Project 3 - Final Project

Self-Checkout vs Cashier Checkout at Meijer

Nicole Pililyan

Systems Simulations

CS4830/CS6830

December 6, 2022

Dr. Mateen Rizki

The Question

Is it faster to go through self-checkout or to go through the trained-employee's register when checking out at the Meijer by Wright State University (Meijer, 3822 Colonel Glenn Hwy, Fairborn, OH 45324)?

Overall, customers feel like self-checkout is faster, but is it really? Reading online, one said it feels faster because customers are doing the work themselves instead of waiting on others [1]. This also leaves more customers feeling satisfied and leaving with a smile on their face rather than a frown, as they weren't having to wait on someone else to get their shopping done. However, there's the chance that the self-checkout makes you stop and wait for an employee to come over, which depending on how busy the employee is at the time (if there's a lot of customers), can make the time it takes to checkout much longer and require the customer to wait, while if they were in the cashier's line, they would already have a cashier with them to either figure it out or would never run into the problem in the first place since they are with a cashier instead of the self-checkout system.

On the other hand, other online resources have said self-checkout *is* faster than checking out with a cashier [2]. However, it's possible that the article means that the speed of checkout is faster at self-checkout compared to checking out with a cashier because when you are in line for self-checkout, you're in line for 6 different machines where you could check out at once, while for cashier checkout, you only have 1 option and everyone in line is waiting for the same place to checkout.

For this project, the question is only addressing the speed of checking out, so the time in line waiting to check out is not included.

The Real World System

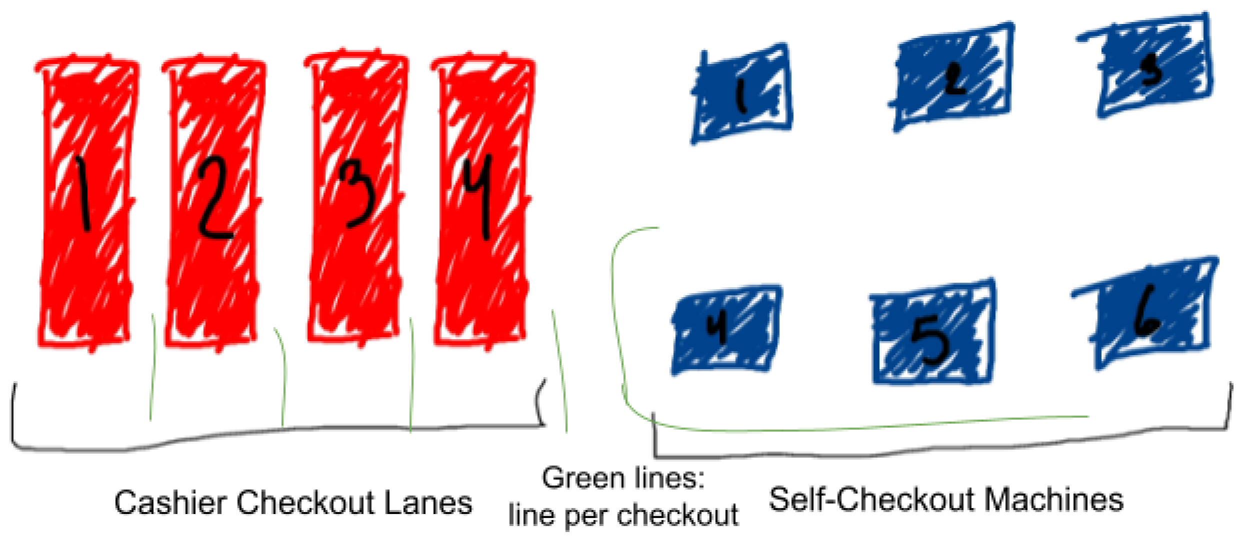

Figure 1: Meijer Checkout Layout

Author: Nicole Pililyan

The Meijer by Wright State University was studied and simulated for this project. The data for this was gathered on weekdays from the times of 6:00 p.m. to 8:00 p.m. At this location, there were four cashier checkout lanes open (although they were more spread out and had others in between them that were closed, unlike the above figure), and six self-checkout machines open.

When going through a cashier checkout line, a customer picks a cash register line to join, then waits until it's their turn to check out. They place their items on the conveyor belt, and the cashier rings up the items and places them in bags. Once they finish ringing up all the items, the cashier asks if the customer has "mperks", which is the

Meijer rewards system, and the customer pays for their items. Once the customer finishes paying, they take their bagged items and leave.

When checking out with the self-checkout machines, a customer can enter their mperks at any time during the process before they pay, but they are not able to ring up items while entering in their mperks ID and pin, and it takes a few seconds for the machine to go back to allowing the customer to ring up items again. The customer may also run into times where they have to wait to keep ringing up items because the machine makes them wait for an employee's assistance, or they may need to ask an employee for help during the checkout process. There are only one or two employees working the self-checkout stations at a time, so the customer has to wait for an employee to be available and aware that they need help, then wait for the employee to do what's needed with the machine before they can return to checking out their items. Depending on how busy the self-checkout stations are and how many people need assistance, the amount of time waiting for an employee to come over can be very short, or can take some time.

Some things I noticed while observing the system was that not every cashier seemed to ask if the customer had mperks, and that a customer can type in their mperks at any time while the cashier is ringing up the items. If a customer types in their mperks while the cashier rings up items, it can save them time while checking out that a customer wouldn't be able to save during self-checkout, since the machine doesn't allow them to ring up items until they finish entering their mperks ID and pin. Another thing I noticed was that different cashiers went at different speeds, which can affect how long it takes to check out. If they're a new employee, it can take them longer, and they may need to ask for help from other employees, which doesn't happen too often, but happened twice when I gathered data.

The Simulation Model

The simulated model for this project follows how the real-world system works with some exceptions. It was hard to tell how long exactly mperks took for the customers to enter in, so a set amount of time, 15 seconds, will be added for simulated customers that have mperks. There is an assumption that the customers who go through the cashier checkout will also get these 15 seconds, and that they enter in their mperks at the end of the cashier ringing up their items, while they pay.

The simulated model also has two variables for self-checkout when a customer needs assistance, one for waiting for assistance, and one for getting assistance, and the time is randomly chosen between the interval of 5 seconds and 45 seconds, the mean time given being 21.2 seconds for waiting, and 20.5 seconds for getting assistance.

Since there were only two data points with the cashiers that required assistance from other employees, a variable for cashiers needing assistance was not included in the simulation model.

The model runs two scenarios, the first being self-checkout and the second one being checkout with the cashier. Both arrive at the checkout lane at a random exponential variate of 3 minutes, or 180 seconds. Both have a random chance of the customer having mperks, based on the ratio from collected data of how many customers had mperks or not. Self-checkout also has an assist variable also based on a ratio from the collected data. Then, both scenarios checkout with a random lognormal variate, and the scenario is finished.

Analysis of Data

How the Data was Gathered and Processed

The data set was collected from the system studied between the hours of 6:00 p.m. to 8:00 p.m. on weekdays. The self-checkout data set and the cashier checkout data set have a column for the total time to check out, whether the customer used mperks or not, if they got assistance, and how long they waited for and got assistance for. To process the data, both data sets were run through a python script that would display the data.

Data Distributions

The Chi-Squared statistical significance test and the Kolmogorov-Smirnov (K-S) goodness of fit test were used for each dataset to make a decision on which theoretical input distribution represented the data best. The theoretical distributions tested were Exponential, Lognormal, Weibull, Gamma, and Uniform. The Uniform distribution was added in because I wasn’t sure if the other distributions would fit the self-checkout data, but in the end the distribution did not get used (see folder “Uniform_Distribution” in the Self-Checkout Data in the folder FitTests).

- Self-Checkout Data Distribution:
    - The input distribution observed to fit best was the lognormal distribution. It did not reject the null hypothesis, and after visual analysis, it was decided to be the best fit. The lognormal distribution gave a shape parameter of 0.361 and a scale parameter of 179.998. Using the Chi-Squared test with 35 bins, the computed Chi-Square statistic was 30.432 with a p-value of 0.596 (see Figure 1 below for a histogram). The K-S test computed a K-S

Statistic of 0.108 and a p-value of 0.985 (see Figure 2 below).

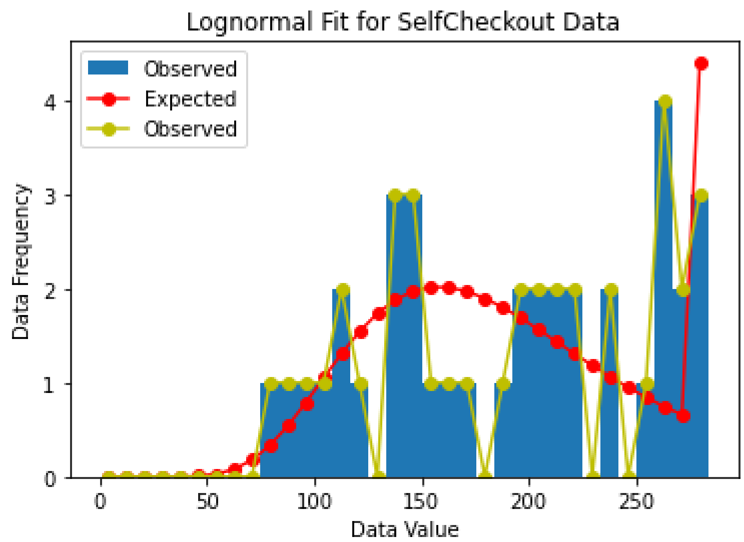

Figure 1: Self-Checkout Data Lognormal Fit using 35 bins.

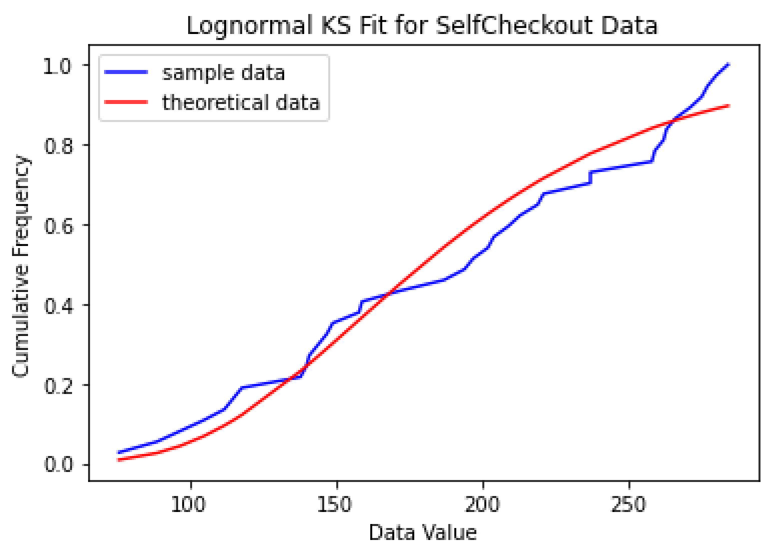

Figure 2: Self-Checkout Data Lognormal K-S Fit.

❖ Cashier Checkout Data Distribution:

➢ The input distribution observed to fit best was the lognormal distribution. It did not reject the null hypothesis, and after visual analysis, it was decided to be the best fit. The lognormal distribution gave a shape parameter of 0.3816 and a scale parameter of 168.2403. Using the Chi-Squared test with 20 bins, the computed Chi-Square statistic was 12.581 with a p-value of 0.816 (see Figure 3 below for a histogram). The K-S test computed a K-S Statistic of 0.0857 and a p-value of 0.9997 (see Figure 4 below).

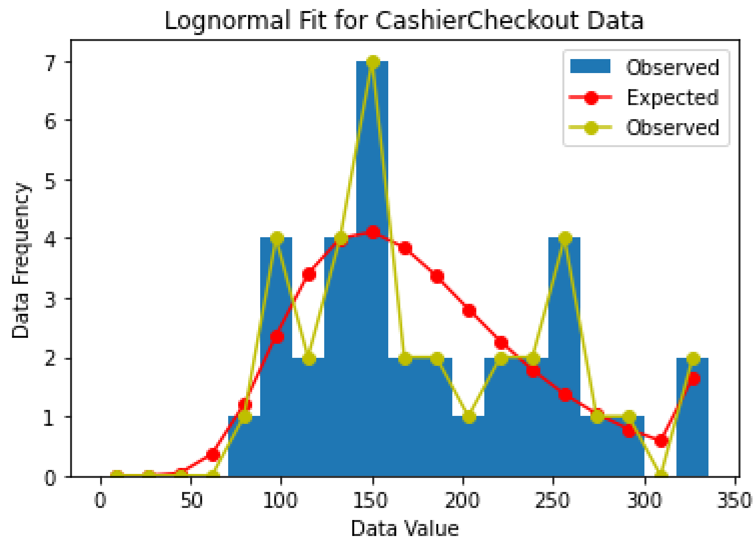

Figure 3: Cashier Checkout Data Lognormal Fit using 20 bins.

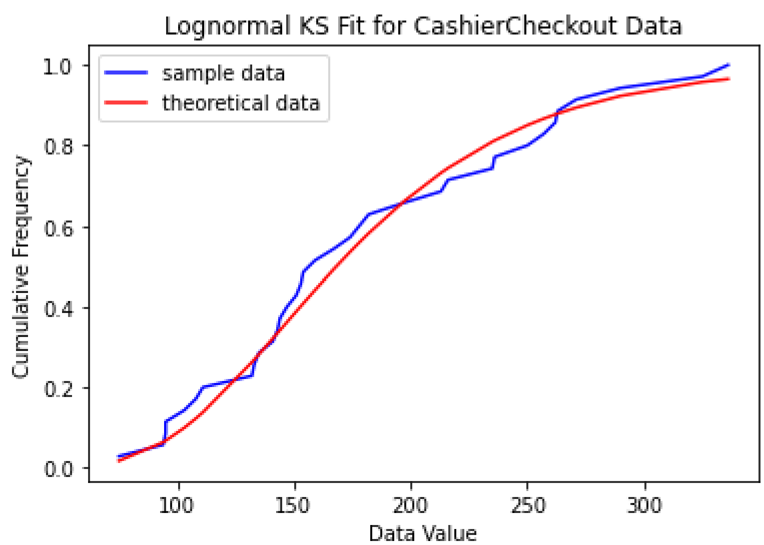

Figure 4: Cashier Checkout Data Lognormal K-S Fit.

Simulation Results

Simulation Implementation

The simulation code for this project was designed similarly to project 2. The complete project uses five python files and multiple directories holding related data. These files include the simulation code, processing the data, creating the data, and analyzing the data. Below are details of what each python file contains, and more details can be viewed in the python files themselves if needed.

- project3.py
  - Simulations and generating data was done here, along with printing results at the end of the file.
- processData.py
  - Processes the raw collected data and returns a list of the combined data from a data set (self-checkout or cashier checkout). Can print out results and plots if needed.
- inputTesting.py
  - Includes multiple methods for fitting and plotting Chi-Squared and K-S tests for the different distributions tested for this project, utilizing data from processData.py.
- createTID.py
  - Creates a theoretical input distribution (TID) reflecting project distributions
- validityTesting.py
  - Analyzes gathered data and compares it to the generated data. Confidence intervals are tested here.

Simulation vs Real System

Shown below are comparisons between the real system and the simulation model datasets. The simulation data was produced from 300 iterations of the simulation, while the gathered data was gathered only by one person, so there is not as much data available as there was possible for the simulation. Thanks to using 300 iterations, data variability was brought down for the simulation as well as higher confidence in the averages and variance metrics produced by the simulation, while the real system data may have more variability in it.

Data-Type	Collected Mean	Collected STD	# of Collected Samples
Self-Checkout	191.216	7.884	38
Cashier Checkout	180.714	8.254	36

Table 1: Real System Collected Data, in seconds, Mean and STD.

Data-Type	Simulated Mean	Simulated STD	Minimum Mean	Maximum Mean	# of Simulated Samples
Self-Checkout	214.702	6.981	182.344	257.251	12039
Cashier Checkout	187.165	7.241	154.709	220.621	12024

Table 2: Simulation System Collected Data, in seconds, Mean and STD.

The plots below show a side by side binned data distribution for self-checkout and cashier checkout datasets. The real-system data is on the left and the simulation model data is on the right.

Although this is to show how the generated data follows a similar distribution as the collected data does, please keep in mind that the collected data had much less data points to go off of compared to the simulation model, so there is much more data variability in the plots.

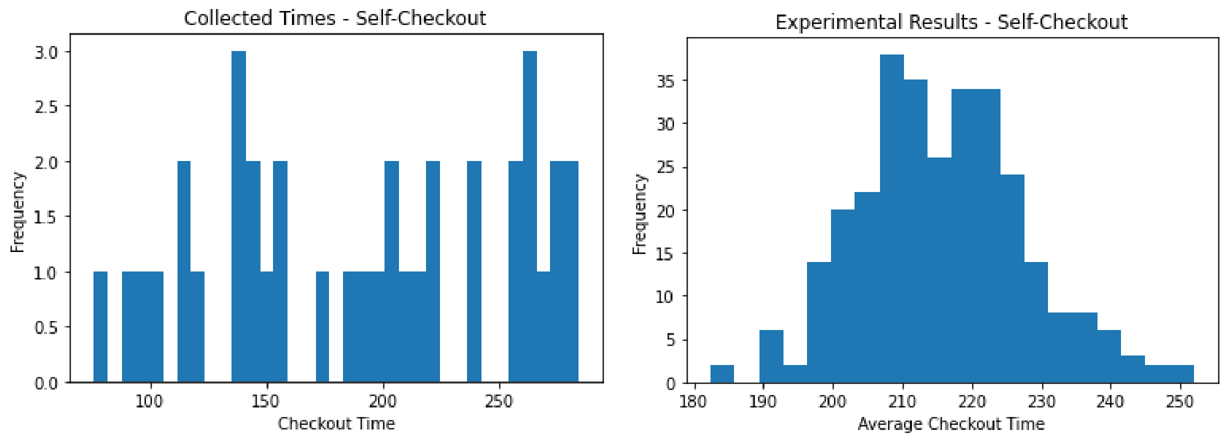

Figure 5: Comparison of Collected vs Generated Self-Checkout Data

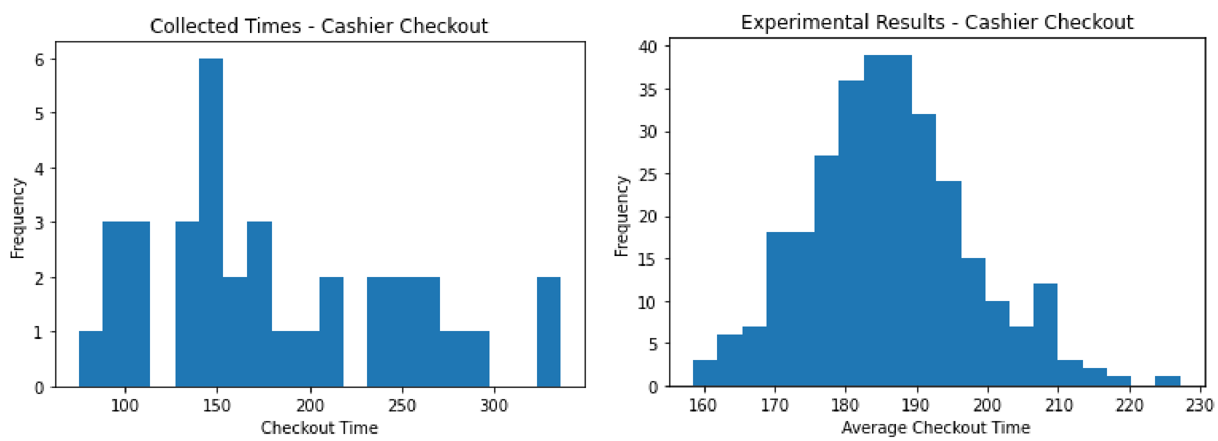

Figure 6: Comparison of Collected vs Generated Cashier Checkout Data

In terms of confidence intervals, a 95% and 99% confidence interval was checked with the datasets. Below are the results.

DataSet	95% Confidence Interval	99% Confidence Interval
Self-Checkout Collected	170.911s to 211.521s	164.531s to 217.901s
Self-Checkout Simulated	191.394s to 238.0098s	184.0703s to 245.334s

Cashier Checkout Collected	170.911s to 211.521s	164.531s to 217.901s
Cashier Checkout Simulated	164.887s to 209.443s	157.887s to 216.443s

Table 3: Collected Data and Simulated Data Confidence Interval Results, in seconds

Conclusion

In this project, I researched and tested whether it is quicker to checkout at a self-checkout machine or with a cashier at the Meijer by Wright State University. Based on the results of this project, checking out with a cashier, not waiting in line included, is faster. With a 99% confidence interval, checking out with a cashier is about 26 to 29 seconds faster, according to the simulation's result data.

Does this mean that overall, going through a cashier checkout will be faster than a self-checkout? I believe that it depends on the situation. It depends on if the cashier is new, if they're experienced, and if there's a line. If the self-checkout line has a machine open without waiting in a line, and the cashier has no line, I believe the cashier should be faster. However, if the cashier has a line of more than one person, and a self-checkout machine is open with no line, self-checkout will probably be faster. If both have a line, it all depends on the length of the line, as self-checkout only has 1 line for 6 different machines, while if you get in a line with the cashier, you must wait for that cashier only.

Overall, the cashier will in general be quicker when it comes to checking out, however with the availability of waiting for six self-checkout machines in one line, one may finish checking out faster that way.

Sources

1. Strong Point. (2017, November 20). *The self-checkout paradox – faster or slower?* iXtenso. Retrieved November 29, 2022, from https://ixtenso.com/technology/the-self-checkout-paradox-faster-or-slower.html.
2. Schmidt-Jacobsen, T. (2020, October 13). *Advantages and disadvantages of 6 retail self-checkout systems*. Sprinting Retail. Retrieved November 29, 2022, from https://sprintingretail.com/blog/retail-self-checkout-systems/#:~:text=Since%20self%2Dcheckout%20is%20typically,comes%20to%20cost%20of%20space.