Nonlinear Pricing under Demand Uncertainty: Transit Passes in Taipei*

Ching-I Huang†

February 22, 2024

Abstract

This study analyzes consumer behavior of the public transit in Taipei. Using smart card usage records, we quantify the causal effect of introducing a monthly pass in April 2018. The estimation result indicates that the policy does increases an individual's transit usage significantly. On average, the value of the increase is 24 – 29 Taiwan dollars per day, depending on the specifications. Moreover, more frequent users have a relatively smaller effect from holding a transit pass. We then construct a structural model to characterize the dynamic decision to purchase a pass, accounting for uncertainty of future transit demand. Based on the structural model, we perform counterfactual simulations to explore alternative pricing policies.

Keywords: transit pass, smart card, public transport, nonlinear pricing

*We would like to thank the Taipei City Department of Transportation for providing the data. We appreciate the suggestions from Mitsuru Igami and seminar participants at Academia Sinica. Financial supports from the Ministry of Science and Technology in Taiwan are gratefully acknowledged. All remaining errors are ours.

†Department of Economics, National Taiwan University and Research Center for Humanities and Social Sciences, Academia Sinica

1 Introduction

Public transport system is an important infrastructure to reduce traffic congestion and lessen environmental pollution in cites. The pricing decision generally involve the consideration of the externality of congestion in private transport and peak loading in the system (Glaister, 1974).

To promote the usage of public transport, many cities offer various transit passes for regular users. A pass holder tends to use public transport more frequently. Nonetheless, to evaluate the effect of these pass programs, it is crucial to control for the self selection issue of buy transit passes. On the one hand, a heavy user has stronger incentive to purchase a pass. On the other hand, since there no extra charge to any additional ride with a valid pass, a pass holder has an incentive to take more rides on the public transport. The aim of the study is to disentangle these two effects.

By offering both single-ride fees and transit passes simultaneously in the public transit system, consumers essentially fact a nonlinear pricing scheme. Traditional economic theories suggest that nonlinear pricing can be use a tool to screen consumers with heterogenous preferences (Wilson, 1993). However, some recent studies indicate that when consumers face the choice of various rate plans, their behaviors are often not fully consistent with the prediction under the rational expectation assumption and propose explanations based on behavior economics. See DellaVigna (2009) for a summary of these findings. I will also construct a structural behavior model to empirically analyze the rationality of consumer behaviors in the context of public transport.

1.1 Literature Review

This paper is related to the broad literature on consumer behavior under nonlinear pricing. In particular, when consumers are offered a menu of contracts, how do they make the choice? In addition to the transit service, similar pricing scheme can be found in many other service industries such as telecom service, health insurance, sport clubs, museums, theme parks, . . . etc. Under such pricing schemes, a consumer needs to make a discrete plan plans, followed by a continuous quantity choice.

Train, McFadden, and Ben-Akiva (1987) estimate a nested logit model in which the plan choice and the quantity choice are simultaneous and both are discrete. Hanemann (1984) proposes a framework to estimate discrete/continuous models in which the discrete plan choice and the continuous quantity choice are linked by the same utility maximization problem. This framework can be used to analyze the simultaneous choices of brand and quantity, but does not account for the time lag between plan choice and volume choice.

Miravete (2002) points out the importance of this time lag. He proposes a theoretical model to justify the nonlinear price schedule commonly observed in the telecom industry. A telephone service company can use a menu of optional calling plans to screen consumers with respect to the ex ante tastes and nonlinear price schedule within each plan to screen them with respect to the interim demand shocks.

My earlier work (Huang, 2008) imposes the rational expectation assumption to analyzing nonlinear pricing in the cellular phone service industry. The primary reason is due to the limitation of data. Without data on each individual consumer, I have to make the rational expectation assumption to infer consumer's preference from the carrier-level data. On the contrary, individual transit usage information is directly observed in the current study.

In contrast to the studies based on the rational expectation assumption, several researches suggest that consumers only have bounded rationality. For example, DellaVigna and Malmendier (2006) study the contract choice in health clubs. Consumers are offered a choice between a flat monthly fee of $70 and a 10-visit pass at $10 per visit. The average attendance number is only 4.3 times per month for those who purchase the flat-fee monthly pass, indicating they could have save money by choosing the other contract. The observed behavior suggests that consumers are overconfident about future attendance rate. As a result, inference based on rational expectation hypothesis could result in biased estimation of consumer preference. In the health insurance market, Handel (2013) shows the effects of inertial on the plan choice and quantifies risk preference.

Lambrecht and Skiera (2006) combine survey data with usage data to investigate the cause of the bias toward a flat-rate pricing scheme for Internet service. They suggest that the potential explanations can be described as (a) insurance effect, (b) taxi meter effect, and (c) overestimation effect. Among these effects, overestimation is the most important factor to explain the bias in plan choice observed in the data.

Similar pattern is also observed in the cellular phone service industry. Grubb (2009) proposes a theoretical model to explain the observed pricing schemes. When consumers may overestimate the precision of their demand forecasts, both monopolists and competitive firms have incentive to offer tariffs with included quantities at zero marginal cost, followed by steep marginal charges. Using a 2002–2004 panel of cellular bills, Grubb and Osborne (2015) empirically analyze the effects on overconfidence. Specifically, they evaluate the FCC regulation which requires carrier to alert consumers when their usage exceeds free allowances. The annual impact on consumer welfare is

$33, indicating that the alerts are informative. They find that consumers are inattentive to past usage and underestimate the variance of future calling.

Additionally, Gathergood et al. (2021) show that consumers make mistakes regarding credit card payment. While automatic payment can guarantee online payment, many consumers activate this service only after they incur an late payment penalty fee. In a recent paper, Bordalo, Gennaioli, and Shleifer (2020) propose a theoretical model to discuss how selective memory and attention to surprise could affect consumer choices.

As for the literature specific to the effects of transit passes, many previous researches use household-level or individual-level survey data to study the demand for public transit. For example, Doxsey (1984) proposes a economic model to characterize transit passes and uses survey data to empirically find out the determinants on the decision to buy a pass. He finds that individual saving is the most important factor on the purchasing decision. Moreover, consumers' response to expected gains and losses from the pass is asymmetric.

Badoe and Yendeti (2007) estimate the effect of transit pass ownership by using survey data in Toronto. They propose a two-step procedure to control for the endogenous decision to own a transit pass. Specifically, transit pass ownership is estimated in the first stage by the binary probit model. They find that owning a transit pass has a significantly positive effect on the number of public transit trips. Although they highlight the importance of owning a transit pass, they does not compute its marginal effect in the paper. Habib and Hasnine1 (2019) also study the demand for transit pass in Toronto with a similar approach. In addition to finding a profound effect of a transit pass on ridership behavior in terms of the daily frequency of transit trips and the total distance travel by transit, they quantify these effects by computing the marginal effects from the estimated econometric model.

Some recent studies rely on data collected from smart card usage in the public transit system. Smart card data provide detailed records on actual transit usage and generally consist of a very large number of observations. Nevertheless, it is not easy to observe a card holder's demographic variables from these smart card data. Zureiqat (2006) use a discrete-continuous model to simultaneously capture the discrete choice on transit pass ownership and the continuous choice on transit usage. He applies the model to study the transit demand in London using Oyster smart card data collected between November 2005 and February 2008. He finds that inertial is an important factor to explain the discrete transit pass choice, indicating the importance of the learning effect. While the big data collected by smart cards become more available for researcher, most previous studies based on smart card data primarily rely on on method in computer science (such as clustering) to provide descriptive analysis of transit pattern. They often abstract away from passenger choice problem. An important goal of this study is to provide economic explanations to those observed transit pattern such that we could provide an causal analyze on the observed pattern.

In a recent paper, Liu et al. (2020) use smart card data from Shizuoka in Japan to analyze the transit pattern. They are able to link the smart card with an anonymous database which consist of the age and gender of card holders. As a result, they demonstrate the heterogeneity of the transit pattern across different demographic groups based on some clustering methods. In particular, females have greater intra-personal variation in the transit pattern than males. The transit patterns are more diverse for elderly people (aged over 65) relative to the younger generation. Egu and Bonnel (2020) combine a clustering method and a similarity metric to explore simultaneously interpersonal and intrapersonal variability of transit usage in Lyon in France. Based on the clustering method on the usage days during the 6-month sample period, they categorize smart card users into three types: consistent users, intermittent users, and low-frequency users. They further categorize consistent users into six groups based on the transit pattern (boarding time, departing time, departing location) and discuss the correlation between transit pattern and demographic status for each group.

2 Background

As the review reported by Kamargianni et al. (2016), the concept of "mobility as a service" (MaaS) has become popular recently. The term MaaS stands for "buying mobility services based on consumer needs instead of buying the means of mobility." Specifically, this concept indicates the provision of integrated and seamless mobility of urban transport.

The monthly pass program in Taipei is based on the concept of MaaS. It was introduced on April 16, 2018.[1] The pass is valid for 30 consecutive days from the first usage at the price of 1,280 Taiwan dollars (TWD). It allows unlimited rides of metro trains and city buses in the Taipei metropolitan area. In addition, it also covers the first 30 minutes of every rental of the public bikes. As a consequent of this integrated program, a pass holder can freely use one of the three transport modes in the Taipei metropolitan area with a single smart card.[2]

Taipei City Government is the main sponsor and proposer of the transit pass. It is a major share holder of the EasyCard Corporation. Therefore, although there were

[1]Before the introduction of the monthly pass, there had been several types of passes targeting tourists in the metro system with durations ranging from one to three days.

[2]Because of a new funding resource from the central government, the pass program in Taipei was substantially expanded on July 1, 2023 to include intercity rail and expand the coverage area to Taoyuan City and Keelung City. The price was also reduced to 1,200 TWD. For the current study, I use data around the introduction of the initial phase of the pass program.

Table 1: Number of Transit Rides in July 2019

	Metro	Bus	Public Bike
EasyCard	58,326,310	60,853,564	2,357,294
iPass	3,980,582	n/a	126,041
iCash	1,905,508	n/a	0
HappyCash	149,782	n/a	0
Rides Paid by Smart Card	64,362,182	n/a	2,483,335
Total Rides (inc. cash...)	66,557,788	66,373,563	2,544,398

Data Sources: Department of Transportation, Taipei City Government

four types of smart cards accepted in the Taipei public transport system (EasyCard, iPass, iCash, and HappyCash) during the research period, the monthly pass could only be purchased as an add-on feature on EasyCard, but not the other three.

3 Data

3.1 Data Source

The smart card data were provided by Taipei City Department of Transportation. The data consist of transaction level records between January 2017 and July 2019 in the three main modes of public transport in Taipei: metro trains, city buses, and public bikes. Specifically, I can observer all records on metro trains and public bikes paying with any of the four smart cards. As for city buses, the data cover all records paying with EasyCard, but not the other smart cards.

Table 1 summarizes the coverage of the data in the last month of the research period. Compared with the aggregate ridership reported in the monthly statistics, the smart card data have covered the vast majority of the transit records. In particular, the data cover 96.7% of MRT rides while the remaining are paid by single-ride tickets, tourist passes, group tickets. Moreover, the data include 97.6% of public bike rentals in Taipei while the remaining rentals are paid through credit cards. Among the four smart cards, EasyCard is by far the most dominant one with a market share of approximately 90% in both metro trains and public bikes. Although we do no have bus records of passengers paying with other smart cards, the share of EasyCard-paying rides is 91.7% among all bus passengers.

To protect the privacy, the card identification code in the research data is a hexadecimal number scrambled from the original identification number. Although we can link the usage records paid with the same smart card, one important limitation is that we can not link the card to an individual user. We cannot observe any socioeconomic variable of card holders. Moreover, although each ride is linked with a unique smart card, a person might use multiple cards in different rides, and a single card might be shared by multiple users in separate rides.

3.2 Data Processing Strategy

To summarize the daily usage pattern, we use a one-dimensional variable as the main outcome variable. "Transit usage" is defined as the total nominal fares of a consumer's transit rides on metro trains, city buses, and public bikes on a given day,[3] taking into account all potential transfer discounts. [4] For a consumer without a pass, her transit usage is exactly the monetary expense on that day. As for a pass-holder, the

[3] Since our focus is the transit service covered by the monthly pass, we only consider the first 30-minute period of each bike ride.

[4] There is an 8-TWD discount for transfers between metro trains and city buses within 80 minutes. Besides, starting from April 1, 2018, there is also a 5-TWD discount for transfer between bike rental and the other two transit modes. Besides, a transfer discount of 8 TWD between two bus rides is limited to "metro bus" and "minibus" routes.

actual monetary payment for her transit usage is zero, and her transit usage equals the money saved from holding the pass. Transit usage is measured in Taiwan dollars (TWD). During the research period, the exchange rate of 1 US dollar is approximately equal to 30 TWD.

We treat each smart card as an independent individual. As Table 1 indicates, there are over 100 million usage records in a typical month. In the current study I only use a subset of the original data for empirical analysis to keep the computational time manageable. The sample is selected based on the first two digits of the scrambled card identification number. Cards starting with “AA” in its hexadecimal identification number are used for the current study. (I need to add a table to show the selected sample is representative.)

This study would focus on the daily usage pattern for each individual card holder. The original data on smart card usage are based on transaction records. Therefore, we need to construct a panel data for the daily usage of each card. Besides, since the majority of monthly pass holders use adult-fare cards, we only include adult-fare cards in the analysis. I will briefly describe the procedure to construct the panel data for our empirical study. More details of the data processing are relegated to the Appendix A.

Because the original data set is based on transaction records, it is straightforward to compute an individual's “transit usage” on a particular day from these records when the daily usage is positive. However, when a card has no record on a particular day, we are not sure whether the card is actually held by some person choosing not to use public transit. It is possible that the card has not been released into circulation or the card has been destroyed. To construct the daily usage panel data for our analysis, we assume that each card has a life span from its first recorded date to the last recorded date in the original data set. If a card has no transaction record on a given day during its life span, we assume that the holder chooses her usage level on that day to be zero. Nonetheless, a card with no transaction record toward both ends of the sample period would be truncated in our constructed panel data. To reduced the truncation problem, we will exclude infrequently used cards from the empirical analysis. We define $freq_i$ to be the frequency of observing a positive usage level during the life span of card i. The full sample of the panel data has 26,942,958 observations with 96,161 unique cards. In the baseline specification, we use the subsample of card with $freq_i \geq 1/7$ (used at least once in a week) for the analysis. It consists of 7,143,004 observations from 36,051 cards.

3.3 Some Observed Patterns

Using the full sample of 96,161 cards, we illustrate the some important patterns in the data. Figure 1 shows the aggregate transit usage during the 31-month research period. Each dot represents the total usage among all card holders in the sample on a particular day. The time trend, shown by the solid curve, is estimated by local polynomial regression. The aggregate usage level was higher on weekdays and lower on weekends. After smoothing out the weekly pattern, there is an increasing trend over time. Moreover, there is some seasonality within a year, with lower usage in early February (lunar New Year holiday), and higher usage in early December. Nonetheless, it is not obvious whether the demand jumped up after implementing the pass policy in April 2018. Next, the number of cards with a valid transit pass is presented in Figure 2. Among the 96,161 cards in our sample, roughly 1% of them had the transit pass attached to the card after the implementation of the policy in April 2018. The adoption rate jumped up in September 2018, and briefly dipped in February 2019. This may

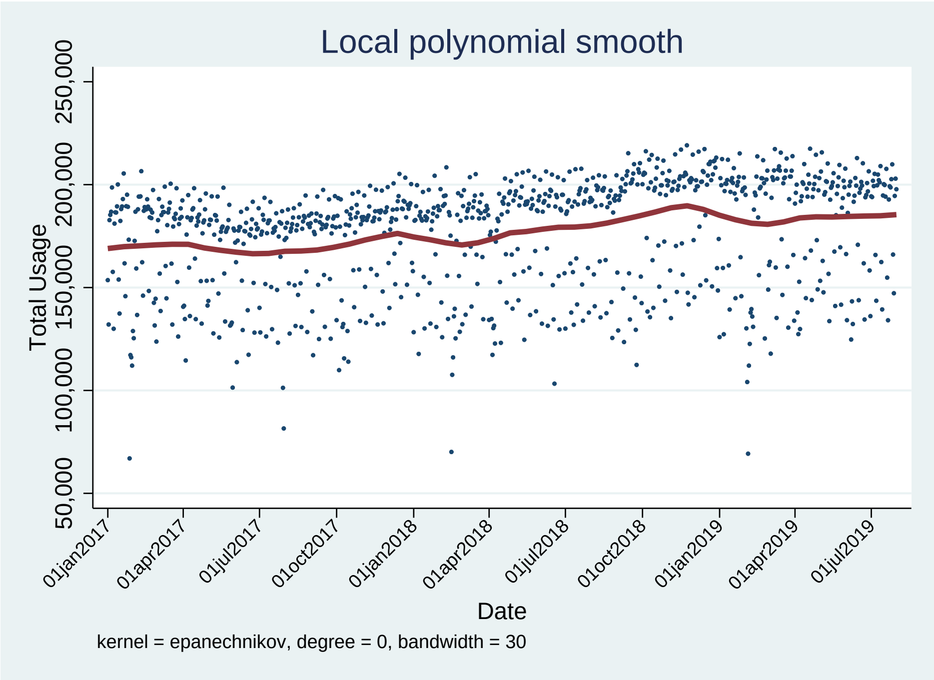

Figure 1: Time Trend of Total Usage

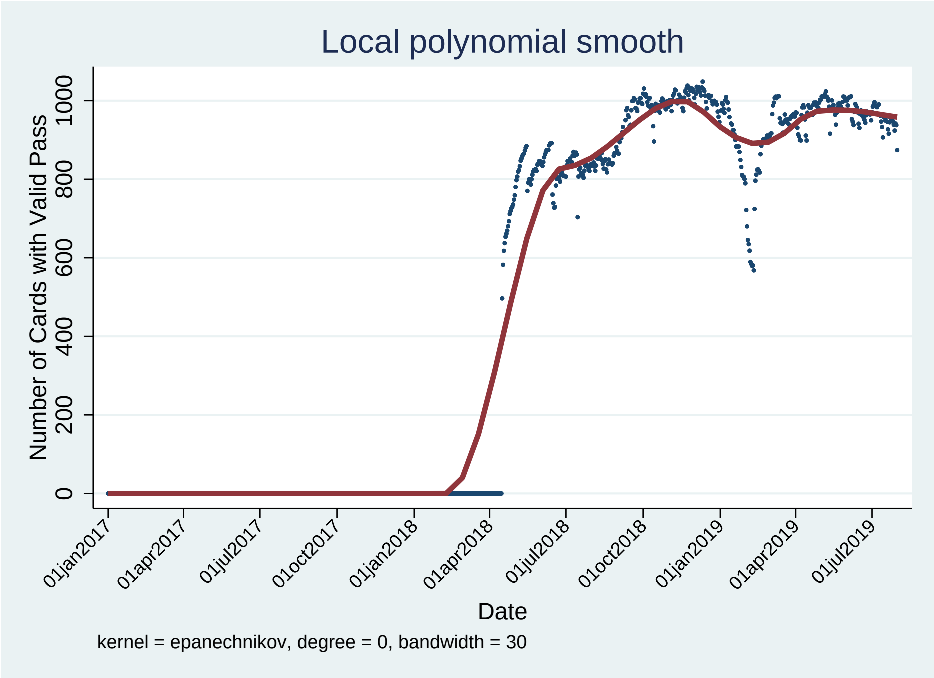

Figure 2: Time Trend of Pass Adoption

Table 2: Activating Time of Transit Passes

Day of Week	Frequency	Percentage
Sunday	1,537	9.22
Monday	3,688	22.12
Tuesday	2,247	13.47
Wednesday	3,091	18.54
Thursday	2,283	13.69
Friday	2,351	14.10
Saturday	1,479	8.87

reflect the seasonality of transit demand, consistent with findings in Figure 1.

As for the pass purchasing decision, a forward-looking consumer is more likely to activate a new transit pass when she expects her demand to be higher in the near future. Table 2 indicates that transit passes are more likely to be activated on weekdays. In particular, Monday has the highest percentage, consistent with higher demand on weekdays. We observed another evidence of forward-looking behavior in Figure 3, which shows the average daily transit usage of all pass holders during the 30-day duration of a pass. Each bar represents the monetary value of the daily transit usage. The average daily usage ranges between 49 and 66 TWD. The highest daily usage occurs on the activating day of a pass. While there is a declining trend of the usage during the 30-day period, we observe a weekly pattern on the daily usage. The observed usage pattern suggests that consumers tend to activate a new transit pass when they expect to have higher demand in the near future. The weekly variation also suggests the transit demand is related to the commuting need on working days.

Figure 4 illustrates the distribution of the actual usage for all pass holders in the 30-day period of a valid pass. The average usage is 1,570 TWD, and the median

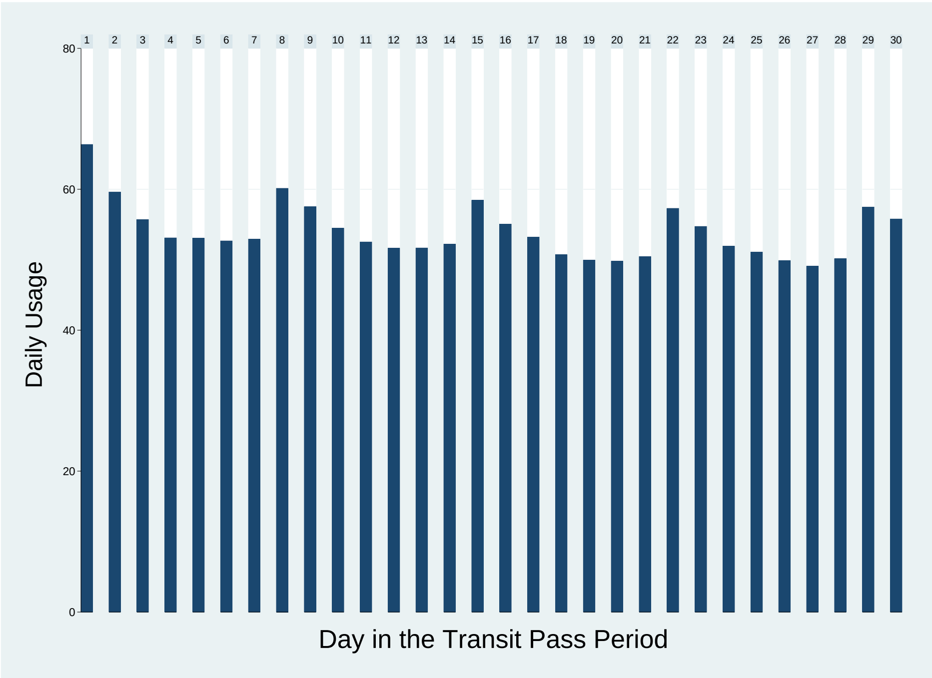

Figure 3: Daily Usage of a Pass Holder during the 30-day Period

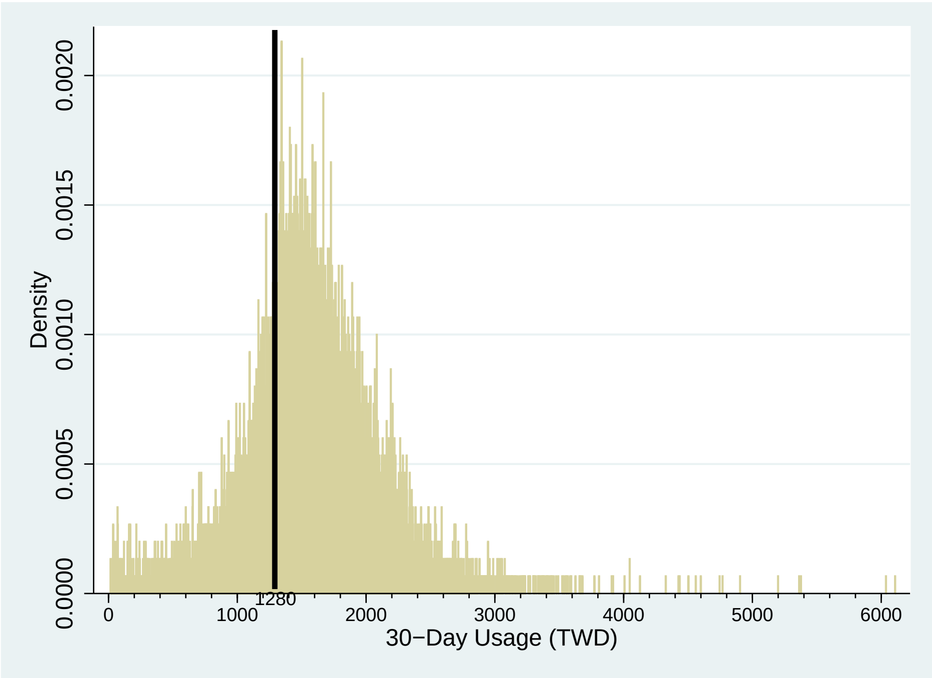

Figure 4: Distribution of Actual Usage among Pass Holders

is 1,549 TWD. The first and the third quartiles are 1,296 TWD and 1,841 TWD, respectively. Out of the 15,014 passes in the sample, 23.55% were used less than 1,280 TWD, indicating these pass holders are not minimizing their expense *ex post*. In this paper, I propose a model with demand uncertainty to explain the puzzle observed in the actual pass usage data.

4 Empirical Analysis of Daily Usage

4.1 Econometric Approach

Let y_{it} denote the daily usage of individual i on day t. It is measured in the monetary value. The variable $HoldPass_{it} \in [0, 1]$ indicates whether this consumer has a valid transit pass on the day. $HoldPass_{it} = 1$ if a pass is valid for the whole day, and $HoldingPass_{it} = 0$ if no pass is valid on that day.[5]

The main regression equation is

$$y_{it} = \alpha HoldPass_{it} + \boldsymbol{\beta}'\mathbf{x}_{it} + f(t) + \xi_i + \varepsilon_{it}. \tag{1}$$

where the vector $\mathbf{x}_{it}$ consists of a constant term and control variables.

As illustrated in the previous section, weekly pattern and seasonality are both important factors in determining the demand for public transit. There are also some longer-term trends over time. Consequently, the control variables include dummy variables for each day of the week, and dummy variables for each month of the year. We also add a polynomial function of the date t, $f(t)$, to control for time trend. More-

[5]For a very small proportion of the data, a consumer holds a pass for a fraction of a day. This is because the consumer might have completed a few rides not covered by a pass before she activated a pass later on that day.

over, we have an individual fixed effect ξ_i for each consumer in some specifications. By adding consumer-level fixed effects, we can control for the selection across consumers. That is, a heavy-user is more likely to purchase the monthly pass. Nonetheless, there still exists selection within an individual over time. A consumer is more likely to buy a new monthly pass when the expected demand is higher in the near future.

Because the introduction of the pass policy is an exogenous event to an individual, we use this policy change to construct an instrumental variable to deal with the endogenous decision of buying a transit pass. The binary variable, $PassAvail_{it}$, equals one for any date t after the introduction of monthly pass on April 16, 2018, and it is zero otherwise. We then use the two-stage least squares (2SLS) method to estimate Equation (1).

The identification assumption is that, except the impact of the pass program, the long-term trend of the transit demand is a smooth function of the date t. Consequently, after using a polynomial function to control for the time trend, the instrument $PassAvail_{it}$ is unrelated to unobserved individual demand shocks ε_{it}. On the other hand, since purchasing a pass is feasible only after the implementation of the program, the pass holding status $HoldPass_{it}$ is positively correlated with the instrument $PassAvail_{it}$.

4.2 Estimation Results

4.2.1 Daily Usage

Table 3 reports the estimated coefficient of $HoldPass_{it}$ for Equation (1) under the baseline specifications. In these specifications, we exclude infrequent users from the sample by requiring $freq_i \geq 1/7$. Hence, a card is on average used at least once per

Table 3: Regression on Transit Usage: Baseline Specifications

	Dep. Var.: Daily Transit Usage Mean: 18.771						
	(A)	(B)	(C)	(D)	(E)	(F)	(G)
$HoldPass$	38.088***	27.128***	2.576*	9.025***	9.543***	28.503***	28.144***
	(0.336)	(0.540)	(1.645)	(1.935)	(1.923)	(1.871)	(1.869)
Day of Week	Y	Y	Y	Y	Y	Y	Y
Month of Year	Y	Y	Y	Y	Y	Y	Y
Card ID dummy	N	Y	Y	Y	Y	Y	Y
Deg. of $f(t)$	none	none	none	1	2	3	4
Estimation Method	OLS	OLS	2SLS	2SLS	2SLS	2SLS	2SLS
Sample Criteria	$freq \geq \frac{1}{7}$	$freq \geq \frac{1}{7}$	$freq \geq \frac{1}{7}$	$freq \geq \frac{1}{7}$	$freq \geq \frac{1}{7}$	$freq \geq \frac{1}{7}$	$freq \geq \frac{1}{7}$
Observations	7,143,004	7,143,004	7,143,004	7,143,004	7,143,004	7,143,004	7,143,004
Unique Card IDs	36,051	36,051	36,051	36,051	36,051	36,051	36,051
R^2	0.1253	0.1231	0.0333	0.0787	0.0828	0.1277	0.1279

Notes: Robust standard errors, clustered at card ID, are given in parentheses. Superscripts ***, **, and *represent significance at 1%, 5%, and 10%, respectively.

week to be included in the estimation.

The OLS result in Columns (A) indicates a positive correlation between holding a valid pass and daily usage. However, the positive relationship could be driven by an individual's endogenous decision to purchase a transit pass. After we add individual fixed effects ξ_i in Column (B), the estimated coefficient substantially drops from 38 TWD to 27 TWD. Moreover, by using 2SLS in Column (C), we find the estimated effect of holding a transit pass on daily usage further reduces to merely 3 TWD.

In Columns (D) to (F) of Table 3, we add a polynomial of date t, $f(t)$ to control for the long-term trend in transit demand. As mentioned in the previous subsection, the instrument $PassAvail_{it}$ would improperly pick up the time trend without this control function. Our preferred specification, listed in Column (F), includes a 3-degree polynomial of t in the regression equation. Adding higher degree polynomial does not

Table 4: Regression on Transit Usage: Alternative Specifications

	Dep. Var.: Daily Transit Usage						
	(A)	(B)	(C)	(D)	(E)	(F)	(G)
HoldPass	28.503***	27.285***	26.747***	24.128***	24.899***	19.741***	12.465***
	(1.871)	(2.011)	(2.008)	(2.055)	(1.661)	(1.632)	(1.941)
Day of Week	Y	Y	Y	Y	Y	Y	Y
Month of Year	Y	Y	Y	Y	Y	Y	Y
Card ID Dummy	Y	Y	Y	Y	Y	Y	Y
Degree of $f(t)$	3	5	6	3	3	3	3
Mean of Dep. Var.	18.771	18.771	18.771	6.183	33.059	38.228	45.487
% Holding Pass	5.94%	5.94%	5.94%	1.61%	13.58%	17.27%	22.61%
Estimation Method	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Sample Criteria	$freq \geq \frac{1}{7}$	$freq \geq \frac{1}{7}$	$freq \geq \frac{1}{7}$	full	$freq \geq \frac{1}{2}$	$freq \geq \frac{2}{3}$	$freq \geq \frac{4}{5}$
Observations	7,143,004	7,143,004	7,143,004	26,942,958	2,449,833	1,358,877	466,057
Unique Card IDs	36,051	36,051	36,051	96,161	23,005	18,675	15,003
R^2	0.1277	0.1274	0.1273	0.1253	0.1522	0.1390	0.0828

Notes: Robust standard errors, clustered at card ID, are given in parentheses. Superscripts ***, **, and *represent significance at 1%, 5%, and 10%, respectively.

have significant impact on the estimated coefficient of $HoldPass_{it}$.

Table 4 compares several alternative specifications. Column (A) replicates our preferred specification. In Columns (B) and (C) we add higher order terms to the polynomial function $f(t)$, and there is little impact on the estimated coefficient. In the remaining columns of Table 4 we change the criteria of the regression sample. In appears that the estimated effect is smaller when we including only highly frequent users. In Column (D) we use the full sample in the estimation. While infrequent users are also included in the estimation, the estimated coefficient on *HoldPass* is slightly smaller. On the other hand, we use higher criteria to select frequent users from Column (E) to (G). In the last column, we only include card used on more than 80% of the days during its life span. We find the estimated effect of transit pass is relatively smaller for more frequent users. To further explore the heterogeneity of the causal

Table 5: Regression on Transit Usage: Alternative Specifications 2

	Dep. Var.: Daily Transit Usage					
	(A)	(B)	(C)	(D)	(E)	(F)
HoldPass	28.503*** (1.871)	36.273*** (3.138)	40.173*** (8.041)	39.187*** (8.070)	35.717*** (9.120)	25.402*** (2.116)
Day of Week	Y	Y	Y	Y	Y	Y
Month of Year	Y	Y	Y	Y	Y	Y
Card ID dummy	Y	Y	Y	Y	Y	Y
Degree of $f(t)$	3	3	3	4	5	3
Mean of Dep. Var.	18.771	14.200	8.874	8.874	8.874	1.839
% Holding Pass	5.94%	3.28%	1.07%	1.07%	1.07%	0.48%
Estimation Method	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Sample Criteria	$freq \geq \frac{1}{7}$	$\frac{1}{7} \leq freq < \frac{2}{3}$	$\frac{1}{7} \leq freq < \frac{1}{3}$	$\frac{1}{7} \leq freq < \frac{1}{3}$	$\frac{1}{7} \leq freq < \frac{1}{3}$	expanded
Observations	7,143,004	5,784,127	3,332,381	3,332,381	3,332,381	90,583,662
Unique Card IDs	36,051	17,367	8,942	8,942	8,942	96,161
R^2	0.1277	0.0991	0.0549	0.0552	0.0549	0.1301

Notes: Robust standard errors, clustered at card ID, are given in parentheses. Superscripts ***, **, and *represent significance at 1%, 5%, and 10%, respectively.

effect of transit pass, Table 5 shows the estimation results for $freq_i \in [1/7, 1/3)$ and $freq_i \in [1/7, 2/3)$. These users appear to have a higher effect than those in Columns (E), (F) of the previous table.

Table 6 compares the estimation results using different time span of the data. Column (A) uses the full duration in the original data. Column (B) drops the first and last month to avoid the truncation problem of non-used cards. Column (C) uses data within one year of the transit pass introduction month (April 2018), and the Columns (D) and (e) further restrict to three months and one month within the introduction month, respectively. The estimated effect of $HoldPass_{it}$ is similar from Column (A) to (D). The results in (E) and (F) are probably less reliable because we cannot properly control for the long-term time trend when using a short time span of the data.

Table 6: Regression on Transit Usage: Alternative Specifications 3

	Dep. Var.: Daily Transit Usage					
	(A)	(B)	(C)	(D)	(E)	(F)
HoldPass	28.503*** (1.871)	26.087*** (2.079)	30.107*** (2.038)	29.814*** (1.778)	23.610*** (1.909)	45.824*** (2.831)
Day of Week	Y	Y	Y	Y	Y	Y
Month of Year	Y	Y	Y	N	N	N
Card ID Dummy	Y	Y	Y	Y	Y	Y
Degree of $f(t)$	3	3	3	3	3	3
Mean of Dep. Var.	18.771	18.562	18.362	18.047	17.712	18.100
% Holding Pass	5.94%	5.85%	5.71%	5.00%	5.38%	4.73%
Estimation Method	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Sample Criteria	$freq \geq \frac{1}{7}$	$freq \geq \frac{1}{7}$	$freq \geq \frac{1}{7}$	$freq \geq \frac{1}{7}$	$freq \geq \frac{1}{7}$	$freq \geq \frac{1}{7}$
Sample Period	17/01-19/07	17/02-19/06	17/04-19/04	17/10-18/10	18/01-18/07	18/03-18/05
Observations	7,143,004	6,765,996	5,905,094	3,134,905	1,682,113	728,831
Unique Card IDs	36,051	33,975	30,242	19,289	13,991	10,477
R^2	0.1277	0.1251	0.1260	0.1234	0.1127	0.1154

Notes: Robust standard errors, clustered at card ID, are given in parentheses. Superscripts ***, **, and *represent significance at 1%, 5%, and 10%, respectively.

Table 7: Regression on Daily Rides

	Dep. Var.: Daily Rides			
	All	MRT	Bus	Bike
HoldPass	1.653***	0.827***	0.769***	0.057***
	(0.099)	(0.064)	(0.058)	(0.013)
Day of Week	Y	Y	Y	Y
Month of Year	Y	Y	Y	Y
Card ID dummy	Y	Y	Y	Y
Deg. of $f(t)$	3	3	3	3
Mean of Dep. Var.	1.004	0.583	0.400	0.021
Estimation Method	2SLS	2SLS	2SLS	2SLS
Sample Criteria	$freq \geq \frac{1}{7}$	$freq \geq \frac{1}{7}$	$freq \geq \frac{1}{7}$	$freq \geq \frac{1}{7}$
Observations	7,143,004	7,143,004	7,143,004	7,143,004
Unique Card IDs	36,051	36,051	36,051	36,051
R^2	0.1074	0.0722	0.0534	0.0005

Notes: Robust standard errors, clustered at card ID, are given in parentheses. Superscripts ***, **, and *represent significance at 1%, 5%, and 10%, respectively.

4.2.2 Effects on Each Transport Mode

Instead of analyzing the effect of transit passes on the total transit usage, Table 7 shows the effect on the number of daily ridership, and further decomposes the effect on the three transportation modes. Most of the increase in transit usage comes from MRT rides.

4.2.3 Heterogeneity of the Effects

We separately estimate the effect of transit passes for weekdays and for weekends, respectively. Table 8 indicates that the effect on the transit usage is higher on weekdays, both in terms of the absolute amount and the amount relative to the mean.

We also estimate the effect on ridership for each transportation mode on weekdays

Table 8: Regression on Transit Usage: Weekdays and Weekends

	Dep. Var.: Daily Transit Usage		
	All	Weekday	Weekend
HoldPass	28.503*** (1.871)	36.326*** (2.231)	11.704*** (1.917)
Day of Week	Y	Y	Y
Month of Year	Y	Y	Y
Card ID dummy	Y	Y	Y
Deg. of $f(t)$	3	3	3
Mean of Dep. Var.	18.771	20.883	13.496
% Holding Pass	5.94%	5.99%	5.81%
Estimation Method	2SLS	2SLS	2SLS
Sample Criteria	$freq \geq \frac{1}{7}$	$freq \geq \frac{1}{7}$	$freq \geq \frac{1}{7}$
Observations	7,143,004	5,101,025	2,041,979
Unique Card IDs	36,051	34,680	31,269
R^2	0.1277	0.1497	0.0442

Notes: Robust standard errors, clustered at card ID, are given in parentheses. Superscripts ***, **, and *represent significance at 1%, 5%, and 10%, respectively.

Table 9: Regression on Daily Rides: Weekdays

	Dep. Var.: Daily Rides on Weekdays			
	All	MRT	Bus	Bike
$HoldPass$	2.064*** (0.119)	1.034*** (0.075)	0.959*** (0.070)	0.071*** (0.015)
Day of Week	Y	Y	Y	Y
Month of Year	Y	Y	Y	Y
Card ID dummy	Y	Y	Y	Y
Deg. of $f(t)$	3	3	3	3
Mean of Dep. Var.	1.115	0.644	0.449	0.022
Estimation Method	2SLS	2SLS	2SLS	2SLS
Sample Criteria	$freq \geq \frac{1}{7}$	$freq \geq \frac{1}{7}$	$freq \geq \frac{1}{7}$	$freq \geq \frac{1}{7}$
Observations	5,101,025	5,101,025	5,101,025	5,101,025
Unique Card IDs	34,680	34,680	34,680	34,680
R^2	0.1177	0.0815	0.0543	0.0005

Notes: Robust standard errors, clustered at card ID, are given in parentheses. Superscripts ***, **, and *represent significance at 1%, 5%, and 10%, respectively.

and weekends, respectively. The estimated effects are shown in Tables 9 and 10. We summarize the estimated effect relative to its mean value by the bars shown in Figure 5. For all transportation modes, the impact of transit passes is higher on weekdays is higher in both the absolute and the relative magnitude. On the other hand, the relative impact is the highest for public bike, even although it has the lowest mean daily ridership. This suggest that the demand of public bike is more responsive to the pass policy.

Table 10: Regression on Daily Rides: Weekends

	Dep. Var.: Daily Rides on Weekends			
	All	MRT	Bus	Bike
HoldPass	0.767*** (0.102)	0.381*** (0.068)	0.362*** (0.058)	0.024* (0.014)
Day of Week	Y	Y	Y	Y
Month of Year	Y	Y	Y	Y
Card ID dummy	Y	Y	Y	Y
Deg. of $f(t)$	3	3	3	3
Mean of Dep. Var.	0.725	0.430	0.278	0.017
Estimation Method	2SLS	2SLS	2SLS	2SLS
Sample Criteria	$freq \geq \frac{1}{7}$	$freq \geq \frac{1}{7}$	$freq \geq \frac{1}{7}$	$freq \geq \frac{1}{7}$
Observations	2,041,979	2,041,979	2,041,979	2,041,979
Unique Card IDs	31,269	31,269	31,269	31,269
R^2	0.0422	0.0263	0.0254	0.0003

Notes: Robust standard errors, clustered at card ID, are given in parentheses. Superscripts ***, **, and *represent significance at 1%, 5%, and 10%, respectively.

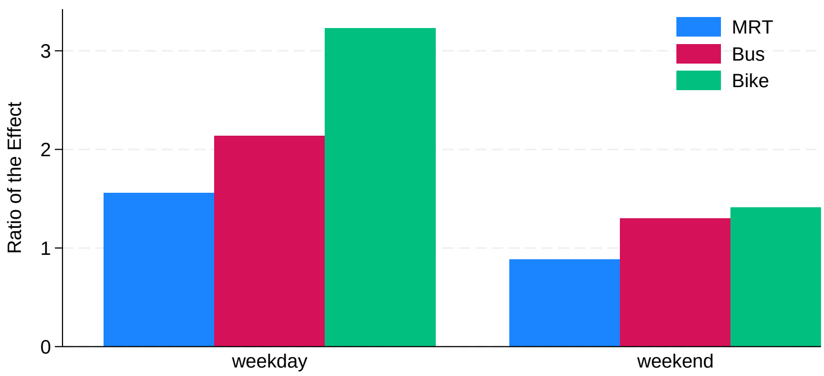

Figure 5: The Estimated Effect Relative to the Mean

5 The Pass-Purchasing Decision

The previous section has shown the effects of transit passes on the actual usage. Our next step is to investigate the decision to purchase a transit pass. In this section, we propose a structural model to analyzed the decision to purchase a transit pass. As illustrated in Subsection 3.3, an individual's decision would depend on the expected usage in the coming days. Therefore, we propose a dynamic model to characterize this decision.

5.1 A Structural Dynamic Model

An individual makes a binary decision on the purchase of the T-day transit pass every day $t = 0, 1, 2 \ldots$. A pass is valid for T consecutive days. The monetary cost of a pass is A. There is some non-monetary benefit of pay-as-you-go relative to holding a transit pass, denoted as η_{it}. Assume η_{it} is i.i.d across individuals i and over time t. For the transit pass in Taipei, $T = 30$ and $A = 1280$.

Individual i's taste for public transportation on date t is characterized by a nonnegative index θ_{it}. It is drawn from a distribution function $F(\cdot; x_{it})$ where x_{it} is a vector of factors (such as days of the week, holiday) observed before the date. The stochastic distribution captures unforeseeable taste shocks realized on date t (such as weather). Assume that θ_{it} is independent over time conditional on x_{it} and also independent of η_{it}. The value of θ_{it} is realized after making the binary decision on purchasing a transit pass. Let q_{it} denote individual i's actual usage of public transportation on date t, measured by the total nominal fares of all the rides taking place on date t.

The time line of events on a given day t is the following:[6]

[6]An alternative model would allow an individual to purchase a new pass after observing the realized

- The non-monetary benefit of pay-as-you-go η_{it} and the state variables x_{it} are observed.

- Unless holding a valid pass for the current day, the individual decides whether to buy a new T-day transit pass.

- The daily transportation taste θ_{it} is observed.

- The individual chooses the transit usage q_{it} for the day.

- A transit pass expires at the end of the T-th day.

Denote consumer i's monetary payment on date t as Pay_{it}. Let $\phi_{it} \in \{0, 1\}$ denote the binary decision to purchase a new transit pass on date t. Her lifetime utility is[7]

$$\sum_{t=0}^{\infty} \delta^t \left[q_{it} \left(1 - \frac{q_{it}}{2\theta_{it}} \right) - \alpha Pay_{it} + (1 - \phi_{it})\eta_{it} \right] \tag{2}$$

where $\delta \in [0, 1)$ is the daily discount factor[8], and $\alpha \in (0, 1)$ captures the marginal disutility of monetary payment. When a pass has been purchased within past $(T - 1)$ days, no new pass is purchased. $\phi_{it} = 0$ whenever $\phi_{i\tau} = 1$ for any $\tau \in [t - T + 1, t - 1]$. On the other hand, a consumer without a valid pass would decide whether to purchase

value of θ_{it}. It seems more difficult to estimate.

[7]The functional form assumption is motivated by Grubb and Osborne (2015) for tractability. This utility form assumption implies the marginal utility of usage q is $1 - q/\theta$. In other words, we normalize the marginal utility from the first unit of usage to one.

[8]In addition to time preference, the discount factor can also captures possibility of losing the smartcard.

a new transit pass. Consequently, there are three possible cases for the payment Pay_{it}:

$$Pay_{it} = \begin{cases} 0, & \text{if } \sum_{\tau=t-T+1}^{t-1} \phi_{i\tau} = 1 \text{ (an exising pass);} \\ q_{it}, & \text{if } \sum_{\tau=t-T+1}^{t} \phi_{i\tau} = 0 \text{ (pay-as-you-go);} \\ A, & \text{if } \phi_{it} = 1 \text{ (a new pass).} \end{cases} \tag{3}$$

Let $v(\kappa, x_{it})$ denote the value function at the beginning of a day (before the realization of η_{it} and θ_{it}) for an individual with state variables x_{it} and holding a pass with $\kappa \in \{0, 1, 2, \ldots, T-1\}$ remaining days.

$$v(\kappa, x_{it}) \equiv E\left[\sum_{\tau=t}^{\infty} \delta^{\tau-t} \max_{q_{it},\phi_{it}} \left\{ q_{i\tau}\left(1 - \frac{q_{i\tau}}{2\theta_{i\tau}}\right) - \alpha Pay_{i\tau} + (1-\phi_{i\tau})\eta_{i\tau} \right\} \middle| \kappa, x_{it} \right]. \tag{4}$$

Suppose that the individual has a valid pass for the day. The transit consumption problem is a static decision:

$$\max_{q_{it}} q_{it}\left(1 - \frac{q_{it}}{2\theta_{it}}\right).$$

By the first order condition, the optimal transit usage is

$$q^*_{it} = \theta_{it}. \tag{5}$$

Plugging into (4), the value function can be written as

$$v(\kappa, x_{it}) = \frac{E[\theta_{it}|x_{it}]}{2} + \delta E[v(\kappa-1, x_{i,t+1})|x_{it}] \qquad \forall \kappa > 0.$$

By induction, for any $\kappa > 0$

$$v(\kappa, x_{it}) = \frac{\sum_{\tau=0}^{\kappa-1} \delta^{\tau} E[\theta_{t+\tau}|x_{it}]}{2} + \delta^{\kappa} E[v(0, x_{i,t+\kappa})|x_{it}]. \tag{6}$$

When $\kappa = 0$, the consumer makes a binary decision ϕ_{it} on whether to purchase a new T-day pass before learning her realized transit demand θ_{it}. Without a pass, the total fare payment is q. The static transit usage problem is

$$\max_q q\left(1 - \frac{q}{2\theta_{it}}\right) - \alpha q.$$

The first order condition implies

$$q^* = (1-\alpha)\theta_{it}, \tag{7}$$

and hence the expected value of not buying a pass is

$$\frac{(1-\alpha)^2 E[\theta_{it}|x_{it}]}{2} + \eta_{it} + \delta E[v(0, x_{i,t+1})|x_{it}].$$

On the other hand, the optimal usage would be $q^*_{it} = \theta_{it}$ if she chooses to buy a new transit pass. The expected value of buying a transit pass is

$$\frac{\sum_{\tau=0}^{T-1} \delta^\tau E[\theta_{i,t+\tau}|x_{,t}]}{2} + \delta^T E[v(0, x_{i,t+T})|x_{it}] - \alpha A,$$

where the first two terms are obtained by setting $\kappa = T$ in Equation (6) and the last term captures the cost of purchasing a pass. To simplify the notations, let

$$S_0(x_{it}) \equiv \frac{(1-\alpha)^2 E[\theta_{it}|x_{it}]}{2} + \delta E[v(0, x_{i,t+1})|x_{it}].$$

and

$$S_1(x_{it}) \equiv \frac{\sum_{\tau=0}^{T-1} \delta^\tau E[\theta_{i,t+\tau}|x_{it}]}{2} + \delta^T E[v(0, x_{i,t+T})|x_{it}] - \alpha A.$$

denote the expected value (net of the non-monetary benefit η_{it}) of the two options, respectively.

The value function for an individual without a valid pass at the beginning of a period is

$$v(0, x_{it}) = E\left[\max\{S_0(x_{it}) + \eta_{it}, S_1(x_{it})\}\right]. \tag{8}$$

We can obtain the following recursive formula for the value function.

$$\begin{aligned} v(0, x_{it}) = {} & \Pr(\phi_{it} = 0|x_{it}) \left[\frac{(1-\alpha)^2 E[\theta_{it}|x_{it}]}{2} + \delta E[v(0, x_{i,t+1})|x_{it}] + E[\eta_{it}|\phi_{it}\!=\!0, x_{it}]\right] \\ & + \Pr(\phi_{it} = 1|x_{it}) \left[\frac{\sum_{\tau=0}^{T-1} \delta^\tau E[\theta_{i,t+\tau}|x_{it}]}{2} + \delta^T E[v(0, x_{i,t+T})|x_{it}] - \alpha A\right]. \end{aligned}$$

Based on the approach proposed by Aguirregabiria and Mira (2002), the recursive formula can be written in a vector form.

$$\mathbf{v}_0 = \mathbf{p}_0 * \left(\frac{(1-\alpha)^2 E[\boldsymbol{\theta}]}{2} + \delta \mathbf{M}_x \mathbf{v}_0 + E^{\phi=0}[\boldsymbol{\eta}]\right) + \mathbf{p}_1 * \left(\frac{\sum_{\tau=0}^{T-1} \delta^\tau \mathbf{M}_x^\tau E[\boldsymbol{\theta}]}{2} + \delta^T \mathbf{M}_x^T \mathbf{v}_0 - \alpha A\right) \tag{9}$$

where the value function $v(0, x_{it})$ and the conditional choice probabilities $\Pr(\phi_{it}|x_{it})$ for all possible states x_{it} are stacked into vectors $\mathbf{v}_0$, $\mathbf{p}_0$, and $\mathbf{p}_1$, respectively; $*$ represents element-by-element product; $\mathbf{M}_x$ is the transition matrix of the states x_{it} with the (a, b) cell in the matrix representing the probability of moving from a-th to b-th state; with slightly abusing the notation, A also denotes a column vector of the value A in each element; $E^{\phi=1}[\boldsymbol{\eta}]$ denotes the vector for the truncated means $E[\eta_{it}|\phi_{it}\!=\!1, x_{it}]$ for each state x_{it}. Rearranging the terms, we obtain

$$\mathbf{v}_0 = \left(\mathbf{I} - \delta \mathbf{p}_0 * \mathbf{M}_x - \delta^T \mathbf{p}_1 * \mathbf{M}_x^T\right)^{-1} \left[\mathbf{p}_0 * \left(\frac{(1-\alpha)^2 E[\boldsymbol{\theta}]}{2} + E^{\phi=0}[\boldsymbol{\eta}]\right) + \mathbf{p}_1 * \left(\frac{\sum_{\tau=0}^{T-1} \delta^\tau \mathbf{M}_x^\tau E[\boldsymbol{\theta}]}{2} - \alpha A\right)\right] \tag{10}$$

where $\mathbf{I}$ is the identity matrix.

The truncated mean $E^{\phi=0}[\boldsymbol{\eta}]$ can be computed from the conditional choice probabilities $\Pr(\eta_{it} = 0|x_{it})$.

$$\begin{aligned} E[\eta|\phi_{it} = 0, x_{it}] &= E[\eta|S_0(x_{it}) + \eta_{it} > S_1(x_{it}), x_{it}] \\ &= E[\eta|\eta > S_1(x_{it}) - S_0(x_{it})] \\ &= E\left[\eta|\eta > F_\eta^{-1}\left(\Pr(\phi_{it} = 1|x_{it})\right)\right] \end{aligned} \tag{11}$$

where F_η is the CDF of η_{it}. When η is normally distributed as $N(\bar{\eta}, \sigma_\eta^2)$, the truncated mean can be computed from the inverse Mills ratio, and it is a linear function of the parameters $(\bar{\eta}, \sigma_\eta)$.

$$E[\eta|\phi_{it} = 0, x_{it}] = \bar{\eta} + \sigma_\eta \frac{\phi\left(\frac{F_\eta^{-1}(\Pr(\eta_{it}=1|x_{it})) - \bar{\eta}}{\sigma_\eta}\right)}{1 - \Phi\left(\frac{F_\eta^{-1}(\Pr(\eta_{it}=1|x_{it})) - \bar{\eta}}{\sigma_\eta}\right)} = \bar{\eta} + \sigma_\eta \frac{\phi(\Phi^{-1}(\Pr(\eta_{it} = 1|x_{it})))}{1 - \Pr(\eta_{it} = 1|x_{it})}.$$

[Maybe we can normalize $\sigma_\eta = 1$ by allowing a coefficient in the first term of the utility function (2)].

5.2 Econometric Implementation

The key component in the value function is $E[\theta_{i\tau}|x_{it}]$. Since $E[\theta_{i\tau}|x_{it}, \kappa > 0] = E[q_{i\tau}|x_{it}, \kappa > 0]$, we could identify this conditional expectation from data directly. Similarly, for consumers without a valid pass $E[q_{i\tau}|x_{i\tau}, \kappa = 0] = (1 - \alpha)E[(\theta_{i\tau}|x_{i\tau}, \kappa = 0]$. Heuristically, α can be estimated from comparing these two results.

Assume that state variables x_{it} consist of two parts, $x_{it} \equiv (\xi_{it}, \xi_{it-1}, \ldots, \xi_{it-6}, \omega_{it})$: demand level within the previous six days $(\xi_{it}, \xi_{it-1}, \ldots, \xi_{it-6})$ and day of week ω_{it}. Demand level on day t is characterized by a discrete variable, $\xi_{it} \in \{H, L, Z\}$ (high, low, or zero), which evolves as a Markov process of degree seven. The day of week $\omega_{it} \in \{Su, Mo, Tu, We, Th, Fr, Sa\}$ is a deterministic process. (Holiday effects is ignored in the current version.) There are $3^7 \times 7 = 15309$ states.

Discretize the usage data to define the demand level ξ_{it}. Assume that the distribution of θ_{it} conditional on $\xi_{it} = Z$ is degenerated at zero; the distribution of θ_{it} conditional on $\xi_{it} = L$ is bounded on $(0, q_m]$, and the distribution of θ_{it} conditional on $\xi_{it} = H$ is on (q_m, ∞). I choose q_m to the median of nonzero usage amount in the data before the introduction of the transit pass, which equals to 40 TWD.

Pin down the value of α from some reduced-form approach. Then, the remaining parameters are the discount factor δ and parameters associated with the distribution of η_{it} (such as $\bar{\eta}$ and σ_η^2 under a normal distribution assumption).

We can use a two-step approach to estimate the parameters in the model. In the first step, estimate the conditional choice probabilities $\mathbf{p}_0$, $\mathbf{p}_1$, and the transition probability $\mathbf{M}_x$ nonparametrically from the data.

1. Estimate α by IV regression of $\log q_{it}$ on a pass dummy, lagged values of $\log q_{it}$ (up to 7 days) and dummy variables for days of the week. By Equations (7) and (5), the coefficient on the pass dummy is $1/(1-\alpha)$, which is estimated as 1.1036. Therefore, $\hat{\alpha} = 0.0939$.

2. Discretize the usage variable. For individuals without a pass,

$$\xi_{it} = \begin{cases} Z, & \text{if } q_{it} = 0; \\ L, & \text{if } 0 < q_{it} \le 40; \\ H, & \text{if } 40 \le q_{it}. \end{cases}$$

   For individuals with a valid pass, the cutoff values are multiplied by the factor $1/(1-\hat{\alpha}) = 1.1036$.

$$\xi_{it} = \begin{cases} Z, & \text{if } q_{it} = 0; \\ L, & \text{if } 0 < q_{it} \le 44; \\ H, & \text{if } 44 \le q_{it}. \end{cases}$$

3. Estimate the transition probability $\mathbf{M}_x$ by the frequency estimator for each $x = (\xi, \omega)$.

4. Estimate the conditional choice probability $\mathbf{P}_1$ by the frequency estimator for each $x = (\xi, \omega)$.

5. Compute the truncated mean $E^{\phi=1}[\boldsymbol{\eta}]$ using equation (11).

6. Using (10) to compute the value function.

7. Express the choice probability as a function of the remaining parameters, $\bar{\eta}$ and

Table 11: Parameters Estimated in the Structural Model

	(A)
α	0.0939
$\bar{\eta}$	1.891
σ_η	46.396
Observations	11,701,775

σ_η.[9]

$$\mathbf{p}_1 = \Phi\left(\frac{\mathbf{S}_1 - \mathbf{S}_0 - \bar{\eta}}{\sigma_\eta}\right),$$
$$\mathbf{S}_1 = \frac{\sum_{\tau=0}^{T-1} \delta^\tau \mathbf{M}_x^\tau E[\boldsymbol{\theta}]}{2} + \delta^T \mathbf{M}_x^T \mathbf{v}_0 - \alpha A$$
$$\mathbf{S}_0 = \frac{(1-\alpha)^2 E[\boldsymbol{\theta}]}{2} + \delta \mathbf{M}_x \mathbf{v}_0$$

8. Apply MLE to the choice probabilities $\Pr(\phi_{it}|x_{it})$ for all the instances to make the binary pass-purchasing decision.[10]

5.3 Estimation Results

Table 11 presents the estimation results in the second stage.

[9] After substituting $\mathbf{v}_0$, the difference of the surpluses between the two choices $\mathbf{S}_1 - \mathbf{S}_0$ is a linear function of $\bar{\eta}$ and σ_η. The computational burden is simply to calculate the choice probability $\Pr(\phi_{it} = 1|x_{it})$ for each possible element x_{it} in the state space.

[10] When computing the sum of log-likelihoods, we can combine all instances of the same value in the state variables x_{it} together.

6 Discussions

6.1 Cost Minimization

When α is close to zero, the usage is not boosted by a pass. In this case, the utility maximization problem is equivalent to cost minimization.

Note that $S_0(x_{it})$ can be written as

$$S_0(x_{it}) = \frac{E[\theta_{it}|x_{it}]}{2} - \alpha E[\theta_{it}|x_{it}] + \frac{\alpha^2 E[\theta_{it}|x_{it}]}{2} + \delta E[v(0, xi, t+1)|x_{it}].$$

Since $\frac{E[\theta_{it}|x_{it}]}{2}$ appears in each time period in both $S_0(x_{it})$ and $S_1(x_{it})$, they cancel out in this binary choice problem. Besides, $\alpha^2 \ll \alpha$ when $\alpha \to 0$, which means the third term in the above equation can be ignored. Hence, when *alpha* is small, the binary pass-purchasing problem can be interpreted as the choice between a lump-sum payment of A for a T-day transit pass and the expected fare payment of $E[\theta_{it}|x_{it}]$ on the current day while deferring the pass-purchasing decision to the next day.

7 Conclusion

This study analyzes consumer behavior of the public transit in Taipei. Using the records from smart card usage, we quantify the causal effect of introducing a monthly pass in April 2018. Reduced-form estimation indicates the flat-rate monthly pass causes an increase in transit usage by roughly 24–29 TWD per day. We construct a structural model to characterize the dynamic decision to purchase a pass, accounting for uncertainty of future transit demand. Based on the structural model, we perform counterfactual simulations to explore alternative pricing policies.

References

Aguirregabiria, Victor and Pedro Mira (2002), "Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models," *Econometrica*, 70, 1519–1543.

Badoe, D. A. and M. K. Yendeti (2007), "Impact of Transit-Pass Ownership on Daily Number of Trips Made by Urban Public Transit," *Journal of Urban Planning and Development*, 133(4), 242–249.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2020), "Memory, Attention, and Choice," *Quarterly Journal of Economics*, 135(3), 1399–1442.

DellaVigna, Stefano (2009), "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature,* 47(2), 315–372.

DellaVigna, Stefano and Ulrike Malmendier (2006), "Paying Not to Go to the Gym," *American Economic Review,* 96(3), 694–719.

Doxsey, Lawrence B. (1984), "Demand for Unlimited Use Transit Passes," *Journal of Transport Economics and Policy*, 18(1), 7–22.

Egu, Oscar and Patrick Bonnel (2020), "Investigating day-to-day variability of transit usage on a multimonth scale with smart card data. A case study in Lyon," *Travel Behaviour and Society*, 19, 112–123.

Gathergood, John, Hiroaki Sakaguchi, Neil Stewart, and Jörg Weber (2021), "How Do Consumers Avoid Penalty Fees? Evidence from Credit Cards," *Management Science*, 67(4), 2562–2578.

Glaister, Stephen (1974), "Generalised Consumer Surplus and Public Transport Pricing," *Economic Journal*, 84(336), 849–867.

Grubb, Michael D. (2009), "Selling to Overconfident Consumers," *American Economic Review*, 99(5), 1770–1807.

Grubb, Michael D. and Matthew Osborne (2015), "Cellular Service Demand: Biased Beliefs, Learning, and Bill Shock," *American Economic Review*, 105(1), 234–271.

Habib, Khandker Nurul and Sami Hasnine1 (2019), "An econometric investigation of the influence of transit passes on transit users' behavior in Toronto," *Public Transport*, 11, 111–133.

Handel, Benjamin R. (2013), "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts," *American Economic Review*, 103(7), 2643–2682.
Hanemann, Michael W. (1984), "Discrete/Continuous Models of Consumer Demand," *Econometrica*, 52, 541–561.
Huang, Ching-I (2008), "Estimating Demand for Cellular Phone Service under Nonlinear Pricing," *Quantitative Marketing and Economics*, 6, 371–413.
Kamargianni, Maria, Weibo Li, Melinda Matyas, and Andreas Schäfer (2016), "A Critical Review of New Mobility Services for Urban Transport," *Transportation Research Procedia*, 14, 3294–3303.
Lambrecht, Anja and Bernd Skiera (2006), "Paying Too Much and Being Happy About It: Existence, Causes, and Consequences of Tariff-Choice Biases," *Journal of Marketing Research*, 43, 212–223.
Liu, Shasha, Toshiyuki Yamamoto, Enjian Yao, and Toshiyuki Nakamura (2020), "Exploring Travel Pattern Variability of Public Transport Users Through Smart Card Data: Role of Gender and Age," *IEEE Transactions on Intelligent Transportation Systems*, 1–10, DOI: 10.1109/TITS.2020.3043021.
Miravete, Eugenio J. (2002), "Estimating Demand for Local Telephone Service with Asymmetric Information and Optional Calling Plans," *Review of Economic Studies*, 69, 943–971.
Train, Kenneth E., Daniel L. McFadden, and Moshe Ben-Akiva (1987), "The Demand for Local Telephone Service: A Fully Discrete Model of Residential Calling Patterns and Service Choices," *RAND Journal of Economics*, 18, 109–123.
Wilson, Robert B. (1993), *Nonlinear Pricing*, New York: Oxford University Press.
Zureiqat, Hazem Marwan (2006), "Fare Policy Analysis for Public Transport: A Discrete-Continuous Modeling Approach Using Panel Data," MA thesis, Cambridge, MA: Massachusetts Institute of Technology.

A Details of Data Processing

This appendix details the steps to construct the daily usage panel data used in our empirical analysis from the transaction records in the original data.

1. Cleaning the data:

    (a) Drop bus routes not included in the transit pass program. We remove highway bus routes (1717), university shuttle routes (NCCU), and private cooperate shuttle routes.

    (b) Remove duplicated records (with identical card ID and usage time).

    (c) In the bus data, there could be either one record or two records for a single ride. For a short-distance ride, a passenger only need to tap the smart card once, so we observe one record for such a ride. For a long-distance ride a passenger needs to tap the card twice, both on boarding and alighting, and the fare is twice (or a higher multiple) of the base fare. We have two records for these rides. We need to convert the data from "tap level" to "ride level". When two consecutive tap records taking place within 80 minutes have identical bus route and identical vehicle plate, We combine them into one observation as a single bus ride.

    (d) The activating date of a transit pass is only reported in the MRT data set. We recover the activating date for bus and bike records using the pass duration observed in the MRT record.

    (e) Eliminate cards with two consecutive pass activated within less than 30 days. This could be caused by coding errors or due to refunds.

2. Computing the monetary value of each ride for pass holders:

    (a) The monetary value is directly available in the MRT data set regardless the holder paid with a transit pass. It has correctly accounted for potential transfer discounts.

    (b) The regular base fare of a bus ride is 15 TWD for a regular route. The fare is 22 TWD for owl routes.

    (c) The fare is 50% off (at 8 TWD) for the "tourist routes" (795, 856, 862) during holiday weekends. The fare is zero for 982 during rush hours.

    (d) Bus fare is determined by the number of zones between the boarding stop and the alighting stop. The longest route has eight zones. The fare is computed as the product of the number of zones and the base fare.

(e) When a bus ride is within 80 minutes of an MRT ride or a bike rental, there is a transfer discount. The discount after an MRT ride is 8 TWD, and the discount after a bike rental is 5 TWD. Discounts are also offered between certain types of bus routes ("metro bus" and "minibus"). The transfer discount is recorded in the bus data set. Consequently, we can directly subtract the discount from the computed bus fare to obtain the monetary value for each bus record.

(f) The month pass only covers the fare for the first 30 minutes of bike rentals in Taipei City, which is priced at 5 TWD. Consequently, the monetary value of any bike rental is 5 TWD regardless the actual rental duration. However, if a bike rental occurs within 80 minutes of a MRT or bus ride, it costs zero due to a transfer discount. Therefore, for such bike rentals, the monetary value of using a pass is zero.

(g) The month pass does not cover bike rentals in New Taipei since the first 30 minutes of a rental is always free during my research period. We do not use bike rental in New Taipei in this study.

(h) Danhai light rail is excluded in the study due to lack of data. Its service commenced in December 2018, and starts to charge in February 1, 2019. Its average daily passenger volume was ... during the research period.

3. Creating the daily usage panel data:

(a) Use 3 am as the cutoff time on each day to compute the daily usage.

(b) The daily usage is simply the sum of the monetary values of all rides on a given day for each card ID.

(c) Fill in the gap between any two observed days in the original data for each card ID. The usage levels are zero for these filled-in observations.

(d) The usage frequency $freq_i$ is computed as

$$freq = \frac{\text{Number of days with positive usage} - 1}{LifeSpan - 1},$$

where $LifeSpan$ is the duration from the first record to the last record of the card in the data.