Exploring Left-digit Bias in Bargained Auto Loans

Yuxuan Wang (yxw209), Yuan Huang (yh2741)

Abstract

The first part of this paper is a replication study of "An Empirical Bargaining Model with Left-Digit Bias: A Study on Auto Loan Monthly Payments", which examines the price bargaining in auto loans where both parties are subject to the left digit bias when processing numbers. We use a synthetic dataset generated by running a Monte Carlo simulation on the original dataset with a few variables excluded. We find that the left digit bias is evident and our replication resembles the original work. The bunching effect for monthly payments is found at both the 9-endings and 0-endings, with 9-endings loans having higher interest rates and 0-endings loans having lower interest rates. Building on the replication, we further explore systematic differences between different payment levels with the same last digit, and whether the middle digits influence the distribution of loans. We found that the majority of the bias comes from 99-endings and 00-endings, and the distribution pattern is different for cheap and expensive loans with 9 as the last digit. This leads to the middle digits having the opposite effect on cheap and expensive loans. For the cheap loans, 99-endings are dominant and for expensive loans, 09-endings are dominant. From 09-endings to 89-endings, the percentages of loans are increasing in cheap loans and decreasing in expensive loans. Overall speaking, this opposite effect makes the middle digits display no impact on the aggregated level.

Section I: Overview and Previous Work

1.1 Overview

Our reference paper empirically studies a bargaining environment in which the results are affected by left digit bias and bargaining power. In the bargaining setting, there are two parties, the consumers and the dealers. The basic setting is that the consumer wants to loan a specific amount of money for the cars through an auto loan dealer. Both the consumer and dealer decide on the length of the loan and monthly payment by bargaining, and the interest rate of the loan payment will be simultaneously determined. Through data analysis, the author finds that consumers and dealers both have the left-digit bias and it drives the outcome of monthly payments further away from what the standard model predicts. The consumer has the left-digit bias to limit payments at 9-endings, while dealers consider 0-endings as a more profitable result. This makes the bargaining result different from when only one or neither of the two parties exhibit left-digit bias. The left-digit bias describes how people have the tendency to focus on the left-most digit of a number while partially ignoring other digits. For example, a number ending in 99 (e.g., $399) may be perceived as much lower than the next whole number (e.g., $400). This is where a dollar increase is more in people's minds than an actual dollar increase. There are various researches examining the left digit bias and we will further introduce it in the following literature review part.

We operate our replication study on simulated data provided by the author, and the results resemble the findings of the original paper. Both find that the percentage of loans with 0-ending and 9-ending payments is significantly larger than those with any other ending digits, and the bias is more significant in 99-endings and 00-endings that cross over $100 in payment. Moreover, the interest rate is higher for the 9-endings compared with the average rate for other endings, while the 0-endings have lower than average interest rates. The lack of variables in the original dataset hinders our replication from going further into the analysis of the demographic features of customers and their bargaining power.

Based on the replicated work, we then explore the systematic difference among different payment levels of the same last digit, and whether the middle digits have an impact on the distribution of loans. For instance, we compare the left-digit bias between 89-endings to

$90-endings and 49-endings to 50-endings to see whether variation in the middle digits results in different percentages of loans. We found that 99-endings and 00-endings account for most of the overall unevenness of distribution caused by the left-digit bias and that the distribution pattern is different for cheap and expensive loans. The middle digits have the opposite effect on these two types of loans, making the aggregated effect neutralized. In addition, we also look at the percentage distribution of payments with different last ending digits in equal-length intervals, for example, payments with 10-20 endings and those with 80-90 endings. We found no systematic difference in distribution patterns and the results are consistent with the reference paper and our replication. If given more information and resources, we could also perform robustness checks on repeatedly simulated data, but unfortunately, this is beyond the scope of this paper.

The rest of this paper is organized as follows. After the literature review below, Section II introduces the work of our reference paper, including the original dataset and model setting. Section III reports the findings from our reference paper about left-digit bias and interest rate patterns for bargained auto loans, as well as our replication of these findings and the similarity or differences between our replication and the original. Section IV introduces our own analysis of the synthetic dataset and it can be seen as extension work. Section V concludes the findings and contribution of this paper.

1.2 Literature review

Left digit bias is a term used in numerical cognition and behavioral pricing that suggests that number size judgments are aimed at the leftmost digit of a multi-digit number. The left digit bias refers to people mistakenly evaluating or perceiving a multi-digit number and overemphasizing the left-most digit. For example, people may think the difference between $9.00 and $7.95 is larger than the difference between $9.05 and $8.00. Left digit bias was first pointed out by Thomas and Morwitz in 2005 in a Journal of Consumer Research article. In the article, the finding of the left digit bias explains why the 9-ending prices are prevalent in the market compared to other numbers as the ending digit. They also find that those 9-ending prices that do not influence the left-most digit will not change the customer behavior. Hence, only the 9-ending digit causes the left-most digit change will cause consumers to perceive a more than 1 dollar change in mind. Therefore, the left-digit bias has implications for cognitive psychology.

There are many kinds of research that prove the existence of the left digit bias in consumer behavior. Stiving and Winer (1997) examined the left-digit effect with scanner panel models. They proposed that 9-ending prices can influence consumer behavior through two different processes: the image effect and the level effect. The image effect refers to those effects where consumers may infer meaning from the number on the right-hand side. The level effect refers to those effects where consumers may undervalue the price. Their results suggest that both of these effects account for the influence of 9 ending prices in the shops. Choi, Lee, and Ji (2012) examined the interactive effects of a 9 ending price and message framing in advertisements. The researchers found that consumers were much more receptive to the advertisements when the 9 ending price was paired with a positive message. This, in turn, increased the likelihood that they would make a purchase decision. Lacetera, Pope, and Sydnor (2012) studied the impact of the left digit bias when reading odometer values in the used car sales market. They found that once the leftmost digit changed at the 10,000-mile threshold, the car's valuation continued to decline (e.g., 19999 miles vs. 20000 miles). This effect was also reflected in the 1,000-mile threshold (e.g., 1999 vs. 2000 miles) but with a smaller effect.

The left digit bias also has an impact on the financial market, public policy, and the medical domain. Bhattacharya, Holden, and Jacobsen (2011) show that the left digit bias can influence the stock-market transactions. They find that there exists excess buying between just below the price ($1.99) and just above the whole number of prices ($2.00). This buy-sell difference leads to a significant change in returns. Olsen (2013) examined how citizens' evaluations of public school districts in the Danish population changed significantly based on the leftmost digit of the average grade. They find that small changes in average grades resulting in a shift in the leftmost digit will cause citizens' reactions to become more dramatic, thus changing their stance on public policy on the issue. Olenski et al. (2020) find that physicians were significantly less likely to recommend cardiac surgery for a whole number of patients (80 vs. 79). Physicians overreact to a patient's age when diagnosing a disease, suggesting that seemingly unrelated factors, such as a few weeks of age difference may influence a physician's treatment decisions and may alter the consequences of life for the patient.

Section II: Works from Our Reference Paper

In this section, we present the dataset, model setting and findings of our reference paper, “An Empirical Bargaining Model with Left-Digit Bias: A Study on Auto Loan Monthly Payments” by Zhenling Jiang. Zhenling Jiang is an Assistant Professor of Marketing at the Wharton School, University of Pennsylvania. This paper is published in 2021 in Management Science Journal.

2.1 Dataset

The original dataset contains 35 million anonymized data on individual credit profiles provided by one of the three major credit bureaus in the United States. It includes all non-subprime auto loans originating from banks or credit unions in the United States from 2011 to 2014. In comparison, our synthetic dataset is generated by running a Monte Carlo simulation on the original dataset, with a few variables excluded and of a much smaller sample size of 1 million. 9 attributes are common in both datasets, 4 of which describe basic information about the loans, another 4 describe information about the consumers and one attribute about the car dealers.

The loan conditions include the length of the loan (term), the interest rate (int), the average monthly payment (payment), and the total loan amount (high credit). Total payment could be calculated based on the first three variables. The analysis from the reference paper and our replication will focus on average monthly payments and interest rates. Information on the consumers was applied by the reference paper to extrapolate the relative bargaining power of consumers. These data include their credit score when the loan was initiated and the percentage of the black and hispanic population in the consumers’ residential areas. In addition, the author also inferences the reservation price of the consumers and dealers, which will serve as upper and lower bounds of the monthly payment of loans.

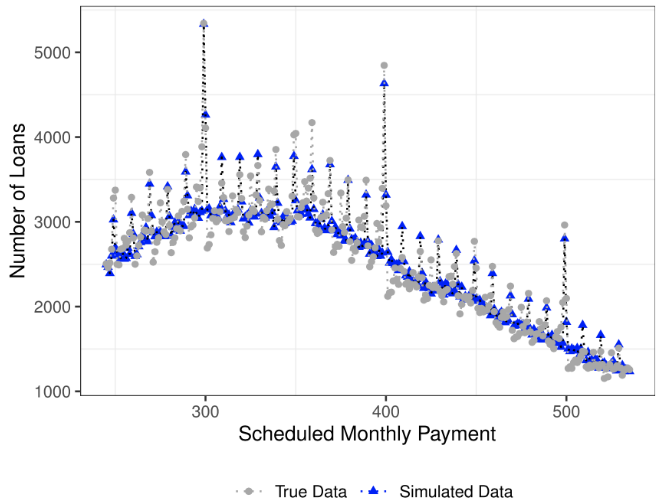


Figure 1: Comparison of the original data and simulated data

2.2 Model Setup

2.2.1 Perceived payment

Considering the left-digit bias people have in the bargaining, the perceived value of the monthly payment may differ from the actual value of the payment. In the reference paper, the author denotes the following equation to capture the perceived payment.

$$\hat{p} = [p]_{100} + (1 - \theta_1)\,[p]_{10} + (1 - \theta_1)(1 - \theta_2)\,[p]_1$$

Where the $[p]_{100}$ is the hundred of a payment, $[p]_{10}$ is the tens of the payment, and $[p]_1$ is the single last digit of the payment. For example, if the payment is \$123, then $[p]_{100} = 100$, $[p]_{10} = 20$, and $[p]_1 = 3$. The θ_1 captures the left-digit bias for the ten digits, and θ_2 allows an increased bias for the single digit. Both θ_1 and θ_2 have the domain from 0 to 1. When $\theta_1 = \theta_2 = 0$, there is no bias, and people's perceived value is the same as the actual payment. When $\theta_1 = \theta_2 = 1$, all the values besides the hundreds digit are ignored. It means that if we back to the example, people have bias with $\theta_1 = \theta_2 = 1$ will view the payment of \$123 and \$100 as the same value. In this case, the perceived value is normally smaller or equal to the actual value. And for the cases that the payment is exactly on the $100 mark, for instance, $200, the perceived value and the actual value are the same no matter whether people have bias or not. From the equation setting, we could find that there is a discontinuity large perceived difference increase if the payment crosses the $10 mark and even larger if it crosses the $100 mark. And it also assumes that for each $10 mark gap, the perceived difference is the same for the same person. This calls for further examination through data exploration in Section IV. Here in Figure 2 we showcase the effect of left-digit bias on perceived payment.

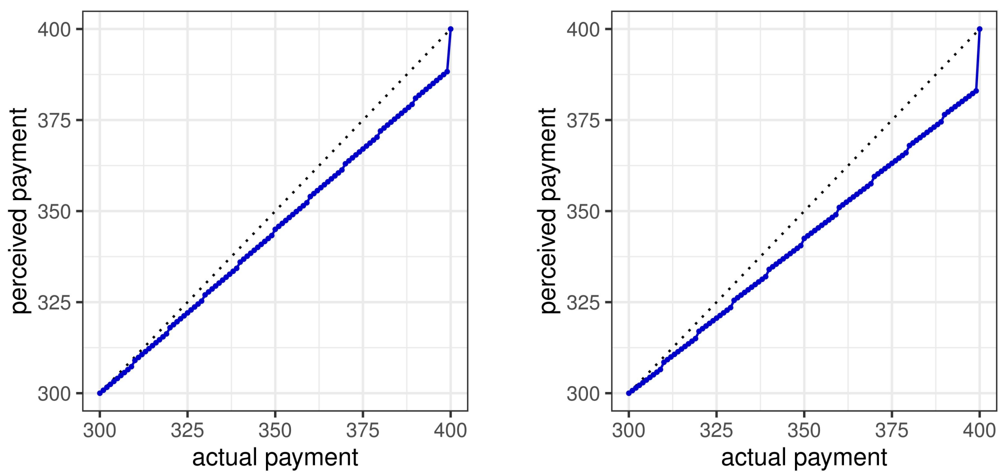


Figure 2: Example of the perceived payment for $\theta = 0.1$ (Left) and $\theta = 0.15$ (Right) (Using synthetic dataste that the author provided)

2.2.2 Nash bargaining with joint-value maximization problem

The author of our reference paper applies the Nash bargaining model to describe how monthly payments are influenced by the relative bargaining power of the two parties. According to the paper, the Nash bargaining solution is a convenient way to describe the results if we do not observe the actual bargaining process or if there is information about cases of bargaining failure. The setting is that the consumer has already determined the total loan amount before the negotiation. This assumption is specific to the auto loan industry setting. The two parties will then negotiate on monthly payments, a process during which the loan interest rate is also determined given the other conditions of loan duration and total amount. Regardless of the outcome of negotiation, monthly payment and the interest rate will be positively correlated. Hence, we assume the monthly payment is the key point for the bargaining process as consumers tend to pay more attention to the decision for the monthly payment instead of the interest rate.

Nash bargaining model assumes that for loan i, the monthly payment p_i will maximize a joint-value function as following:

$$v(p_i) = u_c(p_i)^{w_i} * u_f(p_i)^{1-w_i}$$

Where w_i is the consumer's relative bargaining power, ranging from 0 to 1, and the dealer's corresponding bargaining power is $1 - w_i$. $u_c(p_i)$ represents the consumer's surplus:

$$u_c(p_i) = r_{i,c} - \widehat{p}_i(\theta_c)$$

Where $r_{i,c}$ is the consumer's reservation value which has given in the dataset and $\widehat{p}_i(\theta_c)$ is the consumer's perceived payment as defined in section 3.1. $u_f(p_i)$ represents the dealer's surplus:

$$u_f(p_i) = \widehat{p}_i(\theta_f) - r_{i,f}$$

Where $r_{i,f}$ is the dealer's reservation value which has given in the dataset and $\widehat{p}_i(\theta_f)$ is the dealer's perceived payment as defined in section 3.1.

Also, the bargaining power of the consumer varies across different characteristics of the consumers. For instance, the credit score and the percentage of people of different ethnicities in the consumers' residential areas may potentially determine the bargaining power of the consumers. Hence, we assume the bargaining power of the consumer is defined as:

$$w_i = F(x_i\beta) + \epsilon_i$$

Where F is the logistic function defined as $F(x) = \frac{e^x}{1-e^x}$ and x_i includes a constant and a vector of consumer information about the credit score and the percentage of the black and hispanic population in the consumers' residential areas. ϵ_i is the error term and we assume it follows a normal distribution.

2.2.3 Interest rate estimation with linear regression

With regard to left-digit bias, a reasonable assumption made by the reference paper is that the interest rate of loans is correlated with the ending digits of payments. Beyond the ending digit, the interest rate may also relate to the consumer characteristics. For example, the credit score of the consumers may influence the interest rate substantially. Hence, the reference paper adopted this regression model for interest rate estimation.

$$int_i = \sum_{j=1}^{9} \gamma_j I[d(payment_i) = j] + x_i\beta + \epsilon_i$$

Where int_i represents the interest rate of loan i, and $I[d(payment_i) = j]$ is an indicator variable. The indicator variable equals to 1 if the ending digit of payment is equal to j where j is range from 1 to 9, and the indicator variable equals to 0 if the payment is not equal to j. The variable x_i includes the loan amount, the length of the loan, the characteristics of the consumers such as the credit score and the percentage of the black and hispanic population in the consumers' residential areas. The state and time fixed effects are also considered in the regression. To further relax the nonlinearity of the regression, the quadratic terms of the variables are involved.

Section III: Replication Results and Comparison with Our Reference

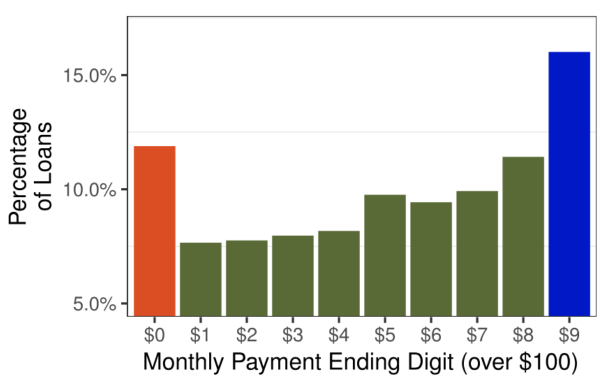
In this section, we present the findings from our reference paper about left-digit bias and interest rate patterns for bargained auto loans, and how consistent or different our replication is from the original one. We focus our replication study on four aspects: the evidence of left-digit bias in our dataset, how does the average interest rate differs among ending digits, what is the estimation from the linear regression model for interest rate, and finally what would happen if one or neither of the two parties exhibit left-digit bias. For each part, findings from our reference paper will be first discussed, followed by our replication results.

We performed our analysis on a synthetic dataset provided by the author of our reference paper. It was unfortunate that due to information security issues, the original dataset could not be provided for replication of findings. Our synthetic dataset is generated by running a Monte Carlo simulation on the original dataset, with a few variables excluded. The synthetic dataset we work on consists of one million observations with a total of 9 attributes, 4 of which describe basic information about the loans, another 4 describes information about the consumers and one attribute about the car dealers. The basic statistics of the attributes from the synthetic dataset are provided in Table1.

variable	mean	sd	min	max
Black population in borrower residential area (%)	0.09	0.14	0	1
Borrower credit ccore	724.61	60.56	620	830
Borrower reservation price	437.14	160.01	179	1732
Dealer reservation price	372.3	146.46	152	1521
Hispanic population in borrower residential area (%)	0.1	0.13	0	1
Interest rate	0.05	0.03	0	0.24
Length of loan	5.65	0.94	3	7
Monthly payment	389.89	148.31	153	1611
Total loan amount ($1000)	22.8	8.33	10	60

Table 1: Summary Statistics of the simulated data

Figure 3 below demonstrates the distribution of loans with different ending digits. The baseline for comparison is that if there is no left-digit bias, then the negotiated payments should follow standard models and display a universal 10% coverage among all the ending digits. Contrary to that, both the original paper and our replication display a bunching effect in loans with 9-ending and 0-ending digits, and correspondingly an under-representation of loans with any other ending digits. According to our reference paper, both the consumers and dealers negotiate under the influence of left-digit bias, so both would parties view the gap between 9-endings and 0-endings as a value difference larger than $1, and for any other gaps, for instance between 3-endings and 4-endings, the perceived value difference is less than $1. Following this bias in perceived value, consumers will have the incentive to keep payment at or below 9-endings, while the dealers will have the incentive to increase payment from 9-endings to 0-endings. What's more, once the dealers succeed with this increase from 9-ending digit to 0-ending digit, there is room for further payment increase since consumers don't mind the increase from 0-endings to 9-endings that much. This is clearly reflected in both graphs and it explains why we observe an increasing percentage from 1-ending payments to 9-ending payments. It is worth noting that one difference between the original figure and our replication is that in the original paper there exists an unusual bunching at 5-ending loans. The author offers an explanation that 5-endings may be perceived as a rounding number similar to 0-endings, which leads to a weaker bunching effect. We don't observe this phenomenon in our replication and a possible reason is the simulation nature of our synthetic dataset.



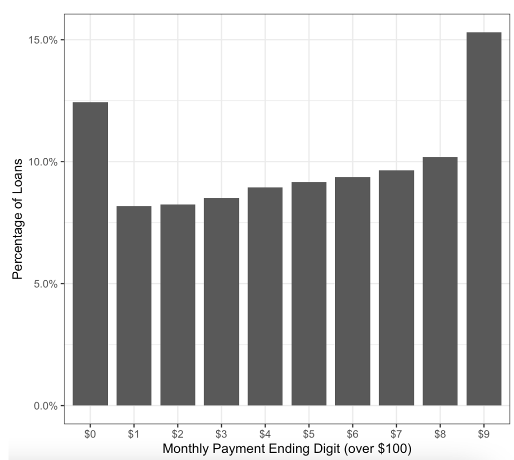


Figure 3: Distribution of different ending digits (original on left, replication on right)

Figure 4 below demonstrates the average interest rate for loans with payment different ending digits. This is a visual representation of the interest rate regression results in Figure 5. It is apparent that in both the original graph and our replication, there is a huge gap in interest rates between 9-ending and 0-ending digit loans. The author further investigates this phenomenon in the paper that this is related to the relative bargaining power of consumers, which is dependent on their income, ethnicity, age and so on. In our replication these information are not available, and we offer a simple explanation that as a result of negotiation, sometimes the payment stays with a 9-ending digit in favor of consumers but has a relatively higher interest rate in favor of dealers, and sometimes the opposite situation occurs where the payment goes up to the next 0-ending digit in favor of dealers, but the interest rate is relatively lower in favor of consumers. Here again, we see the difference between the original graph and our replication occur at the 5-ending loans. While the original study found that the interest rate for 5-endings is lower than those of 4-endings, in our replication we only observe an overall trend that when the ending digit increases, the interest rate also mildly increases. The author attributes this to people perceiving 5-endings as rounder numbers similar to 0-endings and simultaneously perceiving 4-endings similar to 9-endings. This inconsistency may come from the difference between the datasets. Another difference is that in our replication the loans with 7-endings seem to have a lower interest rate, and we suspect it has to do with the excluded variables from the original dataset that causes regression results to differ, which is also shown in the next figure.

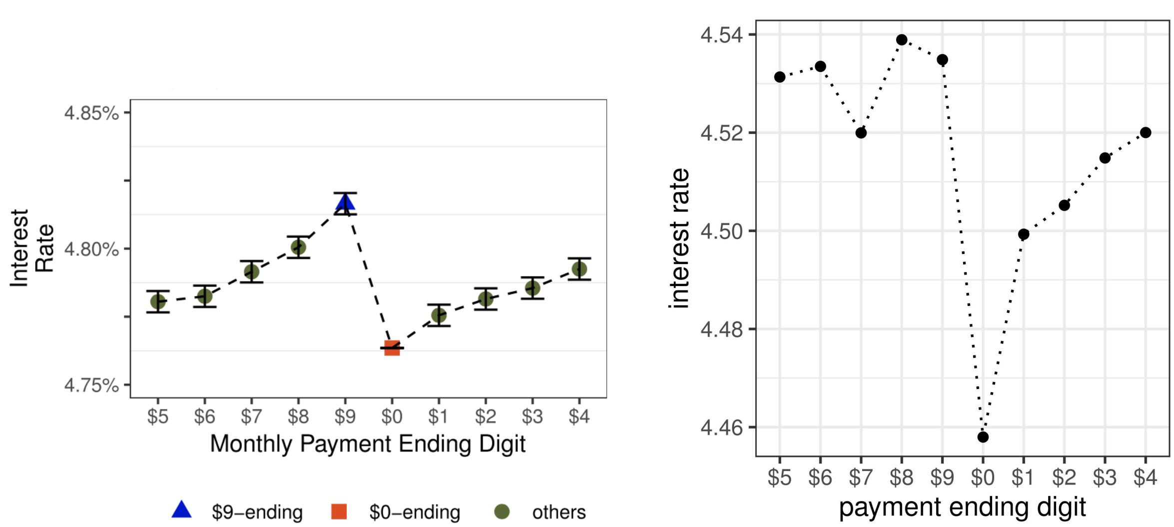


Figure 4: Average interest rate for different ending digits (original on left, replication on right)

Figure 5 below demonstrates the regression results for the interest rate model presented in section 3.3. According to the specification of our reference paper, Columns 1 and 3 use linear and quadratic terms of these variables, while Columns 2 and 4 categorize them into bins and use bin fixed effects. Columns 3 and 4 also include additional consumer characteristics of ethnicity distribution in the residential area of consumers. Agreeing with what Figure 4 implies, in both the original study and our replication, there is a considerable difference between how the 0-ending digit or the 9-ending digit condition affects the interest rate of these loans. The difference is that due to several demographic variables being taken out in replication, the interaction term between ethnicity percentage and the 9-ending digit payments are no longer statistically significant for deciding the interest rate. This shows that ethnicity alone is insufficient to extrapolate how the relative bargaining power is shifted when dealers are negotiating with different consumers.

	Dependent variable:				
	APR (%)				
	(1)	(2)	(3)	(4)	(5)
$1-ending	0.0116***	0.0131***	0.0130***	0.0143***	0.0130***
	(0.0018)	(0.0018)	(0.0018)	(0.0018)	(0.0018)
$2-ending	0.0185***	0.0193***	0.0194***	0.0200***	0.0194***
	(0.0018)	(0.0018)	(0.0018)	(0.0018)	(0.0018)
$3-ending	0.0225***	0.0237***	0.0235***	0.0245***	0.0235***
	(0.0018)	(0.0018)	(0.0018)	(0.0018)	(0.0018)
$4-ending	0.0290***	0.0310***	0.0290***	0.0309***	0.0290***
	(0.0018)	(0.0018)	(0.0018)	(0.0017)	(0.0018)
$5-ending	0.0174***	0.0208***	0.0166***	0.0200***	0.0166***
	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0017)
$6-ending	0.0187***	0.0219***	0.0191***	0.0221***	0.0191***
	(0.0018)	(0.0017)	(0.0018)	(0.0017)	(0.0018)
$7-ending	0.0277***	0.0308***	0.0275***	0.0305***	0.0275***
	(0.0018)	(0.0017)	(0.0018)	(0.0017)	(0.0018)
$8-ending	0.0365***	0.0403***	0.0355***	0.0393***	0.0355***
	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0017)
$9-ending	0.0523***	0.0614***	0.0478***	0.0571***	0.0422***
	(0.0017)	(0.0017)	(0.0017)	(0.0016)	(0.0054)
Age (in 100)			0.3476***	0.3413***	0.3452***
			(0.0029)	(0.0028)	(0.0030)
Income (in $1 million)			–3.371***	–3.3371***	–3.3574***
(zip-level)			(0.0110)	(0.0108)	(0.0116)
African American proportion			1.0671***	0.9855***	1.0618***
(zip-level)			(0.0035)	(0.0034)	(0.0037)
Hispanic proportion			1.4828***	1.3884***	1.4791***
(zip-level)			(0.0056)	(0.0055)	(0.0058)
$9-ending: Age (in 100)					0.0188**
					(0.0081)
$9-ending: Income (in $1 million)					–0.1148***
(zip-level)					(0.0317)
$9-ending: African American proportion					0.0407***
(zip-level)					(0.0091)
$9-ending: Hispanic proportion					0.0291**
(zip-level)					(0.0118)
Covariates *x*	Quadratic	Categorical	Quadratic	Categorical	Quadratic
Date opened fixed effects	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	34,760,946	34,760,946	34,760,577	34,760,577	34,760,577
R^2	0.3002	0.3173	0.3083	0.3246	0.3083

	Dependent variable:				
	interest rate				
	(1)	(2)	(3)	(4)	(5)
as.factor(type)1	0.0413***	0.0416***	0.0406***	0.0409***	0.0406***
	(0.0076)	(0.0075)	(0.0076)	(0.0075)	(0.0076)
as.factor(type)2	0.0472***	0.0483***	0.0469***	0.0480***	0.0469***
	(0.0076)	(0.0075)	(0.0076)	(0.0075)	(0.0076)
as.factor(type)3	0.0569***	0.0588***	0.0568***	0.0587***	0.0568***
	(0.0076)	(0.0075)	(0.0075)	(0.0075)	(0.0075)
as.factor(type)4	0.0620***	0.0637***	0.0618***	0.0635***	0.0618***
	(0.0075)	(0.0075)	(0.0075)	(0.0074)	(0.0075)
as.factor(type)5	0.0734***	0.0748***	0.0727***	0.0742***	0.0727***
	(0.0075)	(0.0074)	(0.0075)	(0.0074)	(0.0075)
as.factor(type)6	0.0755***	0.0777***	0.0754***	0.0776***	0.0754***
	(0.0074)	(0.0074)	(0.0074)	(0.0073)	(0.0074)
as.factor(type)7	0.0620***	0.0628***	0.0626***	0.0634***	0.0626***
	(0.0074)	(0.0073)	(0.0074)	(0.0073)	(0.0074)
as.factor(type)8	0.0809***	0.0805***	0.0806***	0.0801***	0.0806***
	(0.0073)	(0.0072)	(0.0073)	(0.0072)	(0.0073)
as.factor(type)9	0.0769***	0.0778***	0.0766***	0.0775***	0.0729***
	(0.0070)	(0.0070)	(0.0070)	(0.0070)	(0.0087)
vantage	-0.2907***		-0.2907***		-0.2907***
	(0.0007)		(0.0007)		(0.0007)
I(vantage2)	0.0002***		0.0002***		0.0002***
	(0.000001)		(0.000001)		(0.000001)
terms	-1.5821***		-1.5833***		-1.5833***
	(0.0153)		(0.0153)		(0.0153)
I(terms2)	0.2215***		0.2216***		0.2216***
	(0.0014)		(0.0014)		(0.0014)
highcredit	-0.2570***		-0.2570***		-0.2570***
	(0.0009)		(0.0009)		(0.0009)
I(highcredit2)	0.0031***		0.0031***		0.0031***
	(0.00002)		(0.00002)		(0.00002)
black			0.5273***	0.5273***	0.5222***
			(0.0124)	(0.0123)	(0.0133)
hispanic			0.6848***	0.6897***	0.6846***
			(0.0132)	(0.0131)	(0.0141)
black_9					0.0402
					(0.0372)
hispanic_9					0.0019
					(0.0405)
Constant	124.5325***		124.4169***		124.4164***
	(0.2716)		(0.2710)		(0.2710)
Observations	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000
R2	0.5999	0.6073	0.6017	0.6091	0.6017
Adjusted R2	0.5999	0.6071	0.6017	0.6089	0.6017
Note:					*p<0.1; **p<0.05; ***p<0.01

Figure 5: Regression results for interest rate model (original on left, replication on right)

Figure 6 below demonstrates the varying results of loan distribution under different assumptions about whether the consumers or dealers exhibit left-digit bias. The original finding and our replication are highly similar. When both parties exhibit the bias, the results turn out like the right-most graph in Figure 6, which is also consistent with Figure 3. However, when only one party exhibits the bias, we observe that the percentage of 1-ending to 8-ending digits is in accordance with the 10% baseline distribution, while some percentage of the negatively-biased ending digit will fall into the percentage of the favored ending digit. For consumers, 9-ending digits are favored, and for dealers, 0-ending are favored, so the results for them are reversed. In addition, we observe that the relative change between 9-ending and 0-ending digits for both consumers and dealers are very similar, suggesting that professional dealers are not that sophisticated in terms of overcoming left-digit bias, at least in the scope of this paper. The reference paper also shared this interpretation, but through mathematical methods in another section of the paper.

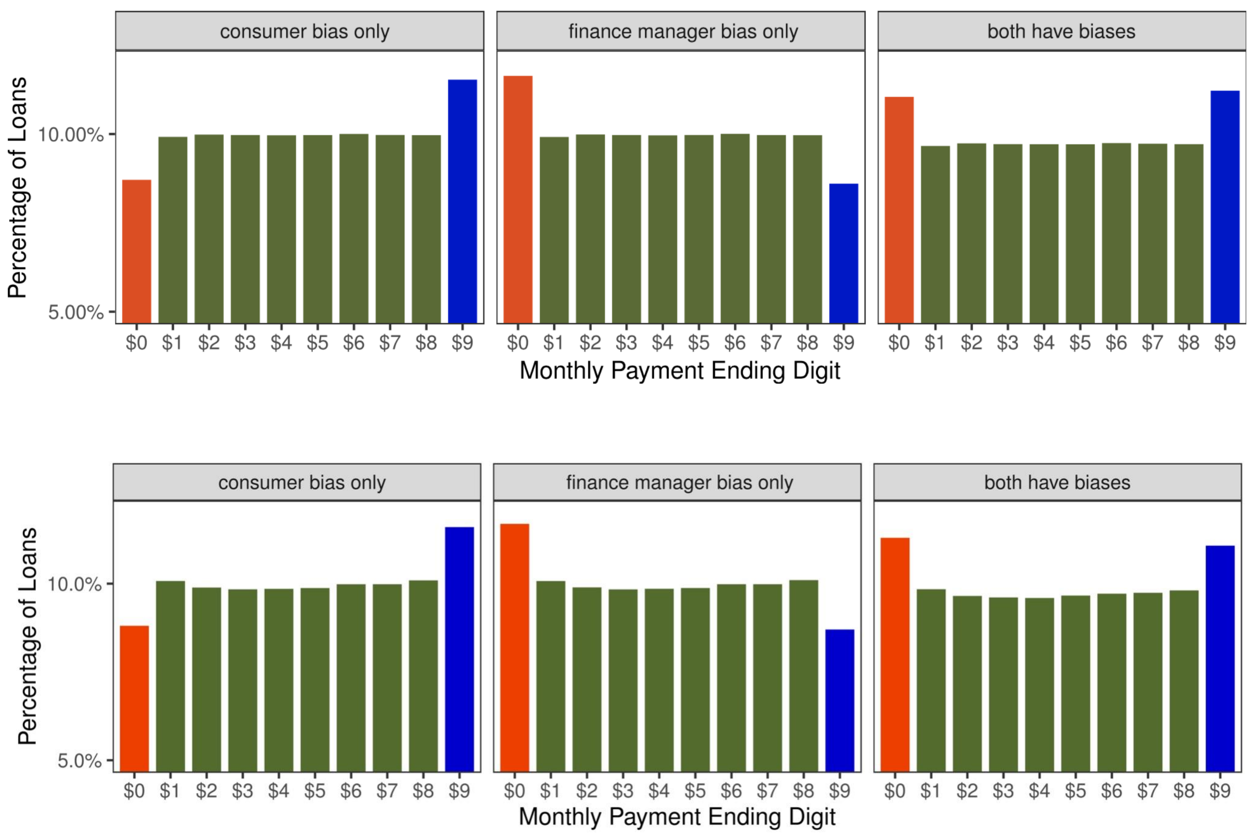


Figure 6: Comparison between different bias settings (original on top, replication on bottom)

Section IV: Innovation and Extended Analysis

In this section, we present our innovative study that builds on the previous replication results, focusing on checking robustness through data simulation and a further examination of how the bunching effect and left-digit bias could vary in different levels of payments with the same ending digits.

The first innovation that we were unfortunately unable to proceed with is to check if the replication results are consistent with the original findings after several rounds of data simulation. The author of our reference paper provided a function to perform Monte Carlo simulation, but that was only for generating graphs instead of simulating the whole dataset, so a lack of simulated variables prevented us from furthering our analysis. In addition, the simulation also involves a characteristic vector containing encoded information on the original dataset. The meaning of this vector is unknown, which makes running multiple simulations impossible. Monte Carlo simulation is commonly adopted to simulate results through probabilistic methods that help demonstrate the effect of random exogenous shocks. If successfully implemented, this could help us verify the robustness of our replication findings as to whether they still hold under the influence of uncertainty.

We were successful with our second innovation to perform analysis on segmented data and we have received rather interesting results. While the reference paper analyzes the aggregated effect of left-digit bias for all payments and studies the bunching effect of the last digit, we wish to investigate whether there's a systematic difference among different payment levels of the same last digit. This leads to us segmenting the data with respect to the last two ending digits, and the inclusion of one more digit can help us distinguish payment levels that involve changes in the left-most digit between payment levels that involve changes in the middle digits. For instance, from Section II we know that 9-ending payments are clear evidence for left-digit bias, but it could be that the bias is not evenly distributed across all levels of payment, that the bias is more prominent for 89-90 ending loans than those loans with 39-40 endings. In addition, we also look at the percentage distribution of payments with different last ending digits in intervals of equal length, for example, payments with 10-20 endings and those with 80-90 endings. We consider our analysis to be an extension of the results of our reference paper.

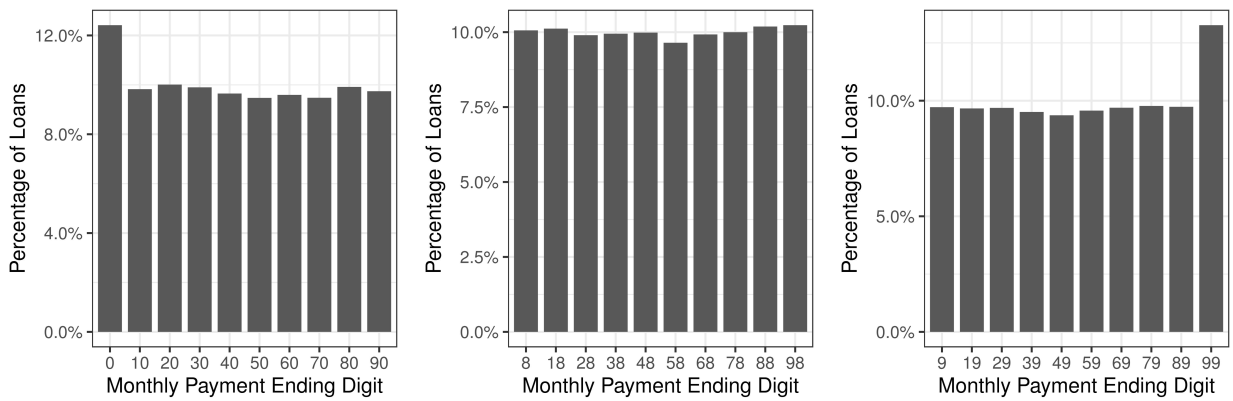


Figure 7: Distribution of loans with the same last ending digit (0, 8 and 9 displayed)

Figure 7 examines if the bunching effect exists for the same last digit in different payment levels. For loans with the last digit ranging from 1 to 8, the distribution is almost uniform across all levels; for those with 0 as the last digit, the bunching occurs at 00 endings; for those with 9 as the last digit, the bunching occurs at 99-endings. This is consistent with what we found with our replication in Section III as 0-endings and 9-endings are where the left-digit bias is displayed. Moreover, it is the 00-ending and 99-ending loans that play a major role in contributing to the bunching effect, which means the bias is much stronger when payment levels go across $100 marks that concerns changes to the left digits. Here it seems that the percentage does not vary among different middle digits, but it will be further examined below.

We wish to be more precise about whether it's only the left-most digit that induces the biases, or does middle digits have an effect as well. This involves further segmentation of the data by whether the monthly payment is over 1000 or not. For loans with payments less than 1000, 99 ending concerns change in the left-most digit, while for loans with payments above 1000, 99 ending only concerns changes in the second-leftmost digit. An additional benefit of selecting 1000 as the benchmark is to separate 3-digit loans and 4-digit loans. The results are shown in Figure 8.

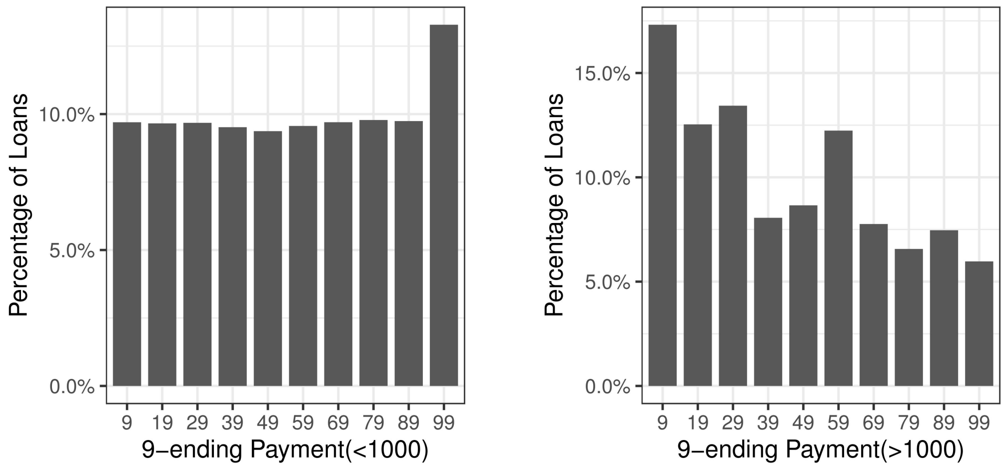


Figure 8: Distribution of loans with 9-ending payments, segmented by $1000

Figure 8 displays the different distribution of loans with 9 as the last digit when compared with the 1000 benchmark. For those loans in the below $1000 section, the distribution is almost identical to the aggregated one in Figure 7. However, for the over 1000 section, the loans with 09-endings become most popular, and the rest of the percentages are not evenly distributed among the other ending digits. We suspect that the reason behind this phenomenon is the scarcity of data and a possible change of pattern. Recall from table1 that the mean and std for payment are 389 and 148 respectively, which means loans with over 1000 payments are located at more than 4 positive standard deviations, so the sample size is rather limited. Another explanation is a change in the loan's distribution pattern when the payment goes across a certain threshold. This is investigated and the results are shown in Figure 9.

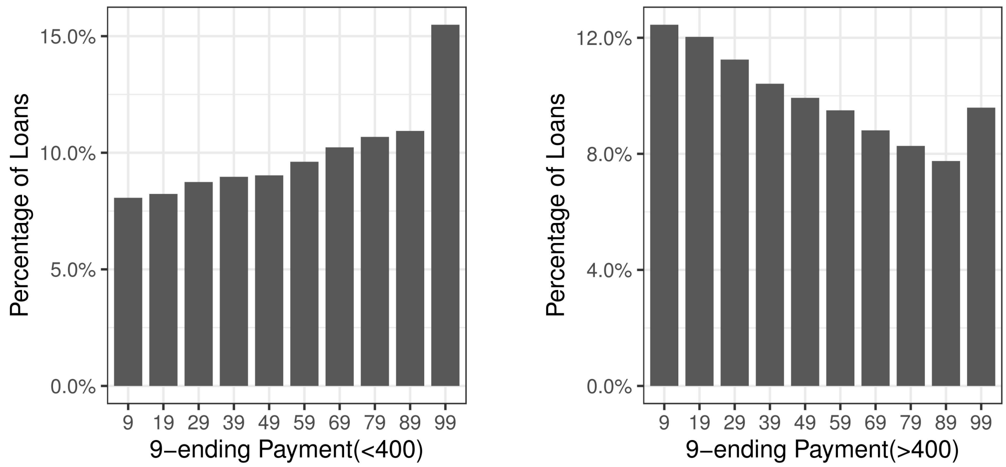


Figure 9: Distribution of loans with 9-ending payments, segmented by $400

Figure 9 displays the different distribution patterns for loans with payment levels over or below 400. We checked for all the possible benchmarks and found that 400 is when the distribution patterns are not only different but also matched with our replication in Section III that there should be a gradual percentage increase or decrease, and at the same time, the 400 mark is close to the mean value of the monthly payments among the synthesis dataset. For payments below 400, 99-endings take up a very considerable proportion, and the other endings have percentages lower than their respective aggregated levels. From 09 to 89-endings, the deficit gradually decreases. For payments over 400, 99-endings only take up a near 10% percentage, whereas 09-endings become popular and the popularity gradually decreases from 09-endings to 89-endings. The gradual decrease of deficit in the below $400 section and gradual decrease of popularity in the over $400 section help balance out the aggregated proportion of 09 to 89-endings, leaving 99-endings as the only dominant distribution in the aggregated situation. It is also important to note that the $400 threshold is very close to the mean of payment in our dataset, so these two distribution patterns also represent cheap loans and expensive loans respectively. We suspect that this difference is due to both the consumers and dealers being subject to left-digit bias. While consumers going for cheap loans have a strong incentive to keep payments at 99-endings, dealers may succeed with lifting payments to 09-endings for expensive loans where consumers may be less price-sensitive. It also shows that middle digits do matter for negotiation outcomes, at least in this setting.

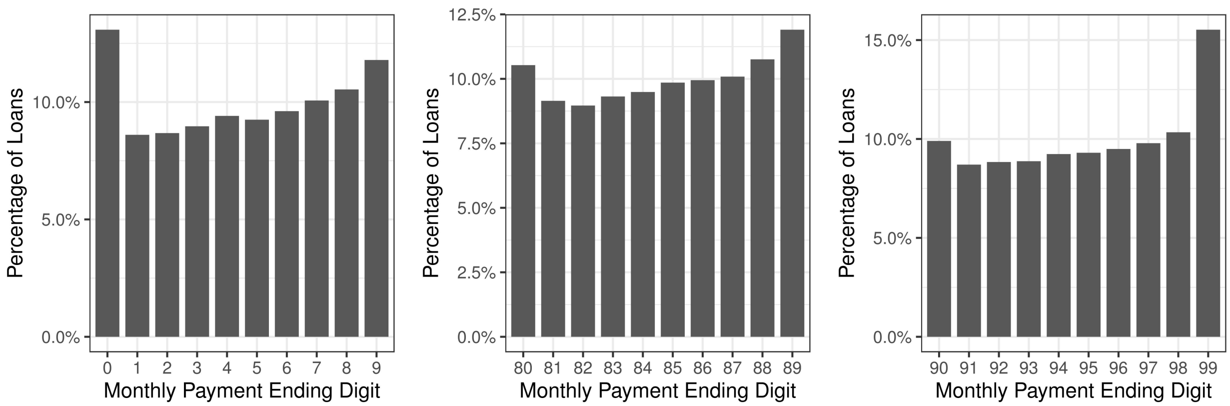


Figure 10: Distribution of loans within the same intervals (00-09, 80-89 and 90-99 displayed)

Figure 10 displays loans whose last two digits belong to the same intervals. We observe that the shape of the distribution is roughly the same for all the intervals, with the middle ones being extremely similar. Again the 00-09 and 90-99 intervals contribute a lot to the high percentage of 99-endings and 00-endings in the aggregate level. The findings are consistent with the replication in section III, and the distribution does not change for different levels of payments.

Section V: Conclusion and Discussion

The main contribution of this paper is twofold. It first replicates the findings of how left-digit bias affects bargaining outcomes in the auto finance market from our reference paper. The proposed model in our reference paper can reflect several phenomena we observe that deviate from what the standard model predicts, and we successfully replicated the results with few differences mainly due to the simulated nature of the synthetic dataset we operate on. Consistent with our reference paper, we find that not only consumers but also dealers suffer from left-digit bias and that dealers are not much better off in terms of the scale of bias. To be specific, the consumer bias leads to the unusually high distribution at 9-ending payments, while the dealers' bias leads to the unusually high distribution at 0-ending payments. The percentage payments of other ending digits are then less than the 10% baseline. Moreover, there is a clear drop in the interest rate of loans with 9-ending and 0-ending digits. The interest rates are influenced by both ending digits and the demographic conditions of the consumers, although the interaction between those terms is less significant in our case compared with the reference paper. It requires both parties to exhibit a left-digits bias for bargaining outcomes from the datasets to occur.

The second main contribution is our extension analysis in Section IV. Based on our replication study, we further segmented the data to check if loans with the same last digit display different patterns at different payment levels. We found that the 99-endings and 00-endings that cross $100 in payment account for most of the bunching effect, and that the distribution pattern is different for loans with payments larger or smaller than the mean of $389. For cheap loans, 99-endings are most dominant and the percentage increases from 09 to 89-endings, while for expensive loans 09-endings are most dominant and the percentage decreases from 09 to 89-endings. For loans that go across the same intervals in different payment levels, we found no systematic difference and the result is in line with our replication.

References

Jiang, Z. (2022). “An empirical bargaining model with left-digit bias: A study on auto loan monthly payments.” *Management Science*, *68*(1), 442–465. https://doi.org/10.1287/mnsc.2020.3923

Olenski, Andrew R., André Zimerman, Stephen Coussens, and Anupam B. Jena. (2020). “Behavioral Heuristics in Coronary-Artery Bypass Graft Surgery.” *New England Journal of Medicine* 382 (8): 778–79. https://doi.org/10.1056/NEJMc1911289.

Olsen, A. L. (2013). “Leftmost-digit-bias in an enumerated public sector? An experiment on citizens' judgment of performance information.” *Judgment and Decision Making*, 8(3), 365–371.

Bhattacharya, Utpal and Holden, Craig W. and Jacobsen, Stacey E. (2012). “Penny Wise, Dollar Foolish: Buy-Sell Imbalances On and Around Round Numbers.” *Management Science* 15, 413-431, 2012. Available at SSRN: https://ssrn.com/abstract=1569922

Lacetera, Nicola, Devin G. Pope, and Justin R. Sydnor. (2012). "Heuristic Thinking and Limited Attention in the Car Market." *American Economic Review*, 102 (5): 2206-36.

Choi, J., Lee, K., Ji, Y. (2012). “What type of framing message is more appropriate with nine-ending pricing?” *Marketing Letters*, 23(3), 603-314. doi:10.1007/s11002-012-9164-7

Manoj Thomas, Vicki Morwitz. (2005). “Penny Wise and Pound Foolish: The Left-Digit Effect in Price Cognition”, *Journal of Consumer Research*, Volume 32, Issue 1, June 2005, Pages 54–64, https://doi.org/10.1086/429600