A Statistical Analysis of Pizza Delivery Tips

By xxx (deleted)

EC 413

Introduction

For people working in service positions such as waiting on tables or delivering food, tips from customers represent a very large percentage of the actual money earned. Many such positions actually pay less than minimum wage salary, which makes tipping the most important factor in making money. In this case, wouldn't it be valuable to have a general idea of which people will tip better than others? Now that doesn't necessarily mean that it would be a good idea to blatantly ignore or skip over certain customers, but it is certainly possible to make better decisions (think prioritizing) with more information. This is the reason that I decided to do my project on tips I've earned delivering pizzas. With over a year of experience, I do have some ideas about tipping tendencies, but with this project I aim to put these generalities into concrete form. Most of the time, I leave the store with more than one delivery to make, so I am faced with a decision about which order to deliver first. I can consider how old the order is, the easiest driving route for me, the predicted marginal cost of time in terms of each tip (will this person's tip change if the food gets there sooner or later?), etc. Having more information about the people I'm delivering to can only help this process.

Data

For each order that is placed, labels are automatically printed out by the computer system to go on each box. Part of this printout is a driver's ticket, which tells the driver the customer's name, location, phone number, etc. Since I have to keep track of these every night to know how much money I am owed when I leave the store, it made a convenient way to collect data.

Dep. Variable - Ltips	**Natural Log of the tip received in $**	**Continuous**
Ind. Variable - Sex	**Gender of customer**	**0=male, 1=female**
Ind. Variable – Race	**Race of customer**	**0=white, 1=non-white**
Ind. Variable – Age	**Approximate age of customer**	**0=under 25, 1=25 and older**
Ind. Variable – Greek	**Greek status of customer**	**0=non-Greek, 1=Greek**
Ind. Variable – PMT	**Form of payment (ACT card, cash, check, C.C.)**	**0=Act card, 1=other forms**
Ind. Variable – Lcost	**Natural Log of the cost of the order in $**	**Continuous**

After each delivery, I made notations on the label to indicate the customer's sex, race, approximate age, and Greek status, as well as how much they tipped. The total cost and the method of payment were both already on the label.

It should be noted that a "1" for Greek was only given in two cases: 1. If the delivery was actually made to a fraternity/sorority house or 2. If it was blatantly obvious the customer was affiliated with a Greek organization (e.g. Greek letters painted on the windows). Surely there were a handful of customers that were misrepresented in this way, but not enough to have a significant effect on the data. Despite their dummy variable status, Race, Age, and PMT were not recorded as simply "0" or "1". In each case it made sense to combine some categories after observing the data. For Race, there were samples for White, Black, Hispanic, and Asian. The last three, however, were simplified into "non-white" due to the fact that only one of 309 samples was Hispanic, and two were Asian. Age was similarly transformed, since only eight out of 309 were over 50. ACT cards were used to pay for 238 of the 309 samples, making it yet another dummy variable candidate. Additionally, 18 instances of $0 tip were changed to $.01 in order to accommodate SAS and the regression analysis.

A breakdown of the 309 samples:

Category	White	Black	Hispanic	Asian	Male	Female
Sample size	258	48	1	2	178	131

Category	<25	25-50	50+	Non-Greek	Greek	ACT Card	Cash	Check	Credit Card
Sample size	277	24	8	259	50	238	40	12	19

Additional descriptive statistics from the tip data:

Max	8.00
Mean	1.93
Mode	1.60
Median	1.60
σ	1.14
σ^2	1.30

Tip Value($)	0.01	0.80	1.00	1.60	2.00	2.40	3.00	3.20	4.00
Frequency	23	21	3	106	16	47	10	7	12

Tip Range($)	0-0.99	1.00-1.99	2.00-2.99	3.00-3.99	4.00+
Frequency	54	121	83	29	22

One note about the Tip Value frequency distribution: since the university takes a 20% cut of every ACT card order (including the tip), I only receive $1.60 out of a $2 tip. This is why an "uneven" number like $1.60 is the most frequent sample. A case could be made to count that as $2, but I chose to count the tip that I actually received.

Formulating and Choosing a Model

Rationale behind the inclusion of each independent variable:

- **Sex –** One explanation I have heard as to why males might tip better than females is that since males typically pay for most things, they know better what is expected in a tip. Now I think that is pretty outdated, but I can see some logic in it as well. This may have very well been true half a century ago when males actually did have most of the money, but there are way too many women that work and support themselves today for that theory to have any weight. Nevertheless, this variable still warrants inclusion on simplicity, if nothing else.
- **Race –** I think of this more along the lines of socio-economic level than actual race, but this was the most convenient variable to include along those lines. Income would be a much better indicator here I think, but that would require much more soliciting than was required for this data collection (which was actually none).
- **Age –** There is every reason to believe that different age groups would have different tipping patterns. I see an elderly veteran who wants things exactly his way, for example, as being much more likely to change his tip based on delivery time than a college student who is just happy to receive his food. The general perception (and general reality unfortunately) is that college students don't have

any money, so why would they tip well, if at all? The ACT card and dining dollars definitely have an impact on that theory, however.

- **Greek –** Since Greek members must pay a fairly substantial fee to be a part of their organization, it is assumed that the average family income for Greeks is higher than that of normal students. That begs the question, do people with more money tip better since they can afford it? I know I have heard other drivers say that Frats/Sororities don't tip very well, so I was very interested to see what the regression analysis would say about this one.
- **PMT –** I think to many students, especially at the beginning of a semester, the $300 in dining dollars that their parents were forced to pay for seems like free money. This would lead me to believe that ACT card tips would be the way to make money, but the fact that UA gets 20% evens the playing field back out. Also, these samples were taken from the middle to the end of the fall semester, so most people had blown through that initial free money euphoria. I actually expected non-ACT card tips to be slightly higher. If nothing else, I feel like there is more volatility in say, cash tips. Most of the time, tips are placed on the ACT card over the phone when they place the order, whereas cash tips seem to be more based on factors that I can control, such as delivering in a timely fashion.
- **Lcost –** I think most people, including myself, at least consider the total when thinking about leaving a tip. It is fairly standard to leave 15-20% at a restaurant, but I don't think there is a standard percentage for delivery. Obviously, though, the $2 tip seems to be pretty standard, as over one-third of the samples were $2 ACT card tips. Personally, I feel like this is more of a word-of-mouth phenomenon than anything else. I think it is very possible that college students

aren't used to tipping delivery drivers before going to college, and thus follow the example of others in this case. There is every reason to believe, however, that a person will tip more on a $40 order than a $15 order. It is simply more trouble to deliver seven pizzas than one or two pizzas.

Initially all six of these independent variables were included in the forecasting model. When the SAS output for this regression was analyzed, however, several of the variables failed to be significant. Therefore two additional regressions were run to attempt to find the optimal model. As is shown in the Table to the right, Model 2 dropped the Sex and PMT variables, while Model 3 additionally dropped the Greek variable. The SAS output for each of these three models is included at the end of the paper. While eliminating variables that don't show significance is a good place to start, it does not justify choosing a model by itself. Therefore, the models were compared using Adjusted R^2, AIC, and F-tests. While it wasn't a unanimous decision, Model 3 seemed to be the best, having the lowest AIC, 2nd highest Adjusted R^2, and best F-test results.

Model 1	**Model 2**	**Model 3**
Sex		
Race	Race	Race
Age	Age	Age
Greek	Greek	
PMT		
Lcost	Lcost	Lcost

	Model 1	**Model 2**	**Model 3**
Adj R^2	.1773	.1806	.1789
AIC	.6318	.6196	.6105
F-test			Chosen over others

Interpretation of Results

Ltips = -1.6229 – 1.3505 **(Race) -** .7275 **(Age)** + .7304 **(Lcost)** + e

This is the final equation, using the coefficients from the Model 3 SAS output, which I will use to make my out of sample forecasts. A brief explanation of the coefficients:

- **-1.6229 –** this is the y-intercept of the linear regression
- **-1.3505 (Race) –** if the customer is non-white (Race=1), then the expected tip decreases by 135%.
- **-.7275 (Age) –** if the customer is over 25 years of age (Age=1), then the expected tip decreases by 72.75%.
- **.7304 (Lcost) –** as the cost of the order increases by 1%, the expected tip increases by .7304%.
- **e –** this is the error term of the linear regression

Additional variables from Model 1 that are not statistically significant, but are still worth mentioning, if only for amusement:

- -.1391 (Sex) – if the customer is female (Sex=1), then the expected tip decreases by 13.91%.
- -.2457 (Greek) – if the customer belongs to a fraternity or sorority (Greek=1), then the expected tip decreases by 24.57%.
- .02133 (PMT) – if the customer pays with a method other than an ACT card (PMT=1), then the expected tip increases by 2.133%.

Evaluation of Results

In order to evaluate the forecasting ability of this model, one random sheet containing 20 samples was held out of the data collection. The attached excel printout shows the breakdown of each of the samples, as well as each of the three models' predictions for the tip (Tip1, Tip2, Tip3). MSE is then found by taking each of the e^2 columns, which represent $(Tip-Tip1/2/3)^2$, and averaging them to find Mean Square Error. Obviously Model 3 is the one I'm most concerned with, but the other two are included just for comparison's sake. Model 3 does rate slightly better than Model 2, and moderately better than Model 1, as expected. This result would depend a great deal on the data, however. Cases with large, unexpected tips, would actually make Model 3 look worse. Due mainly to a larger negative y-intercept, in cases where most of the dummy variables are 0, the third model tends to have a lower estimate than the others. All this means, however, is that the three models are very similar, which is obvious.

Conclusion

Just from glancing at the data, it becomes pretty clear that most of my deliveries were to white college students paying with ACT cards. This particular store is actually the second highest grossing Dominos branch in the entire country, mainly because of all the college students. At any given time during the school year, I would estimate about 75-80% of the deliveries on the screen are to campus. There are actually a couple of great economic reasons that we receive most of the college business. First of all, we stay open until 1am on weeknights and 2am on weekends. There is a huge market segment for late-night food that the fast food businesses compete for, particularly those that deliver. That brings me to the second, and most important, reason why we likely do more business than Papa John's and Pizza Hut combined: dining dollars. Several years ago, the owner of Dominos made a deal with the university to become the only pizza place in town that could accept UA dining dollars. At first this doesn't seem like anything special, but when you consider that every full-time student is required to buy $300 dining dollars each semester, it starts to sound better. Basically there are only so many places you can spend dining dollars, and most of them are in campus dining facilities like the Ferguson Center. So for the average college student who wants food delivered to him late at night, and is out of all money but dining dollars, Dominos is really the only answer. That is what you call completely dominating a market segment.

Examining the coefficient estimates, none of them are surprising in terms of the sign. I must admit, though, I was a little taken aback by the large value of the Race coefficient. I do think there is a very good explanation, however. When I discuss the non-white group, I will mainly be talking about the black samples, since they comprised about 95% of the non-white group, and are much more prevalent in ordering than the other two minorities. First, I think a very large percentage of the non-white deliveries were made to the poorest areas of the city, such as the south side of 10th Ave, and Alberta City. This is much more a socioeconomic issue than a race issue. For the most part, one shouldn't expect a tip on any delivery to those areas, regardless of sex, race, age, or any other factor. These areas, called red zones, are places we don't deliver to after dark because at least one driver has been robbed while delivering there. I think if these areas of poverty were discounted, then the race coefficient would be significantly closer to zero. The best explanation I can come up with for the Age coefficient is that it relates to the Race one. Very few of the red zone deliveries are to college age people. Just like Race, I think there are more powerful factors working against the Age variable than itself. If the red zones were discounted, I actually believe that Age would be a positive coefficient instead of negative. The disproportionate number of red zone deliveries for non-white people over 25 years of age definitely puts the results under question. I think much more in-depth data collection would be needed to correctly estimate these two variables.

There are some interesting ways to continue this study. If I were to collect data over an entire semester, or entire calendar year, I am very confident that there would be evidence of seasonality. Particularly in the summer or over Christmas break when most of the college students leave town, there is much less business on the whole, as well as a larger percentage of red zone deliveries. I would also guess that Dominos does more business during the first two weeks of the semester than at any other point, possibly excluding football home games. Once students burn a hole in their dining dollar balance, however, sales even out somewhat.

One addition I would definitely make if I continued this study would be delivery time. It would be a little more trouble to constantly keep track of that, but it could be done. I'm not sure that it would have an affect on ACT card tips (possibly a lagged affect, but very hard to measure), but delivery time definitely affects cash tips. Location would also be an important addition to an improved model, but I am unsure as to how that could be done. I am afraid this study failed to reveal any groundbreaking economic discoveries, but it was interesting to me personally if nothing else.