 4, a cash price ceiling might shut down the entire market even while it is above the competitive cash price.

Even a regulator that was careful to reduce prices gently enough to avoid a negative impact on quantity (e.g., by setting its ceiling at or above p_B), it might still be harming UBER more than it helps customers because the wait times that result from cash price p_0 are inefficient.[8]

[8] Moreover, in markets where quality is a good substitute for quantity, a price ceiling can increase the quantity traded even while it harms consumers; a price ceiling's quantity impact is not always a good metric of its welfare effects.

III. A Model-Student Answer

(The student's audience is the professor/TAs. He needs only to show that he knows which elements of price theory are needed to answer the question, and skips any pedagogy. Citations to the literature are given only when they condense the communication of the economic setup. E.g., "Cobb-Douglas production function" is not intended to indicate further reading of articles by Charles Cobb or Paul Douglas, but rather to quickly indicate the properties of the production set being assumed. Other example: "Hicksian demand"; "products are differentiated in one dimension as in the Hoteling location model.")

<u>Setup the problem so that is no longer "open ended."</u>

Let t = time that passengers wait for pickup.

$1-\theta$ = the fraction of time that drivers are without a fare.

The function $\theta(t)$, with $\theta'(t) > 0$, reflects the fact that drivers wait longer for a fare when passengers wait less, and vice versa, based on the number of drivers per passenger in the market.

My answer will abstract from questions of scale; by assuming $\theta(t)$ I have a proportional setup.

Supply of drivers: Per minute, including idle time, drivers earn θ*0.8*p, where p is the price per minute paid by passengers, 80% of which goes to the driver. The higher this wage, the more drivers.

Because driving is a one low-skill activity among many, I assume a perfectly elastic supply of drivers. The supply curve is:

$$0.8\theta(t)p = r$$

Demand by riders: The product has these two attributes – price and wait time – that I combine into a single full price u per minute of being driven.

$u = p + wt/m$, where m is the length of a trip in minutes (assumed constant) and w is the passenger waiting cost per minute. Based on the law of demand: the quantity of rides demanded $Q(u)$ falls with u.

a. Does it matter whether the UBER handling fee is paid by passengers or drivers?

No. The equilibrium determines who really pays according to pass-through, cost elasticities, etc. Switching the fee from drivers to passengers will cause the drivers to earn less gross of the fee and the same as they were net of the fee. Passengers will pay less net of the fee and the same as they were gross of the fee.

This is not to say that the fee is irrelevant; just that the legal responsibility for the fee is irrelevant for the quantity traded, the total amount paid by passengers, and the net amount received by drivers.

b. If UBER ran an experiment of raising the posted price for a period of time, what would happen to passenger wait times and the number of rides? Would it matter whether UBER had announced (to potential drivers and passengers) this experiment in advance?

Drivers enter in response to the higher price. Driver entry thereby reduces θ and t, and must reduce θ enough to stay on the supply curve (i.e., $\theta(t)*0.8*p = r$).

The full price u of being driven depends directly on p and indirectly through t:

$$u(p) = p + \frac{w}{m}\theta^{-1}\left(\frac{r}{0.8p}\right)$$

If we begin with an arbitrary p, then we cannot sign the total effects of p on u and Q. This is the non-monotonic "demand curve" drawn in Figure 2.

If we assume that the initial p had been maximizing UBER's revenue, we thereby rule out cases in which p is so low that increasing it increases both the number of rides and revenue per ride because UBER should have already implemented such price increases. i.e., the total effects of p are to increase u and decrease Q.

The amount that riders adjust to the higher u depends on whether UBER had announced the price change. Market participants adjust more to price changes in the long run, when substitutable and complementary activities can be adjusted too, than in the short run.

c. What can you say about UBER's optimal price p and percentage handling fee?

Let τ denote the handling fee (expressed as a share). Assume that UBER has no marginal costs of rides beyond the aforementioned waiting costs incurred by drivers and riders. It therefore wants to maximize its revenue as intermediary τpQ. As noted above, t must be on the driver supply curve with

$$(1-\tau)\theta(t)p = r$$

Also recall that the passenger full price u is $p + wt/m$. Use these two equations to substitute and look at UBER's revenue as a function of u and t,

$$\left[u - \frac{w}{m}t - \frac{r}{\theta(t)}\right]Q(u) = [u - c(t)]Q(u) \tag{11}$$

The two subtractions in the square brackets together are the per-minute production cost $c(t)$. UBER's revenue maximization problem is therefore an example of the classic monopoly problem that is solved in two stages: first production techniques t are chosen to minimize cost $c(t)$ and then the price u is chosen to equate marginal revenue to the minimum feasible marginal cost, as in Figure 3. The monopoly u is above marginal cost.

If there were competition among intermediaries, then u would have to reflect the various costs, including $c(t)$. For example, if the costs of the electronic mapping system were negligible, then the competitive result would be u equal to the minimum feasible $c(t)$. The quantity of rides would be $Q(c(t_{min}))$. Competition results in more rides and a lesser cash price than monopoly does but the same (efficient) wait time.

d. Could regulators use a price ceiling to encourage more efficient outcomes?

In order to get an outcome with more total surplus, the price ceiling must be diligently enforced (that is, without too many loopholes) and be close enough to the unregulated price.

Such a ceiling would increase wait times, which is a second-order social loss because the unregulated wait time was at the minimum. The ceiling would increase quantity, which is a first-order addition to efficiency if the unregulated price were not exactly the competitive one.

It would NOT necessarily enhance efficiency to set the ceiling near the competitive price. If UBER operates unregulated at point U in Figure 4, forcing it to charge p_c would move the market to point R, which involves fewer rides and a higher full price for riders because of the extra waiting time.

IV. Essential Ingredients in the Model-Student Answers

The model-student answers are concise, identify the important price theory concepts, and use the concepts correctly to arrive at an answer.

IV.A. Concise but correct answers

A concise answer is achieved by leaving out unimportant details and unimportant logical possibilities. For example, the drivers are not assumed to have market power. The supply of drivers is assumed to be perfectly wage elastic because for many purposes it is in fact quite elastic.[9] At the same time, it is clear from the answer that the model student knows that he is omitting something.

Model answers can acknowledge when the direction of an effect is ambiguous. E.g., the effect of *p* on the number of rides can go in either direction. But the model answer explains what are the countervailing effects (*p* makes rides more expensive, but low wait time makes them less expensive). An A+ answer also notes when a comparative static becomes unambiguous in the neighborhood of the equilibrium: e.g., maximizing behavior by UBER rules out the case when *p* increases the number of rides).

Another example of an ambiguous answer relates to the price ceiling in part (d). The ceiling's effect, even directionally, depends on the level of the ceiling. Indeed, it would be incorrect here to conclude that the efficiency effect of a price ceiling is unambiguous.

IV.B. Key economic concepts in these answers

Model answers usually consider both supply and demand, and that is the case with answering the UBER questions above. Opportunity costs are often important. Conversely, amateur answers often forget either supply or demand, or neglect opportunity cost, even when they are important and change the answer. A number of students answering a version of the UBER question forgot that the waiting time of drivers is costly because the drivers could be doing something else.

Other economic concepts that came up in the UBER question are listed below.

<u>Setup</u>: Full price is distinguished from cash price. Cost curves are examined (to be used later to come to conclusions about aggregation). If possible, assumptions about the price elasticity of

---

[9] This does not mean that all problem set answers should begin with assuming perfectly elastic supply. The best assumption is a function of the industry being considered and the purpose of the analysis. E.g., for most purposes the supply of land is pretty inelastic. Many other industries have fairly inelastic supply in the short run.

industry-level supply and/or demand are made. E.g., there is on optimal way of producing one ride, but then any number of rides can be produced at the same cost.

Part (a): Supply price is distinguished from demand price. In other words, what the riders pay is different from what the drivers receive. Market equilibrium determines who pays UBER's fee.

Part (b): Goods and services have non-price attributes. In this case, the wait time is a non-price attribute and is accounted for by adding the cost of that time to the cash price to arrive at a full price.

Part (c): The profit function is examined and is related to costs. First-order effects are distinguished from second-order effects. Marginal revenue curves may help to arrive at an answer.

Part (d): Efficiency is defined. Non-price attributes are one of the outcomes to be examined.

V. Appendix: UBER's revenue calculus under the three control scenarios

The problem set examines three scenarios for UBER's setting its business parameters. The scenarios differ in terms of which, if any, of UBER's actions are held constant:

(0) UBER does not react, only drivers and riders do. Algebraically, p is changed by "experiment" and τ is held fixed at 20%. As a function of p, the full rider price is:

$$u(p) = p + \frac{w}{m}\theta^{-1}\left(\frac{r}{0.8p}\right)$$

The derivative is:

$$u'(p) = 1 - \frac{w}{0.8mr\theta'}\left(\frac{r}{p}\right)^2$$

For an arbitrary p, $u'(p)$ cannot be signed. But if p had been maximizing revenue given τ, then $u'(p)$ must be strictly positive. An algebraic proof comes from the expression for UBER's revenue:

$$R(p|\tau = 0.2) = 0.2pQ\big(u(p)\big)$$

The effect of p on revenue is:

$$R'(p|\tau = 0.2) = 0.2[Q\big(u(p)\big) + pQ'\big(u(p)\big)u'(p)]$$

which is positive whenever revenue is positive and $u'(p) \leq 0$. In words, if the number of rides could be increase merely by increasing p, then UBER would have already done it. This is the algebraic version of the above discussion of Figure 2.

(1) UBER reacts with $\tau \in [0,1]$ to the p experiment. Because τ increases wait times at each p, we can equivalently say that UBER reacts with $u \in [p,\infty)$ to the p experiment.[10] Here its handling fee problem is equivalent to:

$$u(p) = \underset{u}{\operatorname{argmax}}\left[p - \frac{r}{\theta\left((u-p)\frac{m}{w}\right)}\right]Q(u)$$

The comparative static of interest is $u'(p)$, because that tells us the change experienced by riders as a consequence of a p experiment with u (equivalently, τ) adjusted to minimize the loss of UBER's revenue. It is found by totally differentiating the first-order condition

[10] We can also ignore the two bounds on τ, and therefore the bounds on u, by reasonably assuming that UBER gets zero revenue at each bound.

with respect to u, which is itself the partial derivative of revenue with respect to u. If p increases u, then it much reduce Q because the demand for rides slopes down.

This is the usual procedure for calculating comparative statics of optimal solutions. Because we have only one choice variable, the second order condition implies that the sign of $u'(p)$ is the same as the sign of the cross derivative of the revenue objective between p and u. In other words, $u'(p)$ is positive if and only if p increases the marginal effect of u on revenue, when that marginal effect is evaluated at $(p,u(p))$.

The marginal effect of u on revenue is the sum of two effects: the effect of u on the square bracket term in the revenue equation above (times the number of rides Q) and the effect Q' of u on the number of rides (times the square bracket term). Recall from above that the square bracket term is the difference between the full price u and the combined time costs c. Holding u constant, p has no effect on Q or Q'. This means that the cross derivative is the sum of only two terms: the cross-derivative of the square bracket term (times the number of rides Q) and Q' times the partial effect of q on the square bracket term. The latter is just the partial effect of q on cost, which is zero at UBER's optimum because it minimizes costs $c(t)$. That leaves only one term, which is proportional to the convexity of costs.

The cross derivative is:

$$\frac{2(\theta')^2 - \theta\theta''}{\theta^3}\left(\frac{m}{w}\right)^2 rQ + \left(w - \frac{r\theta'}{\theta^2}\right)\frac{m}{w}Q'$$

The first ratio is the convexity of combined costs $c(t)$ with respect to t, and is positive in the neighborhood of UBER's optimum by the second order condition with respect to p. The parentheses part of the second term is the derivative $c'(t)$ of combined costs, which is zero in the same neighborhood (see below). p is not a choice here but this shows algebraically that we can sign $u'(p)$ if the p experiment is a deviation from UBER's optimum.

(2) UBER sets both τ and p to maximize its revenue as an intermediary. This scenario is a baseline against which the other two can be compared: e.g., the p experiment is performed starting from revenue-maximizing values for τ and p.[11] UBER's maximization problem is the same as in (1), except that p is a choice variable too. The p choice requires that $w\theta^2 = r\theta'$ and $2(\theta')2 > \theta\theta''$; in words, that p must be set so that wait time t is the value t_{min} that minimizes per-minute production costs $c(t)$.

[11] The revenue-maximizing value for τ happens to be 20% (that's why we observe it).

Figure 1a. UBER Production Costs

As a function of rider wait time t.

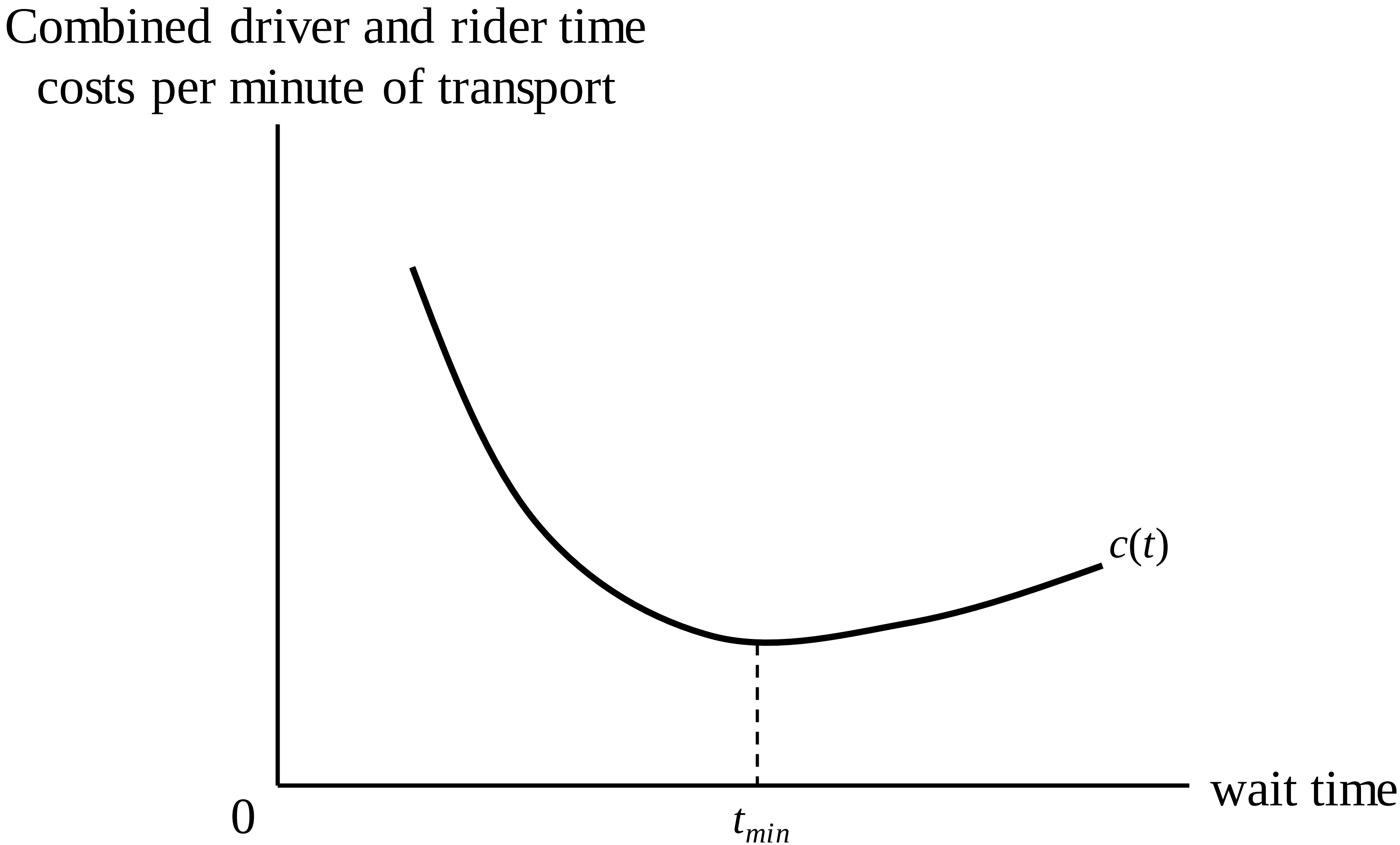


Figure 1b. UBER Costs and Demand

As a function of aggregate rider wait time $t_{min}Q$.

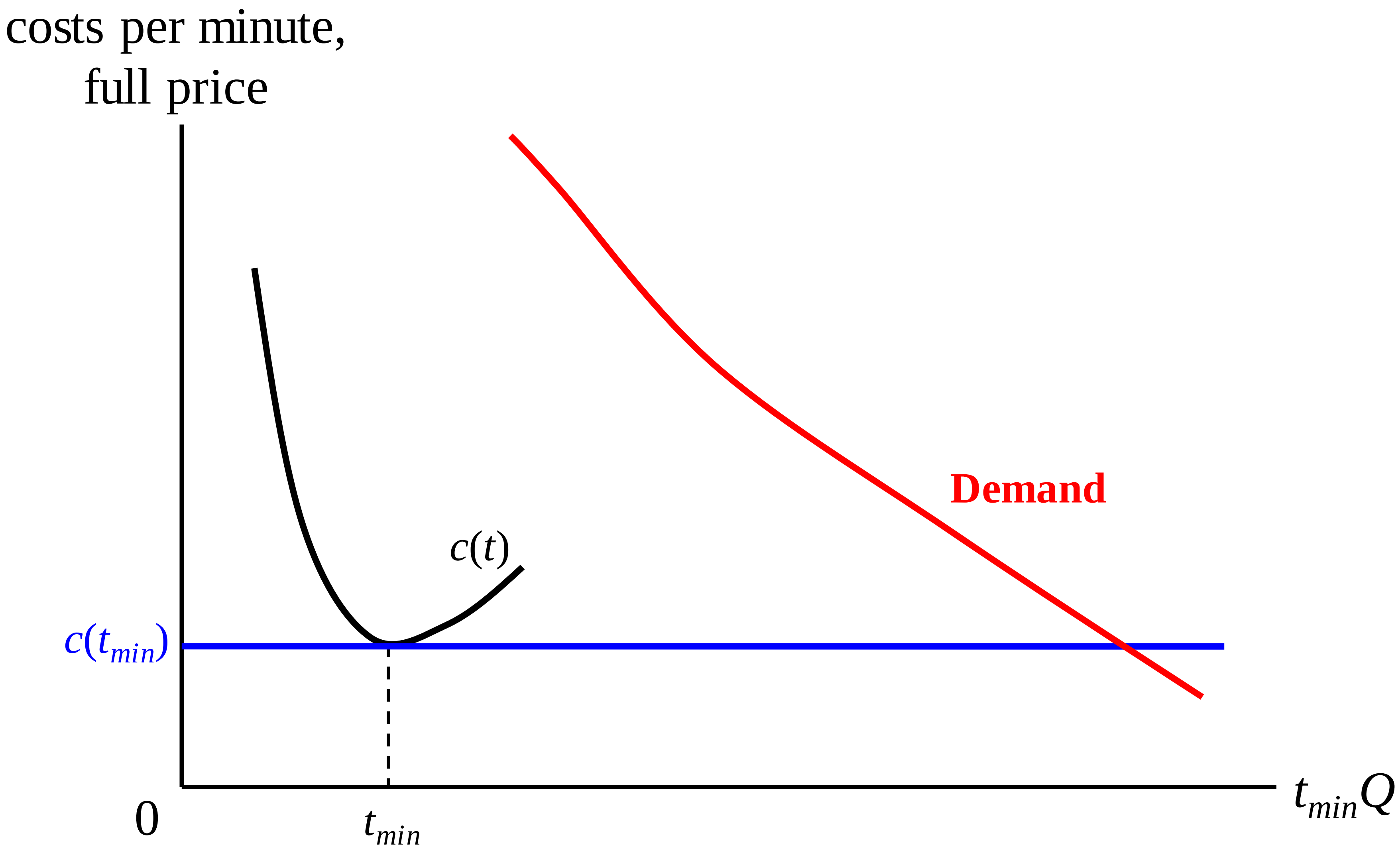


Figure 2. Quantity as a function of cash price
with a 20% handling fee.

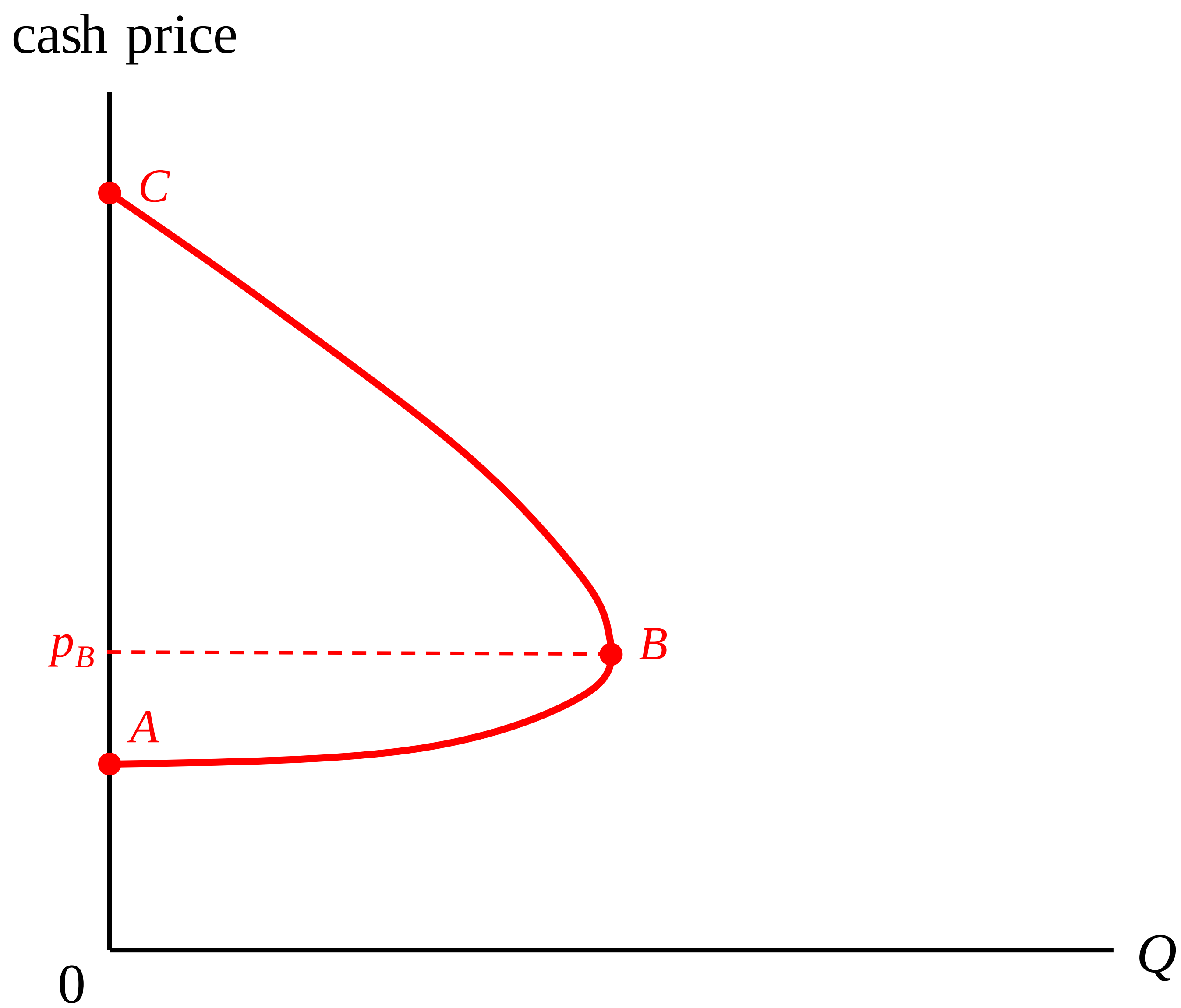


Figure 3. Monopoly full price and quantity

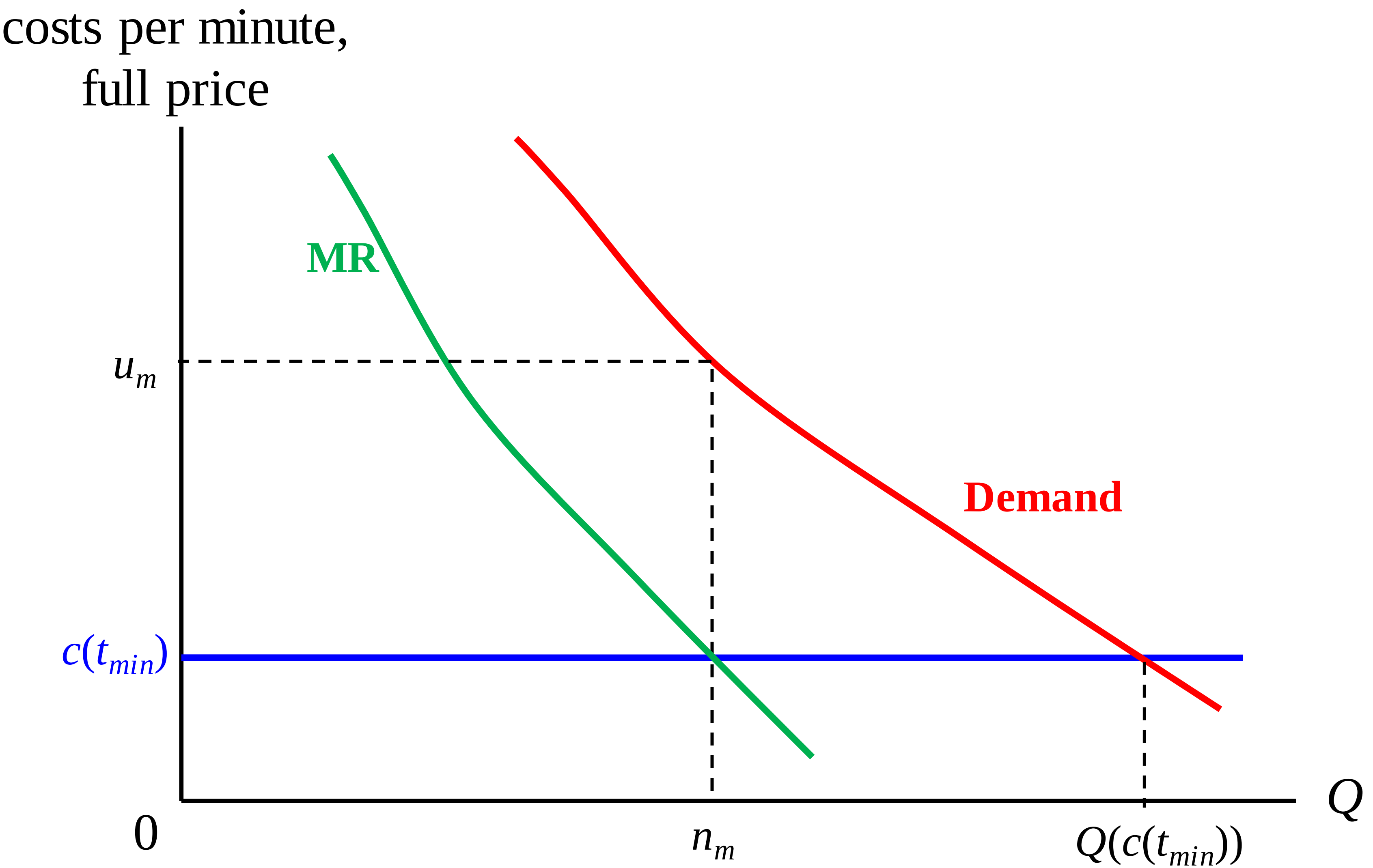
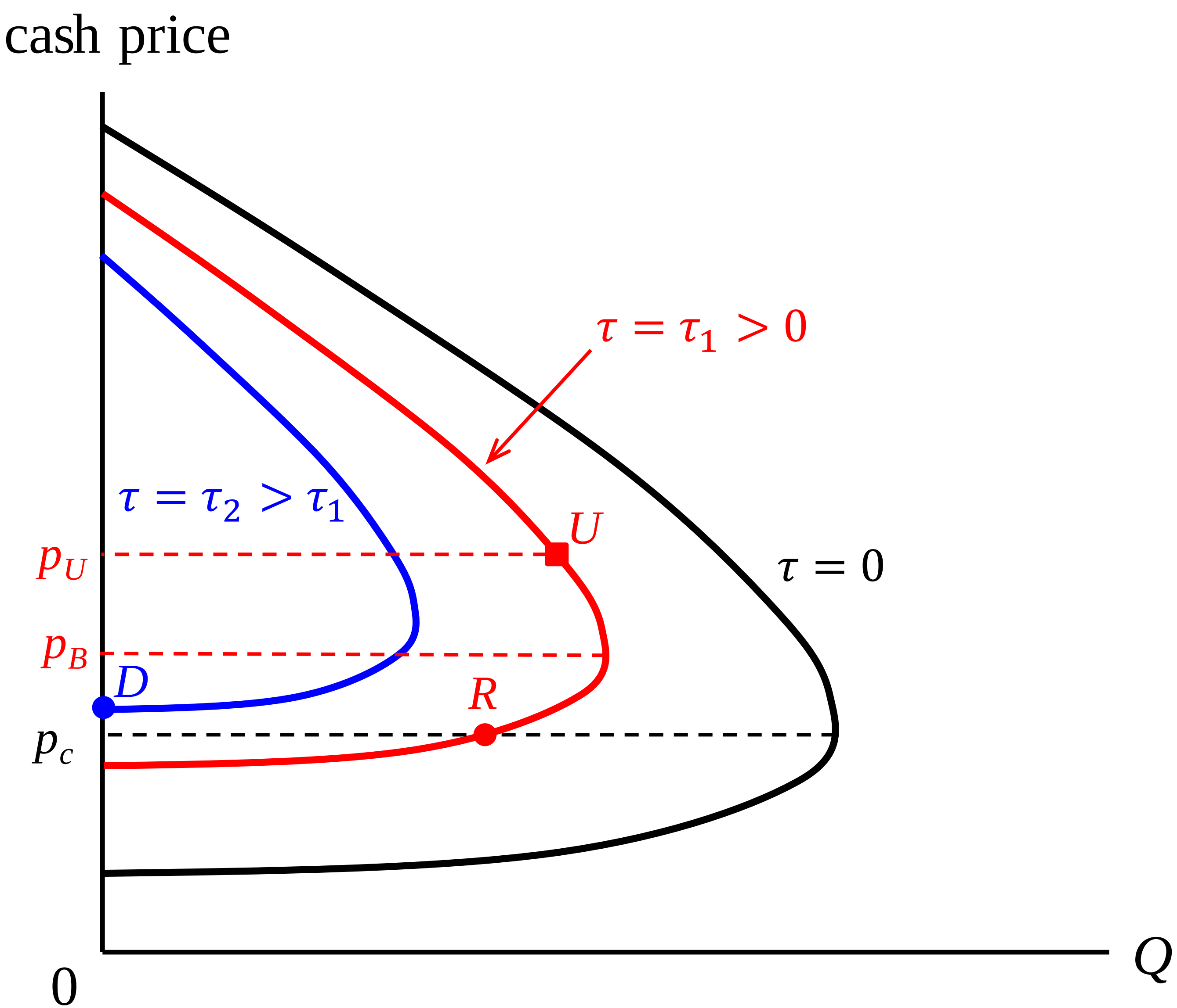


Figure 4. Quantity as a function of cash price
and the handling fee τ.



Bibliography

Becker, Gary S. "A Theory of the Allocation of Time." *Economic Journal* 75, no. 299 (September 1965): 493-517.

Viner, Jacob. "Cost Curves and Supply Curves." *Zeitschrift für Nationalökonomie* 3, no. 1 (February 1932): 23-46.

—. *Lectures in Economics 301.* New Brunswick, NJ: Transaction Publishers, 1930/2013.