The relationship of gender and physical attractiveness for tip*

Thomas Zielinski
September 10, 2014
(thomas.zielinski@online.de)

Abstract

Several experiments (for example Hamermesh and Biddle (1994) or Solnick and Schweitzer (1999)) found a beauty premium in different kind of games or rather different situations of life, like the labour market. Now, this paper analyses if and how physical attractiveness and gender play a role at the tip for waitresses and waiters in restaurants. In a Dictator Game Experiment students and assistants of the University of Passau, had to give virtual gratuity for five different waitresses and waiters at pictures. A small beauty premium was detected. Attractive female and male servers got usually a tip, although their services were really bad. During the further procedure also two gender effects were found: Firstly, women gave essential more gratuity than men, however, not only for waiters, but also for waitresses. Secondly, women usually don`t differ between waitress and waiter, they gave both almost the same tip. But men made in two of five cases a difference: They gave significant more tip to the female than to the male server. Last we researched, if testosterone influences gratuity. But we could not find any relationship for this theory.

*The experiments were conducted in the seminar "Experimental Ecomonics". The author thanks Katharina Werner and Ann-Kathrin Crede for their advice and valuable comments. Special thanks to Manuel Schubert, Marcus Giamattei, Isabelle Rivière, Susanna Grundmann, Maximilian Wolf and Patriz Geiger for helpful comments and support in conducting the experiments.

1. Introduction

For most waitresses and waiters it is an absurdity to talk about his or her gratuity. However, it is interesting to learn something about it. So this paper engages in tip and questions like, "who gets more tip (for the same service)?", "give female- and male guests the same amount of tip?" or "does physical attractiveness act a part?". These and a few other questions will be answered in part **6** of the paper.
But before, in part **2** you get some information about other literature and sources for this work. Then I will introduce the design of the experiment (in **3**) on the one hand and the hypothesis (in **4**) on the other hand. After a few words about the transaction in **5**, I will, as aforementioned, present the results of the experiment (in **6**). Unfortunately the experiment has some limitations, I have to talk about (in **7**), before the paper ends with a conclusion (in **8**).

2. Literature

The following experiment is according to different other experiments with physical attractiveness and gender, the most important will be shown in the next paragraphs.

Tanya S, Rosenblat wrote a paper, that analyzed "Physical Attractiveness and Gender in Dictator Games" (Rosenblat (2008)).
(A Dictator Game (DG) is one of the simplest games. One person, the allocator or dictator, has a fixed stock of money and decides if and how much he wants to share with the second player, the recipient. This decision is binding for both, without any rights for the recipient to reject the offer.) In her experiment Rosenblat found that women, as allocators, always gave more than men. Male subjects always gave almost two out of nine units, no matter how attractive the recipient is. But females that could see and hear the opponent gave "recipients of below-average beauty" only 1.7 out of nine units and "recipients of above-average beauty" 3.3 out of nine units.

Sara J. Solnick and Maurice E. Schweitzer made a research about "The Influence of Physical Attractiveness and Gender on Ultimatum Game Decisions" (Solnick, Schweitzer (1999)).
(An Ultimatum Game (UG), is a simple game, where one person, the proposer, has a fixed stock of money and makes an offer about the distribution of it. A second person, the responder, can accept or reject that suggestion. If he accepts, everybody gets the amount that the proposer suggested, if the responder rejects usually both get nothing.) In their experiment with students of different universities they found neither in offers nor in demands differences between attractive and unattractive people. But they founded a beauty premium for men: attractive men were offered more, and less was demanded. Attractive women were also offered more, but also more was demanded of them.

Another very important paper for my experiment was "Beauty and the Labor Market" by Daniel S. Hamermesh and Jeff E. Biddle (Hamermesh, Biddle (1994)).

On the one hand the authors represented that attractive people earn more than unattractive people. Already while the negotiation (presented by a kind of Bargaining Games)[2] about the wages, attractive people have advantages. Hamermesh and Biddle declared a "plainness penalty" of five to ten percent, mainly men "reach" ten percent difference. On the other hand their experiments showed that unattractive women marry unattractive men with less human capital.

Already these three papers are showing big differences between experiments and results about gender and physical attractiveness. But I had much more sources, that I will present shortly:
Some studies found that men and women differ in their notice of nonverbal signals, like facial expressions in pictures (Hall (1978; 1984); Rosenthal et al. (1984)).
Beauty and gender have been studied in several other papers. Andreoni and Petrie (2008) played public goods games. They found a beauty premium, that only exists during blindness situations and showed that most people consider men as the better leaders. Kahn, Hottes and Davies (1971) or Mulford et al.(1998) made experiments to show the effect of physical attractiveness in Prisoners Dilemma.

3. Design

Basis for the experiment was a DG, I already described above. Theoretically the dictator wants to maximize his payoff. In this game the recipient has no opportunity to reject, so the allocator should keep everything. But because of altruism and fairness most dictators make positive offers (Camerer 2003).
Aim of my experiment were two treatments. Every treatment should consist of five pictures of waitresses or waiters and short different texts for each of them. Always two - one man and one woman - with the same "attractiveness" were a match. That meant one treatment saw the male and the other the female server, each with the same description.
In a first stage, I looked for photos of twelve waiters and twelve waitresses on the internet, printed and cut them out. Then I gave the pictures to 23 subjects at university, thirteen females and ten males. They should find five to seven "pairs", being in a theoretical relationship together.
The probands looked for pictures of waitresses and waiters with quite the same level of sympathy, attractiveness, charisma, facial expression, etc.. I recorded the pairs in a list and went on with the second stage. I gave the seven most frequently mentioned waitress and waiter, in whole 14, to 19 other subjects, eleven men and eight women. That should give an estimation of three categories of beauty of the 14 photos; sympathy, physical attractiveness and facial expression. The estimation took place in a rating from 0 (very bad) to 4 (very good). After this part of the second stage I compared the pairs of stage one with the

[2] Bargaining Games: A Bargaining Game (BG) is played by (at least) two players. The first player has a fixed stock of money and makes an offer about the split of it. If the second player accepts, everybody get the payoff, the first player said. Until this time it is a simple UG. But, if the second player reject the suggestion, the payoff is not zero for both. In a BG it is now the seconds player turn to make an offer to the first player, but with a new, smaller stock. Then the formerly first player can accept or reject. Depending on the number of laps, the game ends now, or the first player can make a suggestion again (if he rejected), of course with a shirking stock.

estimations of stage two. I got five pairs, where both had always quite the same "attractiveness". The smallest average difference between the "attractiveness" of man and woman was 0.34, the biggest 0.71. The most unattractive pair had an average of 1.34, the most attractive pair an average of 3.02.
At that time I had five pairs, so I decided an allocation of three waiters and two waitresses in the first Treatment and two waiters and three waitresses in the second treatment. Now I created five different descriptions to each pair, with specific characteristics: [3]

Pair	Attractiveness (only beauty) Rank	characteristics
1 (Treatment (T) 1 male (m); T2 female (f)	2 of 5 (quiet attractive)	Friendly, Quick, a bit forgetful
2 (T1 f; T2 m)	5 of 5 (most unattractive)	Charming, very quick, Helpful
3 (T1 m; T2 f)	4 of 5 (quiet unattractive)	Quick, Bugged, Incurious
4 (T1 f; T2 m)	1 of 5 (most attractive)	Charming, Slow, Overwhelmed
5 (T1 m; T2 f)	3 of 5 (average attractive)	Nice, Helpful, Obliging, a bit clumsy

Table 1: Attractiveness Rank and characteristics of the different pairs of servers

Basically the experiment was ready for the transaction. But there were some details, that were not considered before:
One discussion was about the amount to be invoiced. Should we take different amounts, or one amount for all? We decided to take one for all, because otherwise we were afraid that different tip consist of different amount invoiced. A guess was, a person who has to pay 20 €, gives automatically less than somebody with a bill of 70 €. The next step was to find a suitable amount to be invoiced. The decision was 41,30 €, because this amount is far enough away from round numbers like forty-five or fifty. In my opinion everybody could round up - or even not - how she or he wanted, without any pressure of rounding.
Another discussion was, if the subjects should only give the tip or the amount invoiced plus gratuity? Finally every subject had in every round hypothetically[4] 20 € only for tip. At the beginning of the experiment I told every subject that they only have to give tip if she or he wants to do, because the bill has already been paid by another person, but without gratuity.
Following the "main experiment", the subjects saw all pictures again. I asked them to do exactly the same like the probands in stage two in order to see if these subjects rate the pictures similar to the others.[5]
Before a small questionnaire about assessments to tip[6] and personality ends the experiment I had one special exercise for the subjects: To trace their left hand, because the digit ratio (2D:4D), means between the index finger and the ring finger, reveals something about the testosterone level of a person. In a nutshell the longer your ring finger compared with your

[3] Compare Appendix 1
[4] Please read about limitations in Chapter 7
[5] Compare Appendix 6 "Screenshots"
[6] Compare Appendix 2

index, the higher is the testosterone level you got at the womb(v. d. Bergh, Dewitte (2006)). Perhaps there is a connection between this and the amount of gratuity.

4. Hypothesis

Earlier papers found that attractive people have some advantages in different situations of life (Eagly, Ashmore, Makhinjani and Longo (1991)). Other studies explained a similar phenomenon: The Beauty Premium, which means that attractive women and men are offered more or can earn more than unattractive people. (Camerer (2003); Doorly, Sierminska (2012)). Consequently, the expectation for this experiment is that pretty waitresses and waiters get more gratuity than less attractive ones.
H1: Gratuities to attractive waitresses and waiters are higher than to less attractive ones when the service is comparable.

Although women are called "the fair sex" compared with men, usually men earn up to 25 % more money; a not fully understood gender gap in wages (Goldin (1990); O`Neill (2003)). But, of course tip is not really to compare with fixed wages and should not depend on them. Nevertheless I try to match wage with gratuity for my next hypothesis, although I know it could sound strange. I think waiters get more tip than waitresses when their service is comparable.
H2: Waiters get more tip than waitresses for a similar performance.

In their book "Women wants more" the Boston Consulting Group (BCG) potters at the phenomenon that women spend more money than men. Globally the average of consume that women shop is around 70 percent (Silverstein, Sayre (2009)). Although the authors do not describe how these 70 percent divide in things like food, the household, other things for the family on the one hand and (luxury-)purchases for the woman herself on the other hand, you could imagine females give also more tip, because they expend more money by nature.
H3: Women give more tip than men.

The next two hypothesis are very similar. For one thing we will research if women give (attractive) waiters more tip than men do. For another thing we will turn around the scenario and detect if men present (attractive) waitresses more gratuity than women do. In other experiments, authors found that women have a bigger spread in such things. While gender and beauty had only small effects on men, women gave more when they dealt with a (nice) man, preferably with a pleasant voice. Males usually think more rationally and wanted to maximize their own pay off (Camerer (2003); Rosenblat (2008)). That implies for my hypothesises that women will give much more tip to waiters and even (a little) more to waitresses than men.
H4: Women give more tip to waiters than men do.
H5: Men give less tip to waitresses than women do.

At last, we will research if there is any connection between tip and the testosterone level that you can work out by comparing the digit ratio (2D:4D), of the index finger (2D) and the ring finger (4D). A very simple explanation: The longer the ring finger compared with the index finger, the higher someone's testosterone level. That usually implies a more dominant and aggressive behaviour. Usually men have a longer ring finger than index finger, while women

have a bigger index finger than ring finger. So for men counts the definition I gave a few lines ahead, and for women counts, the smaller the difference between both fingers, the more dominance and aggressiveness.[7] In an experiment was detected that, the smaller the digit ratio the higher was the rejection of unfair offers in an Ultimatum Game. (v.d. Bergh (2006)). Therefore you could imagine for my experiment, that students with a higher testosterone level, are more aggressive and so they gave less tip.
H6: The higher the testosterone level of somebody is, the less tip he gives.

5. Transaction

My experiment was conducted together with the one of Susanna Grundmann at the University of Passau on June 16th and 17th 2014. Both were played consecutively, so that every subject had to play two times. We changed the succession of our experiments after every session to avoid order effects. Our experiments were played via computer using the Zurich Tool box for Readymade Economic Experiments (Ztree), a freeware that is usually used at the University of Passau. After closing the door, we read a small text, to welcome the subjects, to ask them not to talk or to play with their mobile phones and to act as if they would spend their own money. Then the participants were allowed to start the experiments. My first two pages thanked for coming and explained the "game". For the next five stages every page showed a picture of a waitress or a waiter and a small text describing the server and the whole situation. After considering the photo and reading the description, the students had to decide how much tip they want to give and register the amount (between 0.00 € and 20.00 €) in a small window below the text. After the fifth one, they saw their decisions again and how much money they still would have. The next stage was for explanation again, to declare they will see all servers again and should evaluate them in three categories: sympathy, beauty, and facial expression, each from 0 (very bad) to 4 (very good). After five stages where the subjects had to perform this exercise, three stages appeared with three questions about individual positions in gratuity and a fourth stage that asked the probands for transferring their hand to a paper. At the next stage, the students had to leave their gender, age and what they study. Then you reached the last stage that just thanked again for participation and plead for silence until everybody had finished.

[7] compare <http://www.welt.de/wissenschaft/article13613930/Was-die-Laenge-von-Zeige-und-Ringfinger-verraet.html (Date: 17.09.2014 at 17:36)>

6. Results

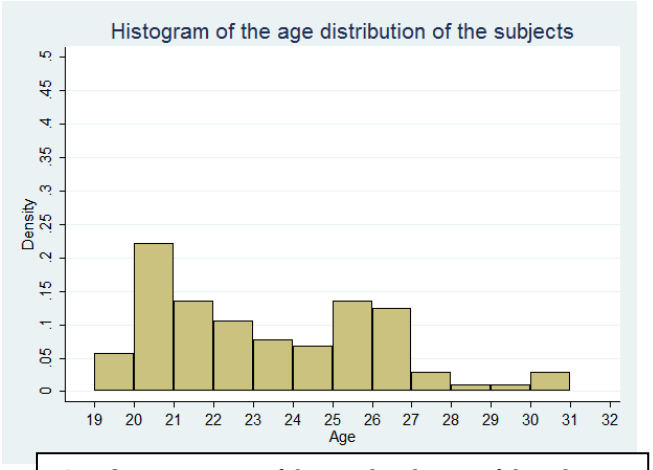
6.1. Key facts

Altogether 104 people took part in the experiment. Most of them were students, the rest assistances of the University of Passau. 70 of them were women and 34 men.

The histogram shows the age distribution of the participants. The youngest subject was 19, the oldest 31 and the overall average was 22.88 years. You can also see, almost one quarter, exactly 23 attendees, were 20 years old.

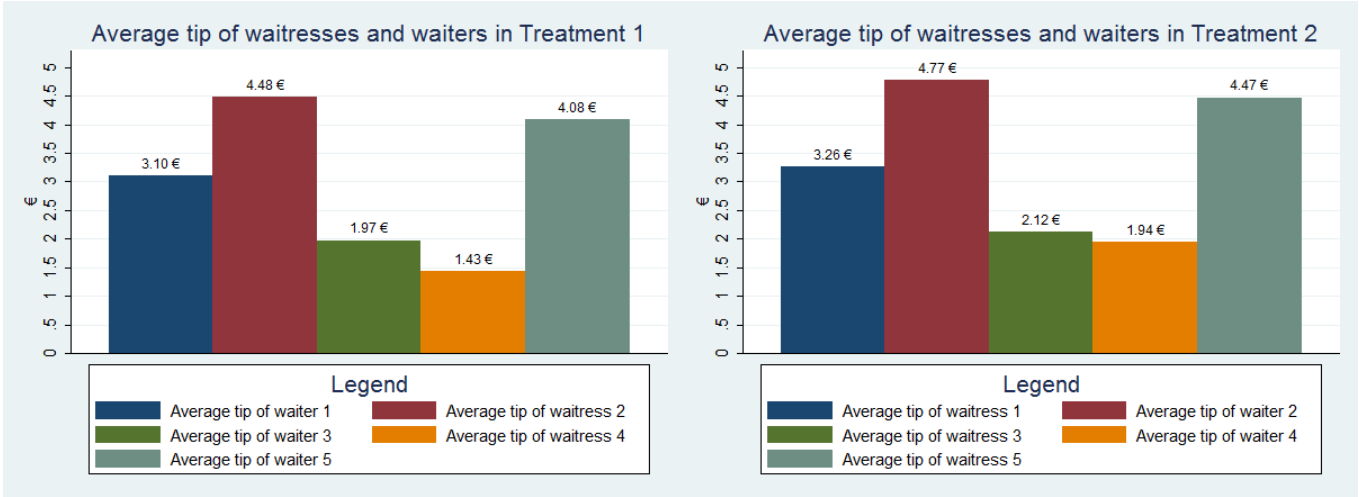
Almost the half, 51, study "Betriebswirtschaft" or "Volkswirtschaft", while the second biggest group "Kulturwirtschaft" consisted only of 18 students.

The first treatment was performed by 53 subjects, 38 of them were women and 15 men. The second one was played by 51 subjects, 32 women and 19 men.



Graph 1: Histogram of the age distribution of the subjects

The second graph shows a comparison between the average tip of both treatments.



Graph 2 shows a compare between the average tips of both treatments

Without any bigger calculation you can notice the following facts:

The overall highest gratuity got waiter two of Treatment 2 with an average of 4.77 €, the smallest tip, waitress four of Treatment 1 with an average of 1.43. That result is interesting because pair two was an old unattractive looking pair but with a good service, while pair four were the most beautiful pair, but they were overwhelmed. We already can assume prettiness is not everything you need for a high tip.

Altogether on average each of the waiters and waitresses got a tip of 3.16 €, whereof the waitresses earned 3.15 € and the waiters got 3.17 €. All servers in Treatment 1 had an average

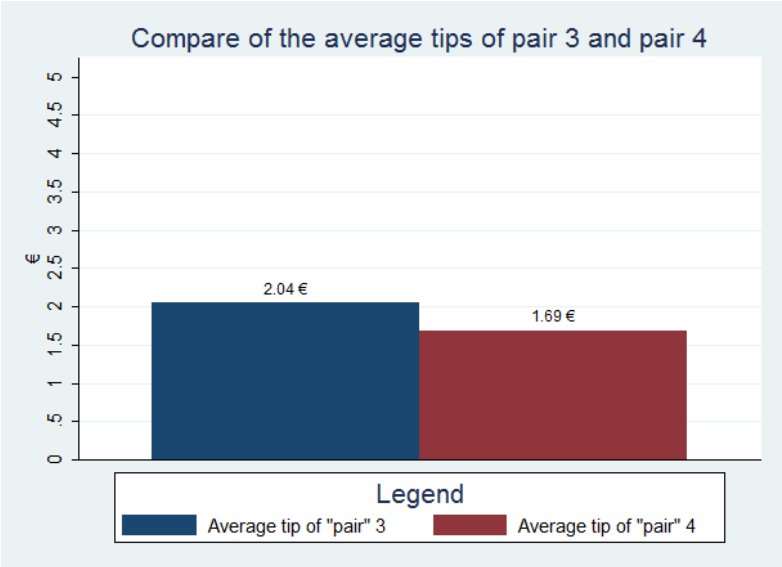
of 3.01 €, in Treatment 2 they had 3.31 €. You just can read off the graphs that every single server, female as male, earned more gratuity in Treatment 2, but there are no indications why.

6.2. Specific Results

Now I will present the results of the experiment and compare them with the hypothesis we postulated above.

a) *H1: Gratuities to attractive waitresses and waiters are higher than to less attractive ones when the service is comparable.*

For answering this question we have to consider the tip of the beautiful pair four and pair three, that fourth out of five at the "Beauty-Ranking". When I designed the experiment, I decided to compare these pairs for H1, because their service was very similar. Pair three was quick but bugged and incurious and pair four was charming, but slow and overwhelmed. So both had one good feature but also two - for servers - very bad features. Take a look at graph 3. It shows the average tip of waitress and waiter 3 (blue) compared with the one of pair 4 (red). While H1, derived from other literature said, pair 4 would have an higher tip, the experiment showed the opposite. While waitress and waiter 3 got 2.04 €, waitress and waiter 4 only got 1.69 €, a difference of 0.35 €. So H1 is obviously wrong. An explanation could be, that sympathy and facial expression of the photos acted also a part. In those categories pair 4 was 5 out of 5, with a big gap to the fourth, in both cases that were pair three.



Graph 3: Tips of pair 3 and 4

To examine if this difference between both is significant, although it is an unexpected one, we took the Wilcoxon Signed-Rank test. Z is 0.0003, that means there is a very significant difference between the tip of pair 3 and the gratuity of pair 4. So we can notice that beauty did not matter in our experiment.

b) *H2: Waiters get more tip than waitresses for a similar performance.*

Above I already revealed that on average waiters got 3.17 € and waitresses got 3.15 €. So it seems that there is not really a difference between the genders. But there is one fact you should not forget. As mentioned before in every case the subjects gave more in Treatment 2 on average. But in Treatment 2 we had three waitresses and two waiters, that means usually the female servers should have an higher average, shouldn`t they?

Table 2 shows the tips of each waitress and waiter again:

Number	Average of waiter	Average of waitress	Difference	z-Data of Mann-Whitney
1	3.10 €	3.26 €	0.16 €	0.9896
2	4.77 €	4.48 €	0.29 €	0.5027
3	1.97 €	2.12€	0.15 €	0.9765
4	1.94 €	1.43 €	0.51 €	0.0573
5	4.08 €	4.47 €	0.39 €	0.2856

Table 2: Differences in gratuity between the pairs and z-Data of Mann-Whitney-Test

Take a look at the differences between the gratuities of waitresses and waiters in detail. You should realise that the gap of only 0.02 € between the genders is not because always both got almost the same tip, but it is more or less a fluke. The z-Data of the Mann-Whitney-Test also can be seen at Table 2.[8] It shows in four of five cases no gender effect. Only when you look at the "Beauty-Pair" number 4. you can guess a small effect. But also this one is not that significant. All in all there are more reasons to reject H1 than not to reject.

c) *H3: Women give more tip than men.*

For this "question" a graph bar is suited again, because you can see some details without calculations again. You can find them in the appendix 5. The graph shows the average tip of each pair, divided into if the subject was female or male. Instantly you should realise that female participants gave more gratuity in all five cases, at this time indifferent if they gave it to a waiter or a waitress. The smallest gap, with 0.35 €, between women and men we observe at pair 3. They got 3.69 € by men and 4.04 € by women. You find the biggest difference at pair 3, with 0.60 €: Men gave them 1.64 € on average, while women donated them 2.24 €.

Now we examine if these numbers are significant. For testing this hypothesis you have to take Mann-Whitney-Test again.
The table right illustrates the situation for pair 3. You can see that 34 men and 70 women were counted. Now is important the last line. If z is smaller than 0.05 you can extinguish, that women give more tip than men to the observed server.

For this case z = 0.0527. So z is a little bit bigger than 0.05. In this case we talk about a small significant difference, we have detected.

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

q1	obs	rank sum	expected
1	34	1507	1785
2	70	3953	3675
combined	104	5460	5460

unadjusted variance 20825.00
adjustment for ties -232.62
---
adjusted variance 20592.38

Ho: amount^3(q1==1) = amount^3(q1==2)
z = -1.937
Prob > z = 0.0527

Table 3: Mann-Whitney-Test for pair 3 according to the question "Give women more tip than men?"

[8] Compare Appendix 4

Comparing the other tests, you will determine that there is no table, where p is smaller as 0.05. That would mean we have to reject our hypothesis again. But while considering more precisely you see that pair 1 and 2 have p-values, that are obviously under 0.20. Pair 1 has a value of 0.0976 and pair 2 has an value of 0.1082.
The value of the other pairs is 0.2803 for waitress and waiter 4, still not too bad and 0.5992 for server 5.
So for $p<0.05$ we have to reject our hypothesis, because only no one observation has a $p<0.05$. But for $p<0.20$ three observations are easily in this window, almost three would be under $p<0.10$. The fourth and fifth values, showed no significance at all. Although, a small significance was detected.

d) + e) *H4: Women give more tip to waiters, than men do.*
H5: Men give less tip to waitresses, than women do.

These hypothesis are very similar, so it would make sense to consider them together. The graph bars for H4[9] shows that women gave waiters in four out of five cases much more gratuity, about one Euro more at average. Only waiter two got by men an incredible tip of 5.21 €[10], while women "only" gave them 4.03 €. Notice 4.03 € was the highest average, women gave over all. To sum up, H4 could be true, we will use the Mann-Whitney-Test again, after we have took a look at the graph bars for H5: First of all we can observe that women gave waitresses in every case more gratuity than men did. The spread of the differences moves from 0.05 € (for waitress 4) up to 1.23 € for female server 2. An interesting fact, you should realise is, that waitress and waiter 2 got each more tip by their own gender, than of the cross one. Remember, by accident those pair was the most unattractive one, but good at their job and very pleasant.

For H4, we only have to check four waiters. We can ignore waiter 2, because we already saw that men gave more gratuity to him at average than women, so it is not possible to find a significance that women would donate him more tip than man. For the other waiters you have a table below.
The Mann-Whitney test shows for waiter 1 (z = 0.0385) and 3 (z = 0.0035) a significant difference between the genders, so women gave them important more gratuity. at average. Waiter 4 (z = 0.1068) still has a small significance for more tip by females. Only waiter 5 (z = 0.3665), and of course waiter 2[11], do not support H4. Nevertheless you can assume of the trueness of H4.

Number of waiter	z-Value
1	0.0385
3	0.0035
4	0.1068
5	0.3665

Table 4: z-Value for Mann-Whitney-test for waiters

[9] compare Appendix 5
[10] First I thought of a mistake, but I calculated it 4 times and always these amount appeared.
[11] Waiter two has a z-Value of 0.0836. So he even has a small significance for getting more by men

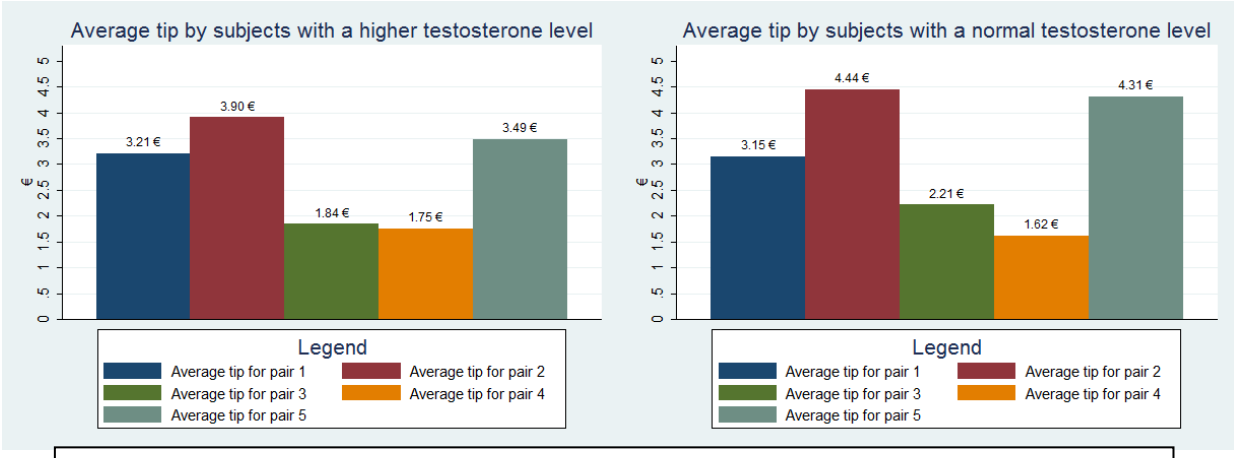
For H5 we will check all waitresses, because the graph bar showed us, women gave waitress more in every case.
In this case we only have a small significance for waitress 2 (z=0.0552). For all the other ones, we cannot find any significant statement for our hypothesis that women would give waitresses significant more tip than men. So we have to reject H5.

Number of waitress	z-Value
1	0.6310
2	0.0552
3	0.9215
4	0.6883
5	0.3673

Table 5: z-Value of Mann-Whitney-test for waitresses

e) *H6: The higher the testosterone level of somebody is, the less tip he gives.*

For this hypothesis, I had to evaluate the hands the subjects drew. Unfortunately, it was not that easy, because some hands were not drawn exactly, other hands seemed to have long fingernails. But I tried my best and found 49 subjects with a higher testosterone level, 22 male and 26 female subjects. The graph bars beyond exposes the results:



Graph 4: Compare of average tip by subjects with higher testosterone level with a normal testosterone level

Left you see the average tip the 49 subjects with a higher testosterone level gave an right the average of the gratuity the female and male servers got by participants with a normal testosterone level. Pair 1 and 4 got a little bit more by the probands with a higher level, the other couples got more by the subjects with a usual testosterone level. So you cannot demonstrate that the testosterone level have had an influence while this experiment, that means H6 can be rejected.

After these presenting these results, there are two more things I shortly want to expose: I worked out two regressions, you find them at Appendix 6. The first regression tries to find a Connection between all shared amounts of T1, gender of the subject, age of the subject and gender of the server. The second one tries the same, but with T2 instead of T1: We cannot find any useful at the second regression, but at the first one the gender of the subject has a p= 0.031. That means: when you consider all mentioned parameters together, the gender of the subject was significantly important for gratuity.

7. Limitations

While the preparation and the transaction you saw our experiments had a few limitations, that could be reasons for other results than we expected. Following I will present the most important limitations of this experiment.

First of all we did not have a real pay out. Neither the subjects got the money, that they did not donate the servers, nor the waitresses and waiters got anything. Whereby we reached another limitation - nobody played (for) the waiters, those were just pictures from the internet. To avoid these problems in a laboratory experiment, first of all you would create a real payout. That need not to be the rest of the 20 € each round, that also could be the rest of the 20 Euros of one random round or a fixed exchange rate for example. The tip for the servers is more difficult, because there are many options. But for every option it is important to tell the subjects what will happen in advance. For example you could take the whole gratuities and donate it for a hospital or anything else. Another idea is to take pictures of real waiters and waitresses in a foreign city and really give them "their" tip after the experiment. But in my opinion the most useful and less problematic idea is to have a couple of subjects - one plays the guest and one plays the waitress or waiter. That it is not too boring for the second player he could guess how much tip he would get in each situation and earn extra money for a right prediction.

A next limitation is that the pictures of the pairs were not comparable. A man and a woman usually cannot have exact the same look, pleasant and facial expression. That is a limitation you always will have at such experiments, but if I had had more pictures and more people that matched the couples, perhaps it would have been possible to have quite perfect couples.

Finally, we almost only played with students that usually do not have much money. In a lavatory experiment you often have more different people, so you can make better statements for an average of all people and not only for students, what we actually did.

8. Conclusion

This paper pottered at tip for waitresses and waiters, while exploring six hypothesis about this theme.

First we noticed that tip does not really depend on beauty and service. There are more things that count: for example sympathy and facial expression. Secondly we analysed if any gender gets a higher tip than the other gender, but we could not demonstrate any significant difference. Then we attended to the question if women give more gratuity than men and realised some smaller gaps. Women really often gave more tip than men. After this discovery we found on average women gave in three out of five cases significant more gratuity to waiters than men did, but men never gave more tip to waitresses than women did. Last we tried to find an influence of testosterone. We wanted to show that people with higher testosterone level give less tip. But we could not demonstrate any relation.

Altogether we found a few things we have guessed before, but it would be interesting if more hypothesis would have come true with more data of different groups of people and a real payout. In my opinion the problem with the only fictive one must not be forgotten. So it would be worth making this experiment again - without the limitations.

9. Literature

- Andreoni, J., & Petrie, R. (2008). Beauty, gender and stereotypes: Evidence from laboratory experiments. Journal of Economic Psychology 29: 73–93.
- v. d. Bergh, B., & Dewitte, S. (2006). Digit ratio (2D : 4D) moderates the impact of sexual cues on men's decisions in ultimatum games. Faculty of Economics and Applied Economics, Leuven (Belgium).
- Camerer, C. (2003). Behavioral game theory. Princeton, NJ: Princeton University Press.
- Davis, W. L. (1971). Cooperation and optimal responding in the prisoner's dilemma game: Effects of sex and physical attractiveness. Journal of Personality and Social Psychology 17: 267–279.
- Doorley C., & Sierminska, E. (2012): Myth or Fact? The Beauty Premium across the Wage Distribution. Discussion Paper 6674.
- Eagly, A. H., Ashmore, R. D., Makhijani, M. G., & ,L. C.. (1991). What is beautiful is good, but: A meta-analytic review of research on the physical attractiveness stereotype. Psychological Bulletin 110(1): 109–128.
- Goldin, C., & Rouse, C.. (2000). Orchestrating impartiality: The impact of "blind" auditions on female musicians. The American Economic Review 90(4): 715–741
- Hall, J. A. (1984).Nonverbal sex differences: Communication accuracy and expressive style. Baltimore, MD: The Johns Hopkins University Press.
- Hamermesh, D., & Biddle, J. (1994). Beauty and the labor market. American Economic Review, 84 (5), 1174–1194.
- Hübener, I. (2011). Was die Länge von Zeige und Ringfinger verrät. <http://www.welt.de/wissenschaft/article13613930/Was-die-Laenge-von-Zeige-und-Ringfinger-verraet.html> (Letzter Zugriff 24.09.2014)
- Kahn, A., Hottes, J., & Hall, J. A. (1978). Gender effects in decoding nonverbal cues. Psychological Bulletin 85(4): 845–857.
- Mulford, M., Orbell, J., Shatto, C., & Stockard, J. (1998). Physical attractiveness, opportunity, and success in everyday exchange. American Journal of Sociology, 103 (6), 1565–1592.
- O'Neill, J. (2003). The gender gap in wages, circa 2000. American Economic Review, 93 (2), 309–314.
- Rosenblat, T. (2008). The Beauty Premium: Physical Attractiveness and Gender in Dictator Games. Negotiation Journal, 465-480.
- Rosenthal, R., Hall, J. A., DiMatteo, M. R., Roger, P., & Archer, D.. (1984). Sensitivity to nonverbal communication: The PONS test. Baltimore, MD: The Johns Hopkins University Press.
- Scharleman, J., Eckel, C., Kacelnik, A., & Wilson, R. (2001). The value of a smile: Game theory with a human face. Journal of Economic Psychology, 22 (5), 617–640.
- Silverstein, M.J., & Sayre, K. (2009). Women want more. New York, NY: Boston Consulting Group.
- Solnick, S., & Schweitzer, M. (1999). The influence of physical appearance and gender on ultimatum game decisions. Organizational Behavior and Human Decision Processes, 79 (3), 199–215.

9. Appendix

1. Text and pictures of all waitresses and waiters

Waiter 1:

Waitress 1:

"Sie wurden von dem Kellner[der Kellnerin] links auf dem Foto bedient. Er [Sie] servierte Ihnen ein 3-Gänge Menü. Leider hat er [sie] dabei übersehen, Ihren Umbestellungswunsch beim Hauptgang an die Küche weiterzugeben, sodass Sie eine Beilage auf dem Teller hatten, die Sie nicht unbedingt essen wollten. Ansonsten ist er [sie] aber sehr um Ihr Wohl bemüht und freundlich. Entscheiden Sie bitte nun, ob und wie viel Trinkgeld Sie von den 20 Euro in ihrem Geldbeutel dem Kellner [der Kellnerin] zukommen lassen wollen."

Waiter 2:

Waitress 2:

"Sie wurden von dem Kellner [der Kellnerin] links auf dem Foto bedient. Kaum haben Sie das Lokal betreten, hat Sie der Kellner [die Kellnerin] mit einem Lächeln empfangen und Ihnen den letzten freien Platz im Restaurant zugewiesen. Trotz des Stresses, den der Ober [die Kellnerin] hat, nimmt er [sie] sich die Zeit Sie nach bestem Gewissen bei der Essensauswahl zu beraten. Schnell und geschickt serviert er [sie] Ihnen Getränke und Vorspeise. Unglücklicherweise dauert die Hauptspeise etwas länger, da das Lokal wirklich voll ist, wofür sich Ihre Bedienung auch aufrichtig entschuldigt.

Entscheiden Sie bitte nun, ob und wie viel Trinkgeld Sie von den 20 Euro in ihrem Geldbeutel dem Kellner [der Kellnerin] zukommen lassen wollen."

Waiter 3: *Waitress 3:*

"Sie wurden von dem Kellner [der Kellnerin] links auf dem Foto bedient. Schnell, aber etwas unfreundlich begrüßt Sie der Kellner [die Kellnerin] und reicht Ihnen die Karte. Nachdem er [sie] Ihnen zügig die Getränke gebracht hat, will er [sie] von Ihnen das Essen aufnehmen. Dabei belehrt Sie die Bedienung, dass die erhofften Umbestellungen nicht möglich seien, da dies in diesem Restaurant nicht üblich wäre. Trotzdem finden Sie sich ein 4-Gänge Menü, auch wenn Sie dabei gern die eine oder andere Beilage ausgetauscht hätten. Der Kellner [Die Kellnerin] bringt ihre vier Gänge in angemessener Zeit, wobei er [sie] Sie aber nicht wirklich zu beachten scheint.

Entscheiden Sie bitte nun, ob und wie viel Trinkgeld Sie von den 20 Euro in ihrem Geldbeutel dem Kellner [der Kellnerin] zukommen lassen wollen."

Waiter 4:

Waitress 4:

"Sie wurden von dem Kellner [der Kellnerin] links auf dem Foto bedient. Mit einem charmanten Lächeln bringt er [sie] Ihnen die Speiskarte mit dem Hinweis, dass das erste Hauptgericht auf der Tageskarte leider schon aus sei. Bis er [sie]dann wieder erscheint, um nach Ihren Wünschen zu fragen, vergehen fast 20 Minuten. Sie bestellen ein 2-Gänge Menü mit Haupt- und Nachspeise und weisen darauf hin, eine der Beilagen bei der Nachspeise bitte wegzulassen, da Sie dagegen eine Allergie haben. Als der Kellner [die Kellnerin] Ihre Getränke serviert, merken Sie, dass der Ober [die Kellnerin] Ihnen ein falsches, ähnlich aussehendes Getränk gebracht hat. Mit großen Entschuldigungen nimmt die Bedienung das falsche Getränk wieder mit, vergisst aber bis zur Hauptspeise Ihnen das richtige zu bringen. Wieder entschuldigt er [sie] sich höflich und bringt nun endlich das richtige Getränk. Als dann die Nachspeise vor Ihnen steht, befindet sich auf dem Teller die von Ihnen abbestellte Beilage, da der Kellner [die Kellnerin] der Küche nicht Bescheid gab. Erst beim Abräumen fällt ihm das Missgeschick auf. Mit hochrotem Kopf entschuldigt er [sie] sich abermals.

Entscheiden Sie bitte nun, ob und wie viel Trinkgeld Sie von den 20 Euro in ihrem Geldbeutel dem Kellner [der Kellnerin] zukommen lassen wollen."

Waiter 5: *Waitress 5:*

"Sie wurden von dem Kellner [der Kellnerin] links auf dem Foto bedient. Etwas in Eile bringt er [sie] Ihnen die Karte mit einem lockeren Spruch auf den Lippen. Der Kellner [die Kellnerin] verspricht Ihnen, dass er [sie] bei Ihrem 3-Gänge Menü statt der eigentlichen Hauptspeise ein völlig anderes Hauptessen servieren kann. Zu ihrer Freude hält er [sie] sein [ihr] Versprechen und Sie bekommen die umbestellte Hauptspeise. Bei der Nachspeise stolpert der Ober [die Kellnerin] kurz vor ihrem Tisch unverschuldet, sodass er [sie] Ihnen ein wenig von der Nachspeise über die Schuhe kippt. Zusätzlich zu einer großen Entschuldigung und einer neuen Nachspeise bekommen Sie als Wiedergutmachung ein Getränk Ihrer Wahl aufs Haus.

Entscheiden Sie bitte nun, ob und wie viel Trinkgeld Sie von den 20 Euro in ihrem Geldbeutel dem Kellner [der Kellnerin] zukommen lassen wollen."

2. Individual positions in gratuity

Question	Answers
"Berechnen Sie Trinkgeld, welches Sie der Bedienung zukommen lassen wollen eher in Prozent oder in absoluten Zahlen?" (n = 104)	in Prozent: 42 absolut: 62
Wenn Prozent: "Wie viel Prozent vom Rechnungsbetrag halten Sie in Deutschland für angemessen?" (n = 42)	5 %: 1x 7 % 1x 8 % 5x 10 % 34x 15 % 1x Ø 9,69%
Wenn absolut: "Wie viel Euro Trinkgeld geben sie durchschnittlich absolut?" (n = 62)	€ < 1: 2x 1 <= € < 2: 22x 2 <= € < 3: 16x 3 <= € < 4: 12x 4 <= € < 5: 6x 5 < =€ < 6: 3x 6 <= €: 1x Ø 2.35 €
"Wie würden Sie sich auf einer Skala von Geizig [0] bis Großzügig [4] einstufen wenn es um Trinkgeld geht?" (n= 104)	Geizig (0) 4x Eher geizig (1) 22x Neutral (2) 54x Eher großzügig (3) 24x Großzügig (4) 0x Ø 1,94 **fast neutral**

3. Table for the Wilcoxon Signed-Rank test of H1

Wilcoxon signed-rank test

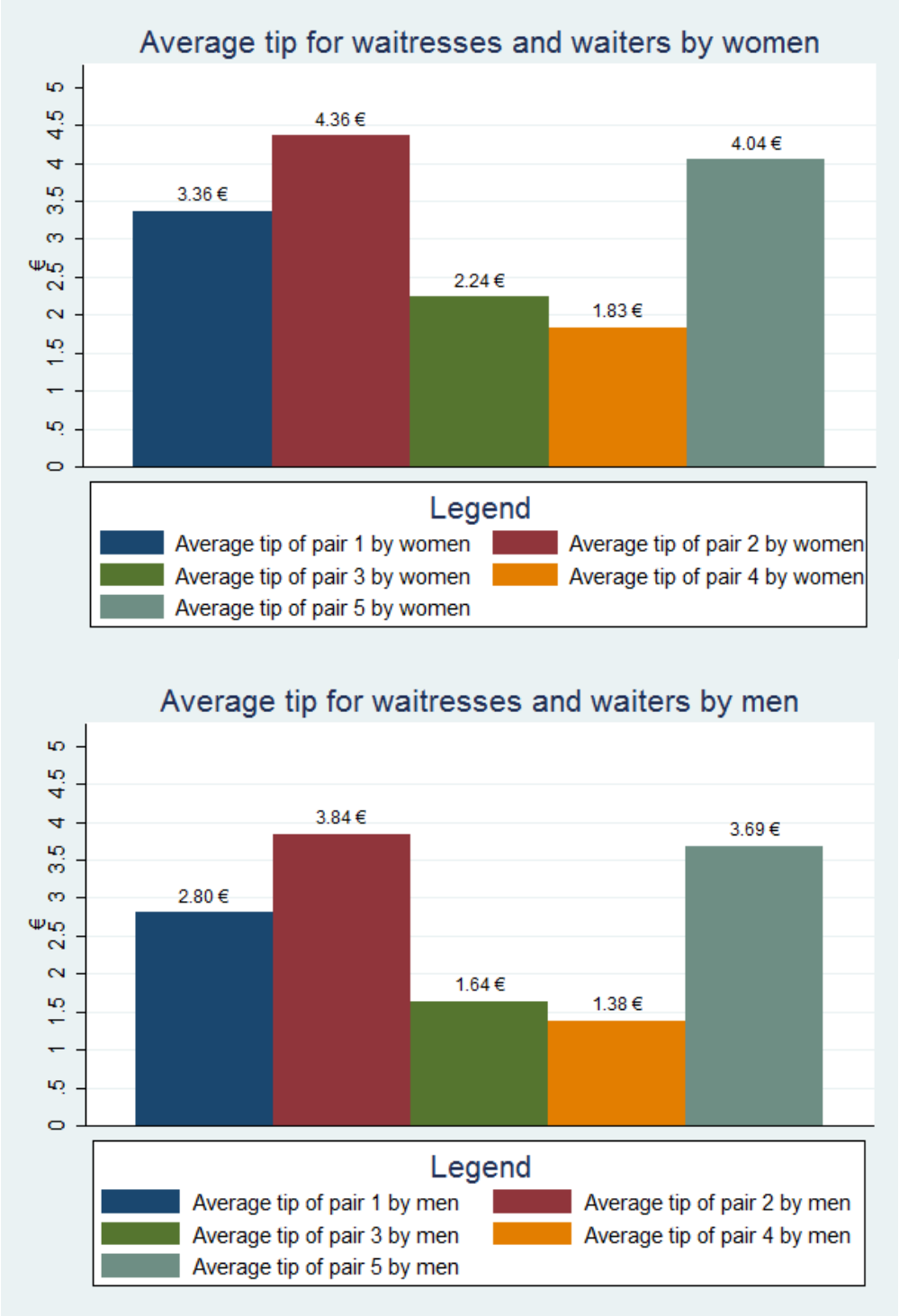
sign	obs	sum ranks	expected
positive	52	3549	2449.5
negative	19	1350	2449.5
zero	33	561	561
all	104	5460	5460

unadjusted variance 95095.00
adjustment for ties -53.00
adjustment for zeros -3132.25
---
adjusted variance 91909.75

Ho: amount_p3 = amount_p4
z = 3.627
Prob > z = 0.0003

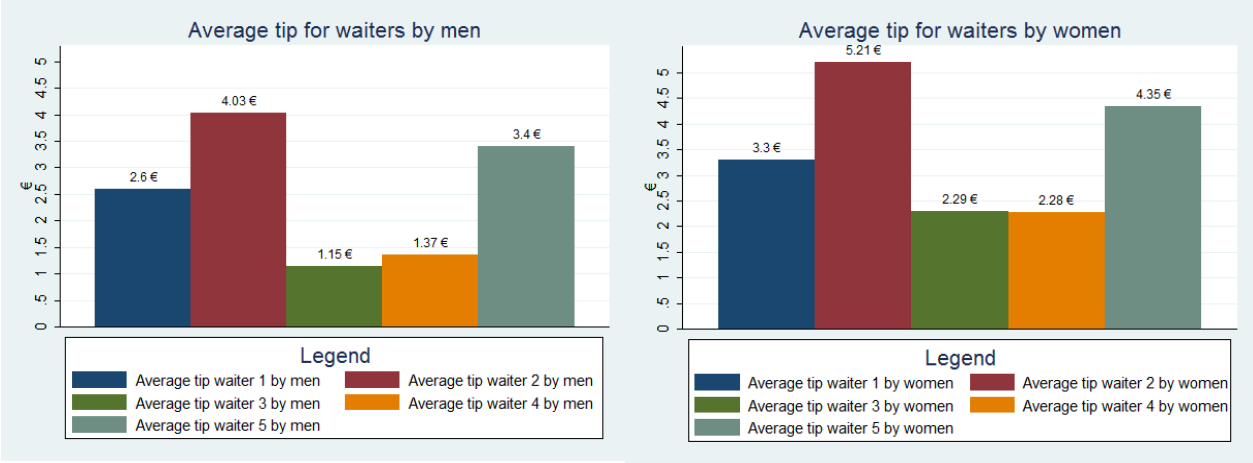
Table 6: Wilcoxon signed-rank test for the question if beauty takes a part for tip

4. Graph Bars according to H3

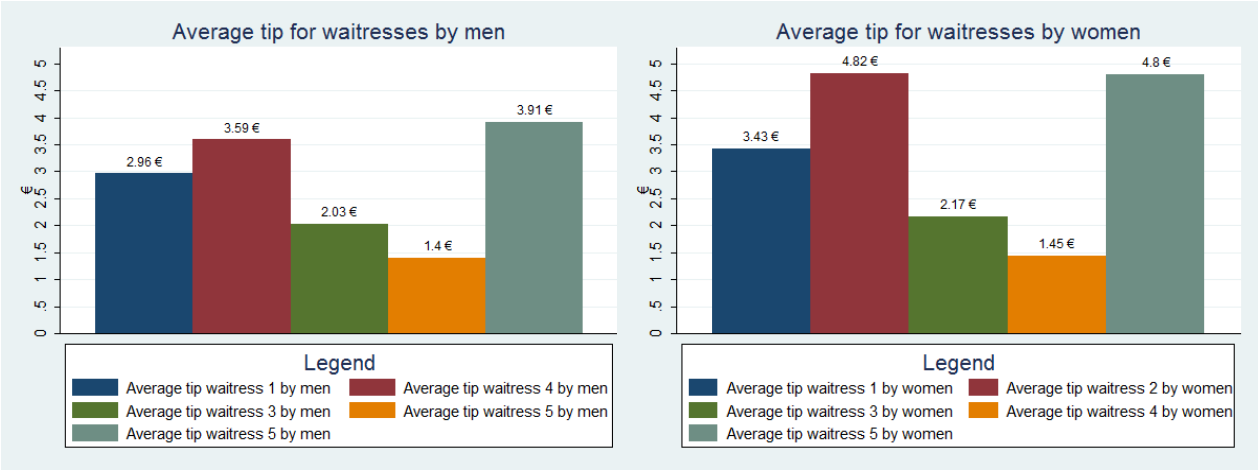


Graph 5: Graph bars to compare the different average of tips by gender according to H3

5. Graph bars according to H4 & H5



Graph 6: Graph bars to compare the different average of tips by gender of guest & gender of server according to H4 & H5.



6. Two Regressions

T1:

regress amount_shared_all_t1 subject treatment gender age gender_bedienung, cluster (subject)

```
Linear regression                               Number of obs =     265
                                                F(  4,    52) =    2.03
                                                Prob > F      =  0.1033
                                                R-squared     =  0.0426
                                                Root MSE      =  2.0437

                                  (Std. Err. adjusted for 53 clusters in subject)
------------------------------------------------------------------------------
             |               Robust
amount_sha~1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     subject |   -.002813   .0057882    -0.49   0.629    -.0144279    .008802
   treatment |  (omitted)
      gender |   .6601634   .2981011     2.21   0.031     .0619798   1.258347
         age |  -.0676689   .0615533    -1.10   0.277    -.1911846   .0558468
gender_bed~g |  -.0964465    .114655    -0.84   0.404    -.3265186   .1336255
       _cons |   3.711176   1.591709     2.33   0.024     .5171775   6.905175
------------------------------------------------------------------------------
```

T2:

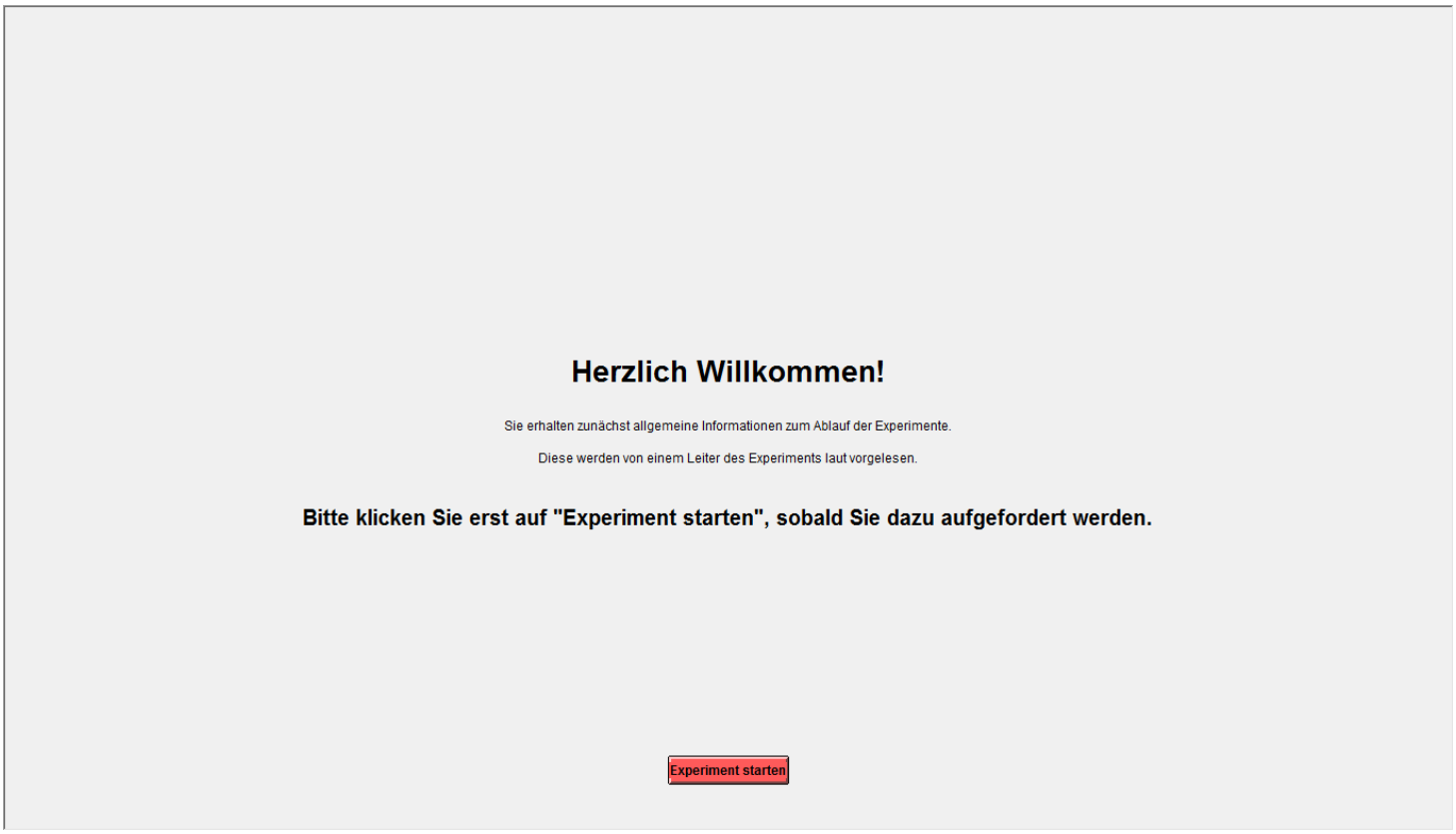
regress amountshared_all_t2 subject treatment gender age gender_bedienung, cluster (subject)

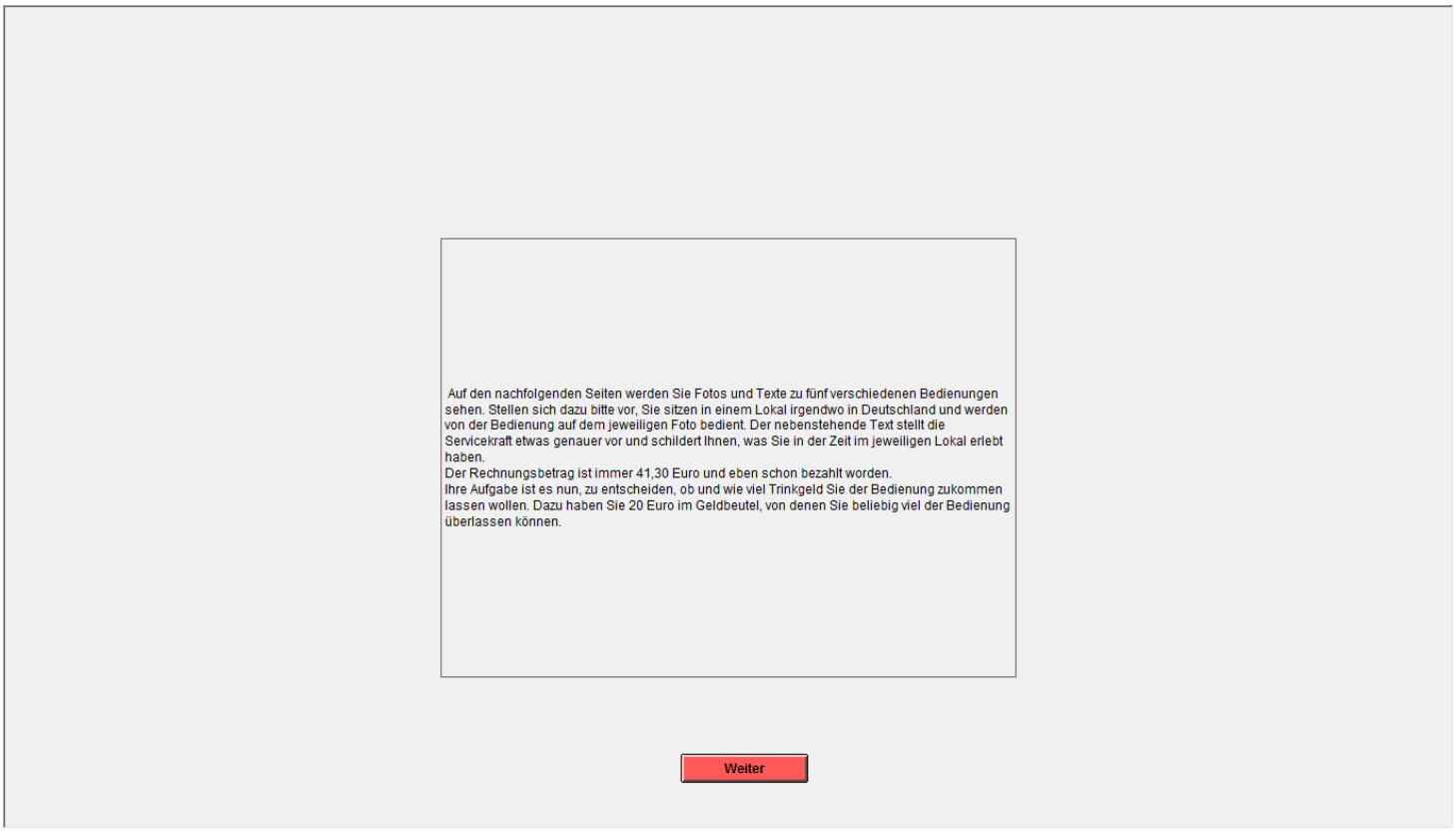
```
Linear regression                               Number of obs =     255
                                                F(  4,    50) =    1.86
                                                Prob > F      =  0.1329
                                                R-squared     =  0.0756
                                                Root MSE      =  2.4511

                                  (Std. Err. adjusted for 51 clusters in subject)
------------------------------------------------------------------------------
             |               Robust
amountshar~2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     subject |  -.0161708   .0074904    -2.16   0.036    -.0312157   -.0011258
   treatment |  (omitted)
      gender |   .6176979   .4183741     1.48   0.146    -.2226312   1.458027
         age |  -.1329879   .0990545    -1.34   0.185    -.3319447   .0659688
gender_bed~g |  -.0742157   .1281428    -0.58   0.565    -.331598    .1831666
       _cons |   6.320811   2.471191     2.56   0.014     1.357278   11.28434
------------------------------------------------------------------------------
```

Table 7: Regressions between all shared amounts of T1 [T2], gender of the subject, age of the subject and gender of the server

7. Screenshots





Sie wurden von der Kellnerin links auf dem Foto bedient. Sie servierte Ihnen ein 3-Gänge Menü. Leider hat sie dabei übersehen, Ihren Umbestellungswunsch beim Hauptgang an die Küche weiterzugeben, sodass Sie eine Beilage auf dem Teller hatten, die Sie nicht unbedingt essen wollten. Ansonsten ist sie aber sehr um Ihr Wohl bemüht und freundlich. Entscheiden Sie bitte nun, ob und wie viel Trinkgeld Sie von den 20 Euro in ihrem Geldbeutel der Kellnerin zukommen lassen wollen.

Trinkgeld Runde 1 0.00

0.00

Rechnungsbetrag: 41,30 Euro

OK

Hier können Sie sehen, mit wie viel Geld in der Tasche Sie nach jeder Runde nach Hause gegangen wären:

Runde 1: 18.80
Runde 2: 18.20
Runde 3: 19.30
Runde 4: 19.90
Runde 5: 17.30

Weiter

äußerst unsympatisch ○ ○ ○ ○ ○ äußerst sympatisch

äußerst unattraktiv ○ ○ ○ ○ ○ äußerst attraktiv

negative Mimik ○ ○ ○ ○ ○ positive Mimik

Weiter

Berechnen Sie Trinkgeld, welches Sie der Bedienung zukommen lassen wollen eher in Prozent oder in absoluten Zahlen? ○ Prozent
○ absolute Zahl

Weiter