Why Do We Tip Taxicab Drivers?

David Flath*

abstract
The leading economic explanation for tipping— that is, explanation why the practice is socially beneficial, not why individuals leave tips even though it is not narrowly advantageous to them— is that it confers an incentive to provide personal services. This fits many instances in which tipping is common but does not fit the taxicab business very well. We propose two novel explanations of tipping that do fit the taxi case. The first is that tipping might represent Lindahl pricing of the services of vacant cabs (essentially, reduced waiting time), a local public good for taxi customers. The second is that the actual fare for a cruising cab is de facto a matter of negotiation between the cab driver and the customer. Tipping thus represents a frank acknowledgment that the posted fare or regulatory determined fare is non-binding; the tip is simply the portion of the agreed fare in excess of the regulated one.

JEL codes: D40, D82, L52, R48.
Keywords: tipping, taxicabs, Lindahl pricing, social norms, regulation

*Dept. of Economics, NCSU, Raleigh, NC 27695-8110
tel. (919)515-4617; email. david_flath@ncsu.edu
Why Do We Tip Taxicab Drivers?

1. Introduction

The recent boom in behavioral economics has induced a boomlet on tipping. Recent literature such as Azar (2004) and Conlin, Lynn and O’Donoghue (2003), has focused on why individuals leave tips even though it is not narrowly advantageous for them to do so. Leaving a tip is an example of conformity to a social norm. The wish to avoid social ostracism, and not the pursuit of economic gain, is what induces customers to leave tips. And yet, it is natural to conjecture that social norms themselves evolve in a way that is economically efficient. For example, unless the tipping practice itself has some economic benefit to society it perhaps would not have become a social norm in the first place. This line of thinking opens the way for economic explanations of the pattern of social norms, even absent a complete understanding of the precise mechanism by which economically efficient norms come to be established, and inefficient ones fail. One simply adopts as a working hypothesis, a conjecture, a guess, that such a mechanism exists. For the sake of argument, if nothing else, let us follow other authors in adopting this premise.

What then is the economic benefit of the social norm of tipping? A leading explanation, first proposed by Ben-Zion and Kami (1977) and further developed by Jacob and Page (1980), is that it confers an economic incentive to provide personal services. So, for instance, the prospect of receiving a tip induces a restaurant waiter to be attentive to each customer, even if the waiter’s employer, the restaurant proprietor say, is unable to monitor the situation or has greater costs of doing so than does the customer. This explanation comports with the fact that nearly all instances in which tipping is common do involve personal services, the quality of which is very much determined by the one who receives the tip. But the explanation does not fit the taxi business particularly well, for the “quality” of the taxi service seems little related to the driver’s behavior. Authors have suggested that the practice of tipping taxi cab drivers confers incentives for cab drivers to attend to the safety of their passengers or to avoid circuitous routes. This is not entirely convincing because even without tipping, cab drivers would still have fairly strong incentives to comply with these very modest expectations of their passengers. After all, the driver’s own safety is at risk in the event of his reckless driving. And such blatantly fraudulent behavior as taking a
circuitous route simply to inflate the fare would seem to justify the passenger’s not paying the fare, and in fact reporting the abuse to the relevant authorities. The passenger is from out of town and so doesn’t know that the driver has taken a circuitous route? Then how would he know to withhold the tip? That tipping affords an incentive to provide personal service fits the case of the restaurant waiter much better than it does the taxi cab driver.

This paper presents two novel economic explanations for tipping—the common practice of customers, at their own discretions, adding amounts to the billed charges for services. One of the explanations applies only to the taxi business. The other might also apply to tipping in restaurants and other venues where the quality of the product has some character of a public good, that is that the quality is the same for all customers and not adjusted to the specifications of each one separately. My novel explanations for tipping may coexist with the leading explanation, that it is to confer an incentive to provide personal services that the customer can better monitor than can the employer of the service provider.

2. Tipping as Lindahl pricing of a local public good

Often, “quality” has the character of a public good. Where this is so, the market will in general not attain the socially optimal level of quality under ordinary pricing. Pourboire—tipping—in which customers voluntarily pay surcharges in amounts of their own individual choosing, is not ordinary pricing. Mightn’t it be the case that the rules customers follow in deciding how much to tip in fact induce the socially optimal quality? In effect tipping might resemble Lindahl pricing of a local public good. This has particular relevance to the cruising taxi cab industry.

The quality of taxi service corresponds to the average waiting time for a vacant cab. Within a particular geographic domain served by cruising cabs or radio dispatched cabs, the average waiting time depends upon the number of vacant cabs. As first noticed by Douglas (1972) and Mohring (1972) this imparts industry-wide economies of scale. Holding quality of service constant means holding the average number of vacant cabs constant. The cost of operating the set number of vacant cabs is a fixed cost independent of the quantity of cab services. But what is the economically efficient number of vacant cabs and will the market attain that? As suggested by Arnott (1996), the
number of vacant cabs and implied average waiting time for a cab has the character of a local public good. The economically efficient allocation thus requires Lindahl pricing of the cab services. That is, each cab customer must be dunned—in addition to a price equal to the marginal cost of the cab while he occupies it—an extra fee that equals his own marginal valuation of vacant cabs. The economically efficient number of vacant cabs is that such that the summed marginal valuations of all cab customers just equals the marginal cost of operating vacant cabs. If the tips of each cab customer correspond to the Lindahl prices of vacant cabs, regulators have then only to set the posted fare equal to the marginal cost of occupied cabs and allow free entry, to assure the economically efficient allocation.

In fact this is a possible basis for tipping generally, not restricted to the cruising taxi case, but in other instances where the argument might apply local monopoly is necessary. The cruising cab example is special in that the quality of cab service is an industry-wide public good even where the industry is itself atomistic. Where quality of a product is a local public good only for the customers of one firm, and there are many firms, then hedonic pricing can induce efficient choices of quality without Lindahl pricing. But where the firm is itself a monopoly there might be some basis for Lindahl pricing à la tipping to induce efficient quality choice by the one firm. Let us consider these arguments a bit more precisely, focusing on the taxicab case.¹

Suppose that the cost of producing and customers’ valuations depend upon the quality of a good as well as the quantity. The socially optimal quality maximizes social welfare:

\[
\max W(q,s) = \int p(x,s)dx - c(q,s),
\]

where \(q\) is quantity, \(s\) is quality, \(p(x,s)\) is the inverse demand function and \(c(q,s)\) is the cost function. The necessary conditions for maximum social welfare are

\[
p - \frac{\partial c}{\partial q} = 0 \tag{2}
\]

and

\[
\int \frac{\partial p(x,s)}{\partial s} dx - \frac{\partial c}{\partial s} = 0. \tag{3}
\]

The first condition is the familiar “marginal cost equals price”. The second condition means that the

¹ Some of this follows the discussion and notation of Tirole (1988), pp. 100-102.
incremental social value of quality enhancement equals the incremental cost. This condition too is familiar; it is the condition for optimal allocation of a public good. Here quality \( s \) has the character of a local public good for all the consumers. The condition means that the summed marginal valuations of quality of all consumers just equal the marginal cost.

Now in the taxi case it is reasonable to let “\( s \)” stand for the average density of vacant cabs and “\( q \)” that of occupied cabs. It costs about as much to operate a vacant cab as it does an occupied one and each is subject to constant returns to scale. Thus

\[
c(q,s) = \bar{c}q + \bar{c}s,
\]

(4)

where \( \bar{c} \) is the constant unit cost. If the regulatory authority stipulates a fare \( p^* \), and there is no tipping, and there is free-entry, then the number of vacant cabs will on average be such that

\[
p^*q = \bar{c}q + \bar{c}s.
\]

(5)

Thus the ratio of vacant to occupied cabs will become

\[
s/q = (p^*- \bar{c})/\bar{c},
\]

(6)

and the number of vacant cabs need not be socially efficient. In particular, the overall allocation cannot be economically efficient unless all customers are alike. If each values quality \( s \) differently, then each must pay differing prices if each’s consumption of taxi services is to be economically efficient. Essentially this same point is made by Beesley and Glaister (1983, at p. 611).

Now suppose that each customer, in addition to the payment of the base fare \( p \), also pays the cab driver a tip that mimics Lindahl pricing of the corresponding vacant cabs. The Lindahl prices of vacant cabs are

\[
\partial p(x,s)/\partial s,
\]

(7)

where \( x \) orders the taxi services \( q \) by marginal valuation (of occupied cabs) of the customers. At these “prices” the consumers all desire the socially efficient quality \( s^* \), and if the regulators have set the fare equal to the marginal cost of operating an occupied cab \( p^* = \bar{c} \), then zero profit is just attained when the allocation of occupied and vacant cabs is efficient. If the regulators set the fare above this, then under free entry there will be too many vacant cabs under reasonable presumptions about demand. The summed tips collected from all customers, under the tips as Lindahl prices regime, equal

\[
s \int \partial p(x,s)/\partial s \, dx.
\]

(8)
If the regulators set the base fare above the unit cost, \( p > \bar{c} \), and the tips continue to mimic Lindahl pricing of vacant cabs even as more cabs enter or exit, then zero profit is reached when the marginal social value of vacant cabs is less than the marginal social cost:

\[
\int_0^q \frac{\partial p(x,s)}{\partial s} \, dx + (p^* - \bar{c}) \frac{q}{s} = s \bar{c},
\]

and so

\[
\int_0^q \frac{\partial p(x,s)}{\partial s} \, dx = \bar{c} - (p^* - \bar{c}) \frac{q}{s}.
\]

\[
< \bar{c}, \text{ if } (p^* - \bar{c}) > 0
\]

In other words, if \( (p^* - \bar{c}) > 0 \), then with free-entry there are more vacant cabs than is economically efficient (and fewer occupied cabs than is economically efficient), under the reasonable presumption that \( p(q,s) \) is convex. If \( p^* = \bar{c} \), then with free-entry the number of vacant cabs is economically efficient, as is the allocation of occupied cabs.\(^2\)

Does this explain why we tip cab drivers? Some readers of an earlier version of this manuscript suggested that the notion that taxicab tips are Lindahl prices of vacant cabs fails minimal standards of plausibility, on the grounds that tips do not vary across customers, or do not vary as much as one imagines that the Lindahl prices of vacant cabs would vary. Well. Taxicab tips vary more than one might have supposed judging from the survey data collected by Ayres, Vars, and Zakaria (2003). Based on 1066 observations of 12 different New Haven, CT taxicab drivers in 2001, they found an average tip rate of 15.8 percent of the fare (an average tip of $1.22), but with a standard deviation of 27 percent of the fare. In fact 23.8 percent of the passengers left no tip at all. It is clearly not the case that every customer pays the same “standard” tip. There is variation in tipping. But what about the pattern of variation?

A principal finding of Ayres et al was the existence of racial disparities in tipping rates.

\(^{2}\)Notice that the average tip rate can be related to the vacancy rate. Specifically, the vacancy rate \( \theta \) is defined

\[
\theta = \frac{s}{q}/(s/q + 1),
\]

so the average tip rate under free-entry and zero profit is

\[
\int \frac{\partial p(x,s)}{\partial s} \, dx / p^* q = (p^* - \bar{c}) / p^* - s \bar{c} / p^* q.
\]

\[
= (p^* - \bar{c} \theta/(1-\theta)) / p^*.
\]

An average tip rate of 15 percent and an average vacancy rate of 46 percent would be consistent with zero profit if the set far were equal to the unit cost of an occupied cab. The actual vacancy rate in New York City averages around 35 percent.
Specifically they found lower tips by minority passengers and lower tips for minority drivers. They speculate on the reasons but reach no definite conclusion. But if, as Ayres et al (2003, at f.n. 103) suggest, some drivers have discriminated against African American passengers by refusing to pick them up, then the lower tip rates of African American passengers comport well with the tips as Lindahl prices story. That is, African American passengers’ valuations of additional vacant cabs would logically be less than those of other passengers, ceteris paribus, because some of the cabs would not stop for them. And so their Lindahl prices of vacant cabs would be less and they would tip less.

All I want to argue here is that the Lindahl pricing explanation for tipping cannot be dismissed offhand, that the argument merits further investigation. However, I think its veracity would be most plausible when two conditions are met. The first condition is that cab customers differ significantly from one another in their valuations of quality, so that the possible improvement in allocation due to Lindahl pricing of vacant cabs is substantial. The second condition is that regulators have set the base fare near the unit cost of operating an occupied cab and have allowed free-entry, so that the actual improvement in allocation (in moving to the tipping regime from the no-tipping regime) is substantial. Both of these points follow from the working premise of this essay: Regimes in which individuals adhere to social norms are viable only if they represent Pareto improvements compared to regimes in which nobody adheres to the social norms.

The conjecture here that tipping might be a form of Lindahl pricing where quality is a local public good can be extended beyond the taxicab example. But in other instances where it might apply, quality is likely to be set by a monopoly supplier. Then if customers’ tips are tied in some way to their marginal valuations of quality—a kind of Lindahl pricing—, the monopolist is induced to supply the efficient quality. See the appendix for a precise representation of this argument.

My own guess is that this explanation for tipping, that it is a kind of Lindahl pricing of quality, might apply to the cruising taxi industry and possibly also to other cases. But for the taxi industry there is yet another explanation for tipping.
3. Tipping to attain market pricing when regulators stipulate a posted price

The taxi industry has been subject to government regulation throughout the world almost since its inception. The economic reasons are not entirely clear. One common thread in the recent literature on taxicab economics is that the fares of cruising cabs are subject to bargaining and negotiation between the cabs and individual customers. The point of regulating fares might then be to avert that process and thus avoid the costs associated with haggling or bargaining (Cairnes and Liston-Heyes (1996)). Another point might be that the outcome of such bargaining reflects local monopoly and its attendant waste, and so there is a possibility that regulation can improve resource allocation (Flath (2003)).

In most major US cities, municipal commissions stipulate the taxi fare structure. Some also limit entry. For instance the medallion system of NYC, a staple of the economics textbooks, limits the number of vehicles that can legally be operated as taxi cabs. A commission also stipulates fares. But where the custom of tipping the cab driver exists, how effective is the regulation of fares per se? Wouldn’t the regulated fare plus the tip amount to the same thing as a laissez-faire price, whatever is mutually agreed between passenger and cabdriver? If the taxi commission stipulates a higher fare, wouldn’t passengers simply reduce their tips by an offsetting amount? In fact, if this is true, then perhaps the existence of a regulatory stipulated fare is the very reason for the tips in the first place.

Consider the situation of a cruising taxi industry in the complete absence of fare regulation. It might seem very natural to model this case as one in which each taxi precommits to a posted price and fully informed demanders then choose to wait for a cruising cab or not. But as Frankena and Pautler (1986, at p. 137), and others have observed, a pure-strategy Nash solution does not exist under this regime. In particular, if all cabs but one set the same price, then the one enlarges its own revenue and profit by slightly raising its price above the others. The unlikelihood that any one demander will hail this cab is small and so its price has a negligible effect on the market demand. Yet the one whose misfortune it is to hail that cab will not wave it off because this would mean waiting for the next vacant cab and waiting is a bad. All cabs reason the same way. So no equilibrium price exists. Under laissez faire, the cab fare paid by each customer is the result of bargaining between him and the one taxi. A model that incorporates this feature is
needed to dissect the laissez faire case. See Flath (2003) for an example of such a model.

My point here is that where the custom of tipping cab drivers exists, it is reasonable to
suppose that the fare including tip reflects bargaining between each cab and customer, even if
regulators stipulate a fare. That is, the portion of the fare labeled “tip” is simply the difference
between the mutually agreed fare and the fare stipulated by regulators. If the regulators stipulate
a fare that is in most instances close to the outcome of bargaining then the custom of tipping the
cab driver would either disappear, or remain solely as a way of setting Lindahl prices on the
vacant cabs, as discussed in the previous section.

Conclusion

That tipping is an example of conformity to a social norm is obvious. In other words it is not
narrowly rational to tip at all. But one presumes that when all conform to the social norm, the
society is better off than if none did. In other words, when all tip according to the customary
rule, resource allocation is improved. But why? The leading answer to this question so far has
been that the prospect of receiving a tip induces diligence by the providers of personal services;
it is an efficient way of monitoring personal servers. This doesn’t fit the taxi business very well.
So maybe the premise is wrong. Maybe society would be better off without the custom of tipping
cab drivers. A recent analysis by Ayres, Vars, and Zakaria (2003) of the tipping patterns in the
New Haven, Connecticut taxi business, found disturbing racial disparities. The authors propose
regulations to disallow tips altogether, so that inequities in payment due to the race of the drivers
or customers would be eliminated.

We have proposed two novel explanations as to why the social norm of tipping taxi drivers
might be socially efficient. The first is that if tips correspond to Lindahl prices of vacant cabs,
it might improve the allocation of cab services. The second is that the price of a taxi ride is de
facto a matter of bargaining between each cab driver and customer. A tip paid to the cab driver
may be nothing more than the portion of the agreed fare in excess of the fare stipulated by
regulations. If the first explanation is true then regulations intended to eliminate tipping of taxi
drivers could damage resource allocation. If the second explanation is true, then regulations intended to eliminate tipping of taxi drives would likely prove futile.

This very speculative essay is intended as nothing more than an opening salvo. The empirical relevance of these ideas is largely untouched here, and may be worth pursuing in future research.
Appendix. Tipping can induce optimal choice of quality by a monopoly.

As in Spence (1975), the monopolist’s choice of quantity and quality fulfills

\[
\max_{q,s} \Pi(q,s) = q \ p(q,s) - c(q,s)
\]

(A1)

The necessary condition for profit maximizing choices are:

\[
p + q \frac{\partial p}{\partial q} - \frac{\partial c}{\partial q} = 0
\]

(A2)

and

\[
q \ \frac{\partial p}{\partial s} - \frac{\partial c}{\partial s} = 0.
\]

(A3)

The first is the familiar condition “marginal revenue equals marginal cost”. The second condition is that the marginal contribution to revenue due to enhancement of quality just equals the marginal cost of quality enhancement. In the special case in which each customer buys at most one unit of the good, it is correct to say that the monopolist gauges whether to enhance quality based only upon the effect on the one marginal customer’s willingness to pay. Whether the monopolist chooses too low a quality or too high a quality from the standpoint of social welfare depends upon whether the marginal customer’s marginal valuation of quality \( \frac{\partial p}{\partial s} \) is higher than the average of all customers’ \( \frac{\int_{0}^{q} \partial p(x,s)/\partial s \ dx}{q^*} \), or lower than the average. Either is possible.

Now imagine that in addition to paying the posted price customers leave tips. Suppose further that each individual’s tip deviates from the standard tip (to be defined shortly) by an amount that equals the deviation of his consumer surplus from the average consumer surplus per unit sold if the quality were socially optimal given the quantity supplied. For example if the standard tip is zero (and negative tips are admitted) the total tips collected by the monopoly supplier are

\[
T(q,s) = \int_{0}^{q} p(x,s)dx - q \ p(q,s) - \int_{0}^{q} p(x,s*)dx + q \ p(q,s*)
\]

(A4)

where \( s^* \) fulfills the condition :

\[
\int_{0}^{q} \frac{\partial p(x,s^*)/\partial s}{\partial s} \ dx - \frac{\partial c(q,s^*)}{\partial s} = 0.
\]

(A5)

The monopolist’s total revenue now becomes

\[
q \ p(q,s) + T(q,s) = \int_{0}^{q} p(x,s)dx - \int_{0}^{q} p(x,s*)dx + q \ p(q,s*).
\]

(A6)
The monopolist’s choices of quantity and quality fulfill
\[
\max_{q,s} \Pi(q,s) = \int_0^q p(x,s)dx - \int_0^q p(x,s*)dx + q p(q,s*) - c(q,s).
\] (A7)

The necessary condition for profit maximizing choices are:
\[
\int_0^q \frac{\partial p(x,s)}{\partial s} dx \cdot \frac{\partial c}{\partial s} = 0,
\] (A8)

and
\[
p(q,s) + q \frac{\partial p(q,s*)}{\partial q} - \frac{\partial c}{\partial q} = 0.
\] (A9)

The monopolist chooses the economically efficient quality \(s^*\) but supplies a smaller quantity than would be efficient. Furthermore, at the monopolist’s optimum \((q^*,s^*)\) the total tips collected equal zero \(T(q^*,s^*)=0\).

The tip rule that induces efficient quality choice is not unique. The stated rule is one of many, indeed infinitely many, but is among the simplest. Here is another, slightly less simple, rule: Each individual’s tip deviates from the standard tip by an amount that equals the deviation of his consumer surplus from the average consumer surplus per unit sold if the quality were socially optimal given the quantity supplied MINUS the deviation of his consumer surplus from the average consumer surplus per unit sold in the regime without tipping. In other words the tip of each consumer measures the enhancement in his own surplus induced by the tipping regime itself. Here if again the standard tip is zero (and negative tips are admitted) the total tips collected by the monopoly supplier are exactly as before and so are the monopolist’s objective function and optimal choices. One feature of this second tipping rule just described is worth noting. It is Pareto optimal compared to the standard pricing, no tipping regime. If all customers faithfully follow that rule and the monopolist responds accordingly by choosing the efficient quality, then all—both the monopolist and all customers— are better off than under standard pricing. And the tipping regime itself thus represents a Pareto improvement in the strong sense (All actually are made better off by it).

The above formulations allow the possibility, indeed require, negative tips. Customers receiving the least consumer surplus actually pay less than the posted price. This obvious departure from reality can be avoided by labeling some portion of the posted price the “standard tip”. More precisely, let us define the standard tip as the difference in payment by a customer who leaves the average tip and one who leaves the smallest tip (when the socially efficient quality is supplied).
References


