Discussion Paper Series

RIEB

Kobe University

DP2026-22

Experience Premium in Consumer Search: Evidence from Drivers' Refueling Behavior

RIEB Junior Research Fellow

Satoshi IMAHIE

June 24, 2026

* This Discussion Paper won the Kanematsu Prize (FY 2025).

Research Institute for Economics and Business Administration

Kobe University

2-1 Rokkodai, Nada, Kobe 657-8501 JAPAN

Experience Premium in Consumer Search: Evidence from Drivers' Refueling Behavior*

Satoshi Imahie†

June 23, 2026

Abstract

Why do uninformed consumers pay more? A long-standing answer is market unfamiliarity. Yet how much it contributes to information frictions, and whether it fades with experience, have not been directly measured. I answer these questions using millions of fuel purchases by Japanese drivers. Identification exploits two kinds of variation in familiarity, from travel and relocation. Comparing the same driver across markets, I find that unfamiliarity accounts for 62% of the price gap between informed and uninformed consumers, while persistent individual differences explain the rest. This unfamiliarity-driven gap (the experience premium) declines with repeated purchases, consistent with learning in consumer search.

Keywords: Consumer Search, Information Friction, State Dependence, Learning, Gasoline

*I am deeply grateful to my advisors Mathias Reynaert and Andrew Rhodes for their patient guidance, constant encouragement, and invaluable advice. I thank Olivier De Groote, Isis Durrmeyer, Heiko Karle, Simon Martin, Mitsukuni Nishida, Kazuma Takakura, Yuta Toyama, Rentaro Utamaru, and Yasutora Watanabe as well as audiences at the 2025 EARIE Annual Conference (Valencia, Spain, August 2025). I acknowledge financial support from the European Research Council (ERC-2019-STG-852815 "PRIDISP"; ERC grant 101088307 "DMPDE") and from the French National Research Agency under the Investissements d'Avenir programme (ANR-17-EURE-0010). I am grateful for the opportunity to work with the E-nenpi data. This paper received the Kanematsu Prize 2025 from the Research Institute for Economics and Business Administration (RIEB), Kobe University. Views and opinions expressed are those of the author alone and do not necessarily reflect those of the European Union, the European Research Council Executive Agency, RIEB, or the data provider. Any errors are my own.

†Toulouse School of Economics; imahiesatoshi1012@gmail.com

I Introduction

Since Stigler (1961), consumer search has been understood as a costly process through which consumers acquire information about sellers and their offers. Empirical work has made progress in estimating the magnitude of these frictions, but much less is known about what generates them. One important component is consumers' unfamiliarity with the markets in which they buy. The idea goes back to Stigler's discussion of why uninformed consumers pay more than informed:

> *"[Inexperienced buyers (tourists)] have no accumulated knowledge of asking prices, and even with an optimal amount of search they will pay a higher price on average" (p. 219).*

This perspective is central to interpreting price dispersion as a measure of market ignorance. Moreover, identifying the source of search frictions matters for policy and managerial practice: different sources call for different responses. Yet whether the unfamiliarity penalty exists, how much it contributes to overall information frictions, and how it evolves with experience have not been directly measured.

The main empirical challenge is the need for data that are rarely available in practice: observations of the same consumers making purchases in markets with varying degrees of familiarity. Without within-consumer variation, market-specific unfamiliarity cannot be separated from time-invariant individual traits. Without variation in familiarity, how it evolves with experience cannot be traced. I utilize individual-level refueling reports collected from a mobile phone application in Japan, comprising more than 50,000 drivers and over 2 million refueling observations across five years. Each transaction reports the date, station, price paid, quantity, and odometer reading; each driver reports a home zip code at registration. Three features make this setting well suited to the questions above. First, the long panel and the prevalence of long-distance trips generate substantial within-driver variation in the familiarity of the market where a refueling occurs: roughly 3% of transactions take place far from the driver's home, in cities visited only rarely. Second, an identifiable subsample of drivers relocates during the sample period, providing variation in market familiarity between a consumer's own old and new home markets. Third, the Japanese retail fuel market provides a natural laboratory for studying consumer search: the product is homogeneous, station locations are highly stable, and there is no national price-comparison platform.

I begin by measuring the overall information friction in this market. Over the long sample period, I observe drivers making temporary visits to areas far from their home city. Together with the large sample size, this allows me to exploit substantial within-market variation in unfamiliarity across drivers. Comparing the conditional mean prices paid by familiar and unfamiliar drivers within the same city–week, I find that unfamiliar drivers pay 2.45 yen per liter (1.9%) more—a magnitude comparable to the "Value of Information" measures derived from price dispersion alone (Varian 1980). This corresponds to the gap between "informed" and "uninformed" consumers and serves as the benchmark.

Next, I isolate the *experience premium*: the extra price the same driver pays in unfamiliar markets. Following Goldman and Johansson (1978), information frictions can be classified into three channels: the opportunity cost of time (tied to income), a cognitive component of search efficiency (tied to education, age, and ability), and an experiential component of search efficiency (tied to familiarity with the specific market). The first two are properties of the consumer and do not vary across the markets she enters; the experiential channel does. Adding individual fixed effects to the within-city–week comparison absorbs the time-invariant part of the first two channels and identifies the third. Whereas the overall information friction compares unfamiliar to familiar drivers within the same market, the experience premium compares the same individual across markets she knows to differing degrees.

The estimated experience premium is 1.52 yen per liter, accounting for 62% of the overall information friction. The remaining 38% reflects stable individual heterogeneity, much of which is captured by variables that proxy for the opportunity cost of time, such as fuel grade and vehicle price. Applied to Japan's roughly 40 billion liters of annual gasoline consumption and the 3% of transactions occurring in unfamiliar markets, the overall information friction implies excess consumer expenditure on the order of 2.9 billion yen per year, of which the experience premium accounts for 1.8 billion yen. The estimate of the experience premium is robust to plausible alternatives: proxies for trip-specific opportunity costs and urgency (long-holiday indicators, tourist density, mileage since the last refueling, purchase quantity), brand-loyalty controls such as price promotions, and more granular time fixed effects all leave the coefficient close to its baseline value.

What kinds of price variation lead to these frictions? Two classic models of price dispersion point to different answers. In Stigler (1961), dispersion is *spatial*: some stations are persistently cheaper than others, so informed consumers gain by directing their purchases toward those persistently cheaper stations. In Varian (1980), dispersion is *temporal*: a given station's price varies transitorily, for example through sales, so the gain comes from exploiting temporary discounts. To relate the measured friction to these two components, I apply the decomposition of Gelbach (2016), which decomposes the estimated friction into the part accounted for by station fixed effects and the remaining residual component. The decomposition of the overall information friction reveals that both components are positive and economically meaningful. I interpret this as evidence that the cost of unfamiliarity cannot be attributed to a single source: informed consumers gain both by knowing which stations are persistently cheap and by exploiting transitory price variation.

How does the experience premium evolve as consumers accumulate market experience? I use two complementary designs. The first extends the binary unfamiliarity measure to the count of repeated visits and shows that the premium declines monotonically with visit frequency. The second exploits residential relocation. Because relocation generates continuous exposure to a new market, time since relocation is a sharper measure of accumulated experience than discrete visit counts, and because the comparison is drawn between a consumer's own old and new home markets, the design is less vulnerable to the trip-specific confounds noted above. An event study yields four findings. First, prices are stable in the months before relocation, consistent with consumers having reached a steady state in their original market. Second, prices jump sharply at the first refueling afterward, indicating that the experience premium emerges immediately upon facing an unfamiliar market. Third, the gap narrows gradually as the driver accumulates experience in the new market. Fourth, it returns to the pre-relocation level within one to two months, reaching a new steady state. Taken together, these results suggest that consumers learn through repeated purchases, implying non-stationarity in consumer search.

Finally, to look behind the prices consumers pay, I turn to the choices they make, tracking how persistently drivers return to the same station and brand around a residential relocation. Persistence is high and stable before a move, collapses immediately afterward, and then recovers only gradually, over three to four months. Because a sequence of independent, history-independent search problems would predict persistence to return to its steady state at once, this slow rebuilding indicates that consumers reconstruct their consideration sets as they accumulate experience in the new market. Together with the price dynamics, this provides further evidence that search in this market is non-stationary and experience-dependent.

This article contributes to three strands of literature. First, it contributes to the established literature that measures information frictions. One strand uses market-level outcomes such as price dispersion, together with equilibrium search models. A second strand embeds search or consideration-set formation into discrete-choice random utility models and uses choice data; see Baye et al. (2006) and Honka et al. (2019) for surveys. I contribute to this literature by offering a new measure of information frictions: the within-market gap in prices paid between drivers familiar (informed) and unfamiliar (uninformed) with the market.[1] More recent work has turned to the underlying sources of these frictions. The closest article is by Nishida and Remer (2018), who explain cross-market variation in search costs using market-level demographics such as household income in U.S. retail gasoline markets. In contrast, I ask how drivers' market-specific experience and other individual-level factors shape search frictions, holding fixed the local market environment.

Second, this study also builds on the traditional view of the role of market experience in consumer search. Stigler (1961) first emphasized the importance of market knowledge in determining search outcomes, but constructing a direct measure of buyer experience has long been challenging (Baye et al. 2006). A few exceptions stand out: Sorensen (2000) exploits prescription drugs, where purchase frequency is directly observable at the drug level, whereas Pennerstorfer et al. (2020) use the local share of commuters versus tourists as a proxy for aggregate consumer experience in a market. Unlike these studies, which link market-level measures of experience to market-level outcomes such as price dispersion, I provide individual-level evidence on how market knowledge translates into the prices consumers actually pay. This approach lets me quantify the monetary cost of unfamiliarity and trace how it evolves as consumers gain market experience.

Third, I contribute to the small but growing literature on consumer search in repeated-purchase settings. Standard consumer search models typically focus on one-shot settings and assume that buyers face a stationary environment, in which the distribution of prices and search costs is taken as exogenous and history-independent.[2,3] These assumptions yield analytical and empirical tractabil-

[1] Other studies also use the prices consumers pay to study the returns to consumer price knowledge (more accurate price beliefs) (Friberg et al. 2025) and the gains from additional search (Ratchford and Srinivasan 1993, Seiler and Pinna 2017).

[2] Several studies allow belief updating within a search episode (Koulayev 2013, Santos et al. 2017, Jindal and Aribarg 2021, Wu et al. 2024), though they retain a one-shot setting; in a repeated-purchase setting, Matsumoto and Spence (2016) document belief convergence with experience in textbook markets.

[3] Some studies, such as Ackerberg (2003), examine learning about experience goods in repeated-purchase settings.

ity, but may not capture key features of real-world markets in which consumers' early experiences shape what they know about the market and how they search in subsequent purchases. Empirical work on consumer search in repeated-purchase settings is limited. A few exceptions are Wu (2017) and Fan et al. (2025), who develop structural models of search and learning across repeated purchases.[4] This article complements these strands by providing reduced-form evidence of non-stationarity in consumer search through the accumulation of market-specific experience.[5]

The remainder of the article is organized as follows. Section II describes the market background, data, and variable construction. Section III examines the overall information friction and the experience premium. Section IV examines how the experience premium evolves with additional experience and its mechanism. Section V presents additional findings on drivers' spatial exploration patterns. Section VI concludes.

II Data

Market Background

The empirical setting is ideal for my analysis for several reasons. First, consumer search plays a critical role in retail gasoline markets.[6] In particular, Japan has no established price-checking system, unlike Australia, France, and Germany.[7] As a result, consumers are likely to face significant search frictions.[8] Second, gasoline is a standard homogeneous product, which allows for a straightforward measure of frictions based solely on variation in prices paid. Third, the product space in fuel markets (i.e., the geographical distribution of fuel stations) is more stable compared to

However, this literature focuses on learning about previously chosen products rather than learning about the search environment itself.

[4]Theoretical work on consumer search in repeated-purchase settings is also relatively limited. Earlier contributions include McMillan and Morgan (1988) and Benabou (1993); more recently, Rhodes and Parakhonyak (2024) analyzes search equilibria in non-stationary environments generated by repeated purchases.

[5]Byrne and De Roos (2022) also document non-stationarity in consumer search, but in a different setting: they focus on the period immediately following the disruption of a stable, coordinated price equilibrium, whereas this article traces non-stationarity arising from individual experience accumulation.

[6]For these reasons, many studies on consumer search have explored the gasoline markets. Examples include Lewis (2008), Lewis and Marvel (2011), Byrne and de Roos (2017), Nishida and Remer (2018), Pennerstorfer et al. (2020), Byrne and De Roos (2022), Wu et al. (2024), Martin (2024), Fischer et al. (2024), Dorsey et al. (2025).

[7]The U.S. also lacks a government-run, publicly accessible station-level fuel price database. Instead, consumers often rely on crowdsourced platforms such as GasBuddy. Japan has a comparable service, but its user base and the volume of price submissions are much smaller than GasBuddy's.

[8]E-nenpi provides station locations and, for a small share of stations, user-submitted prices, but this price coverage is limited and updated infrequently.

other markets such as e-commerce and consumer packaged goods, where frequent product turnover is common. Consequently, I can adequately control for market characteristics using market-level fixed effects.

The Japanese retail gasoline market is characterized by the presence of several major brands and exhibits vertical differentiation at the brand level. There are five dominant oil wholesale companies in Japan, which supply gasoline and other petroleum products to retail outlets. Approximately 80% of retail firms operate under franchise agreements with these wholesalers (i.e., brands). Consequently, their marketing strategies—including advertising campaigns and membership programs—are largely determined by their brand affiliation. For example, each brand offers its own membership card that provides fuel discounts. These brand affiliations are easily recognizable, as most retailers prominently display brand-specific signage. However, retail prices are typically set independently at the station level rather than at the brand level, and stations often adjust posted prices daily in response to crude oil prices and the prices of nearby competitors.

Data Description

I use detailed individual-level data on drivers' refueling behavior. The data are provided by a Japanese company that collects information on refueling and driving through a unique mobile phone application called E-nenpi.[9] This application is freely available and allows users to monitor their driving behavior, particularly their real-world fuel economy, as well as keep track of their refueling history.[10] To use the application, drivers are instructed to submit their odometer reading and details of their gasoline purchases each time they refuel.[11]

For each observation, I can identify the purchase date and time, the gasoline price paid, and the details of the gasoline station where the purchase was made. The dataset also includes information on individual drivers and vehicles. Notably, the inclusion of drivers' home locations (at the zip code level) allows me to distinguish between refueling events that occur in familiar areas and those in unfamiliar areas, based on the distance from home. The dataset spans April 2014 to April 2019.

[9] The same data are used in Tanaka (2020) and Knittel and Tanaka (2021). The online appendix of Knittel and Tanaka (2021) provides further details about the E-nenpi data.

[10] Users benefit from knowing their car's actual fuel economy compared to that achieved by other drivers of identically configured vehicles, as it can help them improve their driving habits and reduce fuel costs.

[11] Users can take and upload photos of fuel purchase receipts to record relevant information, which helps reduce typing errors and improve data accuracy. Furthermore, users are required to indicate "yes" or "no" when asked if any refueling has not been recorded since the previous report.

An important feature of my dataset is that it captures the actual prices paid by consumers, rather than listed prices. This allows the analysis to remain robust to brand-specific promotions or discounts that are not publicly visible. In addition, I am able to control for the brand affiliation of each fuel station in the empirical analysis. Figure A1 in Online Appendix A plots the price trends for actual retail prices from E-nenpi and week-level listed prices from national statistics.

Table 1 reports summary statistics for the key variables in the sample used for the first empirical design. On average, drivers refuel with 33.2 liters of gasoline every 14.1 days and travel 45.2 km per day. Panel B reports statistics at the driver level.[12] There are 14,039 drivers in the dataset whose home address can be identified.[13] On average, each driver in the panel is observed over 991.3 days, with 54.1 recorded refueling events. This long observation period allows me to analyze drivers' repeated purchasing behavior. During this period, I observe an average of 1.6 instances of travel to unfamiliar areas with a substantial standard deviation. In familiar areas, drivers refuel at an average of 3.9 different stations. Panel C reports information about municipalities, which I use as the unit of a gasoline market. In Japan, a municipality is the smallest unit of local government. It is responsible for services such as schools, roads, and waste management. Its closest U.S. equivalent would be a city or township, depending on the state. Hereafter, I use "municipality," "market," and "city" interchangeably. On average, a city has 9.4 gas stations within a 70.4 km^2 residential land area.

Variable Construction

This section explains how I construct the unfamiliarity measure and identify changes in residential location. After removing outlier observations as described in Online Appendix A, I construct variables that capture individual-level unfamiliarity for a city. For the first design, I define the "Unfamiliar Dummy" based on the following criteria: (1) the number of observations in the unfamiliar city is less than 10; (2) the unfamiliar city differs from the driver's residential location; (3) the nearest gas station in the city is located more than 50 km from the consumer's residence; and (4)

[12] Although some drivers own more than one vehicle, the following analysis is conducted at the car level. In what follows, I use the terms "vehicle/car" and "driver" interchangeably.

[13] In total, 56,128 drivers and 2,313,563 observations are included in the dataset, but many lack home address information, which prevents identification of "unfamiliar" areas. These drivers are excluded from the first design. However, these reports are used in the second test.

Table 1: Descriptive Statistics

	N	Mean	S.D.	Min	Median	Max
Panel A: Individual Report						
No. of days b/w refueling (days)	759371	14.08	10.96	1.00	11.00	75.00
Travel distance per day (km)	759371	45.20	41.90	1.00	33.17	300.00
Purchase quantity (L)	759371	33.17	13.33	2.00	31.58	100.00
Panel B: Driver Level Report						
Obs. duration (days)	14039	991.32	619.21	2.00	953.00	1853.00
Num. of obs	14039	54.09	54.98	1.00	37.00	1052.00
Num. of obs with distance >50 km	14039	1.64	3.98	0.00	0.00	89.00
No. of near GS visited (stations)	14039	3.94	4.08	0.00	3.00	72.00
Panel C: Municipality Level Report						
Num. of GS	1594	9.42	10.92	1	6.00	86.00
Residential Land Area (km^2)	1594	70.41	72.15	2.46	45.67	805.24

Note: This table reports summary statistics for key variables. The sample is the same as that used in the first analysis (i.e., the set of observations for which the "Unfamiliar Dummy" can be defined). "Num. of obs with distance > 50 km" is the average count of trips classified as "Unfamiliar Dummy," as defined in Section II. "No. of near GS visited" is the number of gas stations a driver has visited within 50 km of their home.

the city does not contain the refueling station most frequently used by the driver.[14] Additionally, I construct the variable "Total Number of Visits to the Unfamiliar City", defined as the total number of visits made to that city.

Table 2 provides an illustrative example of variable construction. For example, consider a driver i who resides in Osaka City and was observed from May 1st to November 1st in 2016. He refueled on the first day of each month during this period. He visited Kyoto City, Hiroshima City, and Toyota City, in addition to his hometown. The driver refueled at Kyoto City, which is located only 43 km from his residence, on October 1st. He also refueled at Hiroshima City, located 282 km away, on June 1st and November 1st, and Toyota City, located 157 km away, on September 1st. The remaining refueling instances in May, July, and August occurred in his hometown. In this case, both Hiroshima City and Toyota City satisfy $\mathbb{1}_{\text{Unfamiliar}_{ic}} = 1$, whereas Kyoto City satisfies $\mathbb{1}_{\text{Unfamiliar}_{ic}} = 0$, because it is less than 50 km from his hometown. Additionally, Hiroshima City has $n_{ic} = 2$, where n_{ic} denotes the total number of visits to the unfamiliar city, and Toyota City has $n_{ic} = 1$. In contrast, Kyoto City has $n_{ic} = 0$.

Note that, in the main specification, I rely on cross-market variation in experience. Specifically, n_{ic} is defined as the total number of visits to city c over the whole sample, so it does not depend

[14]For some drivers, the most frequently visited station is located far from their recorded residence, suggesting the residential location variable may be measured with error. To address this, drivers whose most frequently visited station lies more than 20 km from their registered home address are excluded from the analysis.

Table 2: An Example of Variable Construction: Visit Summary for Each City

	Osaka (Hometown)	**Kyoto**	**Hiroshima**	**Toyota**
Distance	2 km	43 km	282 km	157 km
5/1	✓ (1)			
6/1			✓ (1)	
7/1	✓ (2)			
8/1	✓ (3)			
9/1				✓ (1)
10/1		✓ (1)		
11/1			✓ (2)	
$\mathbb{1}_{\text{unfamiliar}_{ic}}$	0	0	1	1
n_{ic}	0	0	2	1

Note: This table shows an example of variable creation. The numbers in parentheses indicate the cumulative number of visits to each city in the dataset. Note that the distance is calculated between the driver's home address and the fuel station's address. Therefore, even if a driver refuels in their hometown, the distance is not necessarily zero.

on day d and, in particular, includes visits occurring after a given transaction. It is therefore not a running stock of experience as of day d, but a market-level measure of how rarely driver i visits c, which I use as a proxy for (un)familiarity. If I were instead to use within individual-by-market variation, the cumulative number of visits over time within a single market would suffer from an initial state problem. That is, the econometrician cannot observe whether the first recorded refueling in the data is actually the driver's first refueling in that area. Therefore, I rely on across-market variation, under the assumption that fewer refueling events in a market indicate lower consumer familiarity with that market. This approach follows Sorensen (2000), which similarly defines market-level knowledge using across-market variation.[15]

For the second empirical design, I focus on individuals who have changed their residential location, which allows me to examine within-market variation in experience levels. Unfortunately, I cannot directly identify such cases based on registered address updates, as very few drivers in my sample revise their residential information after initial registration. Instead, I use an indirect method to identify potential movers.[16] First, I define each driver's most frequently visited city and their second most frequently visited city.[17] Then, I apply the following selection criteria: (1) The geographical distance between the two cities exceeds 50 km. (2) The first refueling date in

[15] Although my main estimate relies on across-market variation, Online Appendix B shows the results that allow both within and across-market variations where I use cumulative number of visits to the cities as the key independent variable.

[16] The two designs' samples are not nested: the first excludes drivers whose zip code cannot be identified, as required to construct the "Unfamiliar Dummy," whereas the second does not need the home zip code and so uses a larger sample.

[17] Note that the most frequently visited city and the second most frequently visited city may not correspond to the driver's actual residential locations.

the second most frequently visited city is later than the last refueling date in the most frequently visited city, or vice versa. This ensures that, within the observed period, the driver did not travel between the two cities before the inferred relocation. (3) The number of refueling observations in the second most frequently visited city is over 5. After applying these criteria, 278 drivers remain in the sample.

Price Dispersion

Although my data include individual prices paid and do not contain station-level listed prices, they are still useful for examining price variations. A large body of economic literature has documented price dispersion in retail gasoline markets as partial evidence of consumer search, and my data reflect a similar pattern. Table 3 reports summary statistics of weekly price distributions across markets. Specifically, for each market and week I compute the minimum, maximum, mean, and standard deviation of regular and premium gasoline prices, and then take the average across all market-week observations. I also calculate "Value of Information (VOI)" following Varian (1980) as $E[p] - E[p_{min}]$. For example, for regular gasoline prices, the average weekly market-level maximum is 128.8 yen per liter (USD 4.38 per gallon), whereas the average minimum is 120.2 yen per liter (USD 4.10 per gallon), the mean standard deviation is 3.1 yen per liter (10.6 US cents per gallon), and the VOI is 3.4 yen per liter (11.6 US cents per gallon).[18] These values are comparable to those reported in previous research on the U.S. gasoline market (Nishida and Remer 2018). The second column shows the values for premium gasoline. Significant price dispersion still exists. Online Appendix A further investigates the sources of price dispersion in my setting, following the approaches of Chandra and Tappata (2011) and Nishida and Remer (2018). The results show the markets in my setting exhibit both spatial and temporal components of price dispersion.

Descriptive Evidence

Table 4 presents the average price paid in unfamiliar and familiar areas. "Unfamiliar Dummy" ($\mathbb{1}_{\text{Unfamiliar}_{ic}} = 1$) has 23,079 observations, which account for 3% of the total valid observations. The rich variation in the unfamiliar areas strengthens identification and allows me to perform a decomposition analysis based on fixed effects. The prices paid in unfamiliar areas statistically differ

[18] The exchange rate of JPY/USD = 111.08 corresponds to April 1, 2014, which is the base date for my deflator.

Table 3: Price Dispersion Summary Statistics

	Regular	Premium
Min Price (Yen/L)	120.2	129.9
Max Price (Yen/L)	128.8	135.7
Mean Price (Yen/L)	123.6	132.6
Price SD (Yen/L)	3.1	2.4
VOI (Yen/L)	3.4	2.7
Observations	183064	69237

Note: This table presents summary statistics of prices. For each market and week, I compute the minimum, maximum, mean, and standard deviation of gasoline prices, and then take the average across all market-week observations. The value of information (VOI) is computed as $E[p] - E[p_{\min}]$, following Varian (1980). Observations refer to the number of market-week pairs with more than two refueling report entries.

from those in familiar areas, suggesting that drivers tend to pay higher prices when refueling in unfamiliar areas. Moreover, the price difference is economically significant. According to Table 3, a difference of 3.9 yen per liter (13.3 US cents per gallon) exceeds one standard deviation.

However, these price differences may be partly explained by market-level factors or differences in the timing of refueling. Therefore, in the regression that follows, I account for market-level and temporal heterogeneity by including city-by-week and day fixed effects.

Table 4: Market Familiarity and Price Paid

Variable	Purchase Price (Yen/L)	Observations
0: Familiar	126.2	736292
1: Unfamiliar	130.1	23079
Price Difference	3.9	
Welch's t test	Reject (t = 35.84)	

Note: This table shows the average price paid, conditional on the status $\mathbb{1}_{\text{Unfamiliar}_{ic}}$, which is defined according to the criterion outlined in Section II.

III Refueling in Unfamiliar Areas

Empirical Model

I begin by measuring the overall information friction in this market. I define the information friction as the differences in prices paid between familiar drivers and unfamiliar drivers within the same city-week. This price gap reflects both unfamiliarity and stable differences across drivers, such as the opportunity cost of time; separating the two is the goal of this section and the next. Specifically, I estimate the following regression for driver i on day d of week w

$$y_{ijd} = \alpha_1 \mathbb{1}_{\text{Unfamiliar}_{ic}} + \mu_{cw} + \mu_d + \epsilon_{ijd} \tag{1}$$

The outcome variable is the price paid or the logarithm of the price paid at the station j, located in city c. The coefficient α_1 is the parameter of interest. The dummy variable $\mathbb{1}_{\text{Unfamiliar}_{ic}}$ equals one if the city c is unfamiliar to individual i, as defined in Section II. The regression includes city-by-week fixed effects μ_{cw}, and day fixed effects μ_d. This interpretation follows Pennerstorfer et al. (2020). If I interpret drivers who refuel in familiar areas as "informed" and those who refuel in unfamiliar areas as "uninformed," the resulting price difference can be interpreted as the gain from information. An advantage of this measure is that it does not rely on assumptions about the search model or firms' pricing strategies. More recently, Friberg et al. (2025) also exploit the relationship between price knowledge and individual prices paid as a measure of information frictions.

Next, I isolate the experience-based component of information friction. Following Goldman and Johansson (1978), information frictions can be classified into three channels: the opportunity cost of time (tied to income), an information-processing or cognitive cost of search (tied to education, age, and ability), and an experience-related component of search efficiency (tied to familiarity with the specific market). The first two are properties of the consumer and do not vary across the markets she enters; the experiential channel does. Adding individual fixed effects to the within-city–week comparison therefore absorbs the time-invariant part of the first two channels—in particular, persistent cross-consumer differences in the opportunity cost of time and in cognitive search ability—and identifies the experiential channel. Whereas the overall information friction compares unfamiliar to familiar drivers within the same market, the experience-based component compares the same individual across markets she knows to differing degrees.

Specifically, I estimate the following regression for driver i on day d of week w

$$y_{ijd} = \alpha_2 \mathbb{1}_{\text{Unfamiliar}_{ic}} + \mu_i + \mu_{cw} + \mu_d + \epsilon_{ijd} \tag{2}$$

The outcome variable is the price paid or the logarithm of the price paid at the station j, located in city c. The coefficient α_2 is the parameter of interest. The dummy variable $\mathbb{1}_{\text{Unfamiliar}_{ic}}$ equals one if the city c is unfamiliar to individual i, as defined in Section II. The regression includes individual fixed effects μ_i, city-by-week fixed effects μ_{cw}, and day fixed effects μ_d. In words, α_2 captures the difference in prices paid when an individual refuels in unfamiliar rather than familiar areas, after controlling for market- and day-level differences. I interpret α_2 as an experience-related information friction and refer to it as the **"experience premium"**. Appendix A explains how α_2 can be mapped within a search-theoretic framework.

In econometric terms, Equation (1) is subject to upward bias arising from a selection problem: unobserved individual-level traits may be correlated with the Unfamiliar Dummy. For example, individuals who travel more frequently tend to have higher incomes and lower price sensitivity—that is, higher search costs. Because such drivers are over-represented among refueling events in unfamiliar areas, the unconditional comparison overstates the pure effect of unfamiliarity. Individual fixed effects, as included in Equation (2), eliminate this bias. As noted above, however, the bias itself is of interest, because it captures the contribution of stable individual heterogeneity to overall information frictions. Appendix A formalizes this selection problem.

Main Results

Table 5 reports the relationship between the price paid and refueling in unfamiliar areas. Columns (1) and (5) correspond to Equation (1), where the dependent variables are the price paid and its logarithm, respectively. The coefficients α_1 are positive and statistically significant at the 1% level, indicating that drivers who refuel in unfamiliar areas tend to pay higher prices within a given city-week. Therefore, the estimated information friction is 2.45 yen per liter, which accounts for 1.91% of the price paid. My estimate is also close in magnitude to information-cost measures derived from market-level price dispersion. For example, as shown in Table 3, the value of information implied by Varian (1980) is 3.4 yen for regular gasoline and 2.7 yen for premium gasoline. These figures are broadly consistent with my estimated information friction of 2.45 yen. The two measures are not identical—the VOI is the within-market gap between the mean price and the expected minimum, whereas α_1 is the gap in mean prices paid between unfamiliar and familiar buyers—but both are designed to capture the monetary value that better information confers on a buyer.

Columns (2) and (6) correspond to Equation (2), estimating the experience premium α_2. The coefficients remain positive and statistically significant at the 1% level, confirming that even after controlling for individual-level heterogeneity in baseline search costs, drivers pay higher prices in unfamiliar areas. Hence, the estimated experience premium is 1.52 yen per liter, which accounts for 1.19% of the price paid.

To provide a sense of the economic magnitude, consider the following back-of-the-envelope calculation. Japan consumes approximately 40 billion liters of gasoline annually.[19] Using the share of unfamiliar-area refueling events in my sample (approximately 3%; see Table 4) and my preferred estimate of 1.52 yen per liter in excess prices attributable to the experience premium, I obtain an implied 1.8 billion yen (16.2 million dollars) per year in excess prices paid by consumers who refuel in unfamiliar markets. As Friberg et al. (2025) note, becoming better informed raises consumer welfare through two channels: a direct effect, whereby an informed consumer pays lower prices facing a given price distribution, and an equilibrium effect, whereby a larger share of price-sensitive informed consumers intensifies competition and lowers posted prices (Varian 1980). Like Friberg et al. (2025), I cannot speak to the equilibrium channel with my reduced-form data and methodology; my calculation quantifies only the direct gains, holding current prices and quantities fixed, and measures what these consumers would save if their unfamiliarity were removed, taking the price distribution as given. Following the same logic as Friberg et al. (2025), this direct figure should be read as a lower bound on the total consumer savings from reducing unfamiliarity, since the equilibrium response of prices would generate additional savings.

Comparing columns (1) and (2) shows that the experience premium accounts for 62% of the overall information friction (1.52 of 2.45 yen per liter). The remaining 38% appears to be explained by individual-level traits. The 38% residual reflects the first two channels in the Goldman and Johansson (1978) classification—opportunity cost of time and cognitive costs. Nishida and Remer (2018) investigate their relative contributions to search costs by regressing cross-market variation in estimated search costs on observable market-level demographics such as household income (a proxy for the opportunity cost of time), and conclude that the opportunity cost of time, rather than information-processing or cognitive costs, is the dominant driver. My data do not contain a credible proxy for the cognitive channel, so a full decomposition is infeasible; I can, however, examine the role of the opportunity cost of time. To do so, I omit the individual fixed effects and replace them with two observable proxies for consumer heterogeneity: fuel grade and car price.

[19]This figure is drawn from the Ministry of Land, Infrastructure, Transport and Tourism's Survey on Motor Vehicle Fuel Consumption for fiscal years 2014–2019.

Table 5: Price Paid in Unfamiliar Areas

	Price Paid				Log(Price Paid)	
	(1)	(2)	(3)	(4)	(5)	(6)
Unfamiliar Dummy	2.4487	1.5166	1.6895	1.6756	0.0191	0.0119
	(0.1184)	(0.0488)	(0.0659)	(0.0660)	(0.0009)	(0.0004)
Luxury Car Dummy				−0.1267		
				(0.0817)		
Observations	759371	759371	759371	753062	759371	759371
$R^2_{within,adj}$	0.01	0.01	0.01	0.01	0.01	0.01
FE: Individual		Yes				Yes
FE: Date	Yes	Yes	Yes	Yes	Yes	Yes
FE: Oil Type			Yes	Yes		
FE: City-by-Week	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table shows the relationship between the price paid at the time of refueling and whether the refueling occurred in an unfamiliar area. The dependent variable is the price paid in Columns (1)–(4) and the logarithm of the price paid in Columns (5)–(6). Each observation is at the car-station-day level. The key independent variable, "Unfamiliar Dummy," is a binary indicator equal to one if the individual refuels in a city that is unfamiliar to them, as defined in Section II. Columns (3) and (4) replace individual fixed effects with oil-type fixed effects , which distinguish high-octane (premium) from regular gasoline, and, in Column (4), a luxury-car indicator that equals one if the car's price exceeds the sample median and zero otherwise, as proxies for the opportunity cost of time. Standard errors are clustered at the car level. $R^2_{within,adj}$ represents the adjusted within R-squared.

Following Chandra and Tappata (2011), I interpret these variables as income-related correlates that proxy for the opportunity cost of time. Columns (3) and (4) of Table 5 report the results. The two proxies absorb most of the variation previously captured by the individual fixed effects, consistent with the opportunity-cost channel being a leading candidate.

A natural follow-up is whether the two components are independent, or whether the experiential channel varies systematically with the same traits that drive the residual component. To address this question, I conduct a heterogeneity analysis by interacting the unfamiliar dummy with proxies for the opportunity cost of time, which is also a correlate of income.

Columns (1) and (2) of Table 6 report the results from regressions of the price paid. The interaction terms are not statistically significant, indicating that the experience premium does not vary systematically along these observable dimensions. This null finding carries two implications. First, it speaks to the distributional incidence of informational interventions. Because the experience premium is uniform across income proxies, the benefits of policies that resolve unfamiliarity-driven frictions—such as price-comparison platforms—would be distributed in a manner that is broadly neutral with respect to income. By contrast, as shown in Table 5, the residual friction attributable to the opportunity cost of time remains concentrated among consumers with higher opportunity costs. Eliminating the experience premium component therefore does not generate progressive redistribution on its own. Second, the absence of heterogeneity supports the interpretation that the two components of information frictions identified above, the experiential channel and the stable individual heterogeneity channel, are additive. This separation has implications for both policy design and managerial strategy. On the policy side, interventions such as price-comparison platforms are likely to mitigate the experience-based information frictions, but the opportunity-cost component will persist whenever such tools themselves require non-trivial access costs in time or attention. The two channels thus call for different instruments. On the managerial side, the separation suggests that firms targeting consumers on the basis of income or other stable individual traits, and firms targeting them on the basis of prior purchase experience and market familiarity, are operating on distinct margins whose effects do not substitute for one another. Consumer-level data that combine both dimensions—such as loyalty-card histories linked to vehicle or demographic information—may therefore allow firms to address each margin with a tailored instrument.

Column (3) of Table 6 also reports a heterogeneity analysis with respect to gas station density. Intuitively, in denser markets, drivers face more alternatives within a short distance and can therefore search more easily. However, the interaction between the unfamiliar dummy and log gas station density is statistically indistinguishable from zero, indicating that the experience premium does not vary systematically with local market thickness.[20]

Robustness

Higher trip-specific opportunity costs in unfamiliar areas may partly explain the experience premium. For instance, drivers may be less willing to search while traveling, leading them to pay higher prices in unfamiliar areas. To examine this possibility, I conduct several additional analyses. First, I interact the unfamiliar-area dummy with an indicator for long holiday periods, defined as stretches of more than three consecutive holidays. Travel during such periods is more likely to be leisure-related and may therefore involve higher opportunity costs. Second, I incorporate variation in touristic intensity by interacting the unfamiliar-area dummy with an indicator for above-median

[20] Notice that Premium Dummy and Luxury Car Dummy are dropped because they are collinear with the individual-level fixed effects. In contrast, Log(GS Density Per Area) is identified from within-market-level changes.

Table 6: Heterogeneous Effects of Refueling in Unfamiliar Areas

	(1)	(2)	(3)
Unfamiliar Dummy	1.5567	1.5454	1.4939
	(0.0550)	(0.0690)	(0.1197)
Unfamiliar Dummy × Premium Dummy	−0.1371		
	(0.0978)		
Unfamiliar Dummy × Luxury Car Dummy		−0.0424	
		(0.0892)	
Log(GS Density per Area)			−1.4543
			(0.4620)
Unfamiliar Dummy × Log(GS Density per Area)			−0.0116
			(0.0579)
Observations	759371	753062	759371
$R^2_{within,adj}$	0.01	0.01	0.01
FE: Individual	Yes	Yes	Yes
FE: Date	Yes	Yes	Yes
FE: City-by-Week	Yes	Yes	Yes

Note: This table shows the heterogeneous effects of refueling in unfamiliar areas. Each column augments the baseline specification of Equation 2 with an interaction term between the unfamiliar dummy and an individual- or market-level characteristic. Column (1) interacts the unfamiliar dummy with an indicator for premium-gasoline vehicles that equals one if the car uses high-octane (premium) gasoline and zero if it uses regular gasoline. Column (2) interacts the unfamiliar dummy with a luxury-car indicator, a proxy for vehicle price that equals one if the car's price exceeds the sample median and zero otherwise. Column (3) interacts the unfamiliar dummy with the log gas station density per residential land area; the level of density is also included. Standard errors are clustered at the car level. $R^2_{within,adj}$ represents the adjusted within R-squared.

tourist intensity, measured by the ratio of tourists to commuters using private cars. Third, to capture differences in urgency during non-routine travel, I interact the unfamiliar-area dummy with indicators for above-median mileage since the last refueling and above-median purchase quantity. If the coefficient on the unfamiliar dummy remains close to the baseline estimate, this would suggest that differences in opportunity costs or urgency do not fully account for the experience premium. As shown in Table A4 in Online Appendix B, the estimates are broadly similar to the baseline results, indicating that the experience premium is not primarily driven by heterogeneity in trip-specific opportunity costs and urgency.

As an additional robustness check, Table A6 in Online Appendix B reports the results with various control variables. Specifically, I include more granular time fixed effects, such as market-by-day fixed effects and market-by-day-by-time fixed effects. I also replace market-by-week fixed effects with the mean city price and the variance of city prices. In all specifications, the results are nearly identical to those from the baseline model.

I interpreted α_2 as an experience-related information friction. However, the effects may capture the effect of loyalty programs offered by familiar gas stations. In Japan, in particular, price discounts for membership card holders are common. Unfortunately, my data do not include information on promotions. Instead, I include a variable indicating whether the driver has previously used that brand as a proxy for familiarity with brands.[21] Table A7 in Online Appendix B shows the results. The estimated results are robust to controlling for it.

Interpreting α_1 and α_2 as information frictions rests on treating gasoline as a homogeneous good. When the product is identical across stations, any lower price an informed consumer secures is a pure return to search. If stations instead differ in convenience or quality, however, paying a higher price need not reflect a search failure—it can simply be the going price for a more convenient or higher-quality station. Part of the gap between unfamiliar and familiar drivers would then reflect what they buy rather than how well they search, so α_1 and α_2 would overstate the informational friction. Two forms of differentiation are relevant. First, spatial differentiation within a city: stations at convenient or high-traffic locations, such as those near highways or tourist destinations, are persistently more expensive, so a driver who refuels there is paying for convenience rather than failing to search. To absorb this component, I add 3 km mesh fixed effects, which control for persistent price differences across micro-locations within a city. As Table A9 in Online Appendix B shows, these area differences account for only about 20% to 30% of α_1 and α_2 and do not fully explain them. Second, vertical differentiation across brands: major brands differ in perceived quality and ancillary services, and because consumers recognize this commonly ranked quality, a driver who chooses a higher-quality brand pays more by choice. Table A10 in Online Appendix B shows that the premium remains robust to controlling for brand-level vertical differentiation. The remaining gap is therefore not a compensating differential for convenience or brand quality, but a residual informational disadvantage.

[21] Ideally I would use station-by-day-by-oil fixed effects, but only 0.2% of observations vary within that cell. Table A8 in Online Appendix B shows experience-related price differences do arise within it, and the brand-familiarity proxy absorbs them.

Gelbach Decomposition

The information friction measured in Equation (1) captures the realized gains that better-informed consumers obtain, but it is silent on what kinds of price variation lead to these frictions. Theory points to two distinct sources, corresponding to the two classic models of price dispersion. In the search model of Stigler (1961), dispersion is *spatial*: some stations are persistently cheaper than others, and the gain to informed consumers is the gain from directing purchases to persistently lower-priced stores. In the model of sales of Varian (1980), dispersion is *temporal*: the gain comes not from knowing which store is persistently cheap, but from exploiting intertemporal price variation—timing purchases to coincide with temporary sales rather than paying regular prices.To trace the measured friction to these two underlying components, I apply the decomposition of Gelbach (2016) to Equation (1).

I decompose the overall information friction into (i) the contribution of station-level fixed effects—which I refer to as the *spatial* component—and (ii) the contribution of the remaining covariates, which I refer to as the *non-station* (residual) component. Intuitively, the spatial component captures the extent to which informed drivers shift their purchases toward stations that are persistently cheaper. The non-station component captures the extent to which, holding the station fixed, unfamiliar drivers still pay more—for example, by failing to time their purchases to days on which a given station is relatively cheap.

1. I estimate the full model by adding station-level fixed effects to the baseline specification in Equation (1). The resulting coefficient, α_1^{other}, captures the non-station residual component, after absorbing permanent price differences across stations.

$$p_{ijd} = \alpha_1^{\text{other}} \mathbb{1}_{\text{Unfamiliar}_{ic}} + \mu_{cw}^{\text{other}} + \mu_j^{\text{other}} + \mu_d^{\text{other}} + \epsilon_{ijd}$$

   where μ_j^{other} denotes station-level fixed effects.

2. For each observation, take the estimated fixed effect $\hat{\mu}_j^{\text{other}}$ of its station and regress it on the covariates in the baseline specification.

$$\hat{\mu}_j^{\text{other}} = \alpha_1^{\text{spatial}} \mathbb{1}_{\text{Unfamiliar}_{ic}} + \mu_{cw}^{\text{spatial}} + \mu_d^{\text{spatial}} + \epsilon_{ijd}$$

Let $\hat{\alpha}_1^{\text{base}}$ denote the estimate of the baseline model of Equation (1). Then the following holds

$$\hat{\alpha_1}^{\text{base}} = \hat{\alpha_1}^{\text{other}} + \hat{\alpha_1}^{\text{spatial}}$$

As shown in Table A2, my market-level price distributions exhibit both spatial and temporal price dispersion. A theoretical challenge is to rationalize these two patterns within a single framework. One possible approach is to treat the cross-station ranking as well known to consumers. Under this interpretation, consumers observe the vertical differentiation across stations without engaging in costly search, and costly search concerns only the temporal component of prices: consumers may know the persistent cross-station ranking but still need to search to learn which station offers the lowest price at a given point in time (Wildenbeest 2011). Nishida and Remer (2018) apply this framework to the U.S. retail gasoline market.[22] If the cross-station ranking were a known dimension of vertical differentiation rather than an object of search, any gap between tourists and commuters in the stations they choose—and hence in the persistent component of the prices they pay—would have to be attributed entirely to taste differences. Absent such taste differences, the spatial gap is more naturally read as evidence that the cross-station ranking is itself partly an object of search. I treat the "spatial" component as part of information friction. Under this interpretation, my decomposition indicates that both the spatial and non-station components contribute to the realized gains from information.

The first and second columns of Table 7 report the decomposition. Both components are positive and economically meaningful, so each of the spatial and non-station channels accounts for a substantial share of the information friction. The information friction of unfamiliarity therefore cannot be attributed to a single source: informed consumers gain both by sorting toward persistently cheaper stations and, holding the station fixed, by paying less at a given station—consistent with timing their purchases to transitory low prices.

Friberg et al. (2025) contribute to this discussion. They show that consumers with more accurate price knowledge pay lower prices by timing their purchases to coincide with temporary sales, thereby providing direct empirical support for Varian (1980). My findings are consistent with theirs in showing that the non-station channel is an important source of gains from information. However,

[22] Wu et al. (2024) likewise interprets learning as operating over the temporal component alone in gasoline markets.

my results also point to an additional source: the spatial dimension. In this sense, my measure of unfamiliarity captures a more upstream source of information friction than the price-recall-accuracy measure employed by Friberg et al. (2025). Whereas their results speak to heterogeneity in discount utilization among consumers who are already exposed to a given market, my results indicate that market exposure itself is a fundamental determinant of consumer information.

I remain agnostic about how consumers acquire information about the spatial and temporal components of prices. As emphasized by Lach (2002), temporal price dispersion and mobility in stores' positions in the price distribution may make it difficult for consumers to identify stores that consistently post low prices. One possible mechanism is that consumers update their beliefs about station-level prices through repeated purchases, as in Wu (2017).[23] I also leave open the question of how the information frictions faced by drivers map into the spatial and temporal components of equilibrium price dispersion. Clarifying this mapping is a theoretical question for future work.

I apply the same decomposition to the experience premium in Equation (2). Because this specification includes individual fixed effects, it absorbs stable cross-driver heterogeneity and isolates the component of information friction associated with market-specific inexperience. Columns (3) and (4) of Table 7 show that both the spatial and non-station components remain positive and economically meaningful. Thus, experience-based information friction operates through the same two broad channels as overall information friction.

The relative importance of these channels, however, differs across the two measures. The spatial component accounts for 55% of the experience premium, compared with 47% of overall information friction. This comparison is informative because overall information friction combines market-specific inexperience with stable cross-driver heterogeneity, whereas the experience premium removes the latter through individual fixed effects. Because adding stable heterogeneity reduces the spatial share from 55% to 47%, stable heterogeneity must be less spatial-intensive than market-specific inexperience. Put differently, this result suggests that inexperienced drivers appear to lose money mainly because they fail to sort toward persistently cheaper stations. By contrast, stable differences across drivers seem to operate relatively more through the non-station margin. This pattern suggests that different sources of information friction may work through different channels.

[23] Ursu et al. (2024) studies how consumers' prior information affects product-level beliefs and search behavior for differentiated products.

Table 7: Gelbach Decomposition: Mechanism

	Information Friction		Experience Premium	
	Non-Station	Spatial	Non-Station	Spatial
Unfamiliar Dummy	1.3063	1.1424	0.6864	0.8303
	(0.1242)	(0.0943)	(0.0366)	(0.0657)
Observations	759371	759371	759371	759371
$R^2_{within,adj}$	0.00	0.00	0.00	0.01
FE: Individual			Yes	Yes
FE: City-by-Week	Yes	Yes	Yes	Yes
FE: GS	Yes		Yes	
FE: Day	Yes	Yes	Yes	Yes

Note: This table reports the Gelbach decomposition (Gelbach 2016) of the overall information friction and the experience premium into a spatial and a non-station (residual) component. The first two columns ("Information Friction") decompose α_1 estimated from Equation (1), and the last two columns ("Experience Premium") decompose α_2 estimated from Equation (2), which adds individual fixed effects. In each case, the "Spatial" column reports the contribution of the gas station fixed effects, and the "Non-Station" column reports the residual component net of permanent station-level differences; the two sum to the corresponding baseline coefficient. $R^2_{within,adj}$ represents the adjusted within R-squared. Standard errors are clustered at the gas station and individual level.

IV The Role of Additional Experience

Repeated Visits

Although Equation (2) identifies the price penalty associated with unfamiliarity, it is silent on how this penalty evolves with market-specific experience. To examine whether additional experience reduces the penalty, I estimate the following regression for driver i on day d of week w

$$y_{ijd} = \sum_{k=1}^{4} \beta_k \mathbb{1}_{(n_{ic}=k)} + \beta_5 \mathbb{1}_{(10>n_{ic}>4)} + \mu_i + \mu_{cw} + \mu_d + \epsilon_{ijd} \tag{3}$$

where the outcome variable is still the price paid and its logarithm. n_{ic} denotes the total number of visits to unfamiliar city c. The regression also includes individual fixed effects μ_i, city-by-week fixed effects μ_{cw}, and day fixed effects μ_d. The dummy variable $\mathbb{1}_{(n_{ic}=k)}$ equals one if n_{ic} equals k, where n_{ic} denotes the total number of visits defined in Section II. Also, $\mathbb{1}_{(10>n_{ic}>4)}$ equals one if n_{ic} is greater than 4 and less than 10.

Housing Relocation

In this section, I introduce a new identification strategy for the experience premium. Specifically, I employ an event study design that uses housing relocation as a treatment event. This approach identifies the dynamic effect of experience level changes. After moving, consumers have limited knowledge of the new market and are not subject to the initial state problem. Furthermore, by tracking the gradual accumulation of experience, I examine how long it takes individuals to converge to the post-relocation steady-state level of experience.

My baseline empirical specification follows the event study model:

$$log(p_{ijd}) = \sum_{k \in \{\leq -4, -3, \ldots, 0, \ldots, 3, 4 \leq\}} \gamma_k D_{i,m-k} + \mu_i + \mu_{cw} + \mu_d + \epsilon_{ijd} \tag{4}$$

The outcome variable is the logarithm of the price paid. $D_{i,m-k}$ is an event dummy variable that equals one if month m corresponding to day d is exactly k periods relative to unit i's event time (where $k > 0$ indicates post-event periods and $k < 0$ indicates pre-event periods), and zero otherwise.[24] The coefficients γ_k capture the dynamic treatment effects relative to the event period, where γ_{-1} is set to 0 as the reference point. Because a large relative period from the moving date may introduce noise into the estimates, the treatment window is restricted to 10 months before and after the treatment. That is, $k \in [-10, -4] \cup \{-3, -2, -1, 0, 1, 2, 3\} \cup [4, 10]$. Observations corresponding to refueling in cities other than the most familiar and the second most familiar cities, as defined in Section II, are also excluded. Unlike the standard two-way fixed effects event-study design, my specification includes individual fixed effects, city-by-week fixed effects, and day fixed effects, which together control for both market-level and temporal heterogeneity in price distributions. As the control group, I use non-movers observed in the same city-by-week as the movers; these non-movers share similar characteristics with the movers and reinforce the parallel-trends assumption.Recent econometric work shows that the standard OLS estimator can be biased in staggered-adoption settings like mine when treatment effects are heterogeneous across cohorts and/or over time. Here, I also report the estimator proposed by Sun and Abraham (2021) that uses never-treated units as controls.

[24] I define $D_{i,m-k}$ at the month level to preserve sample size. The first post-move refueling is coded $k = 0$; other same-month refuelings are coded -1 if before the move and 1 if after.

This design has several advantages compared to the first design. First, due to the concern of the initial state problem, the first design focuses on the across-market variations.[25] However, the second design can focus on within-market variation apart from the initial state problem. Just after moving to a new market, consumers have limited knowledge of it; hence, I can observe the change in market-level knowledge before and after treatment. Within-market variation in the second design is more interpretable for understanding the role of experience evolution, enabling me to clearly answer the key question of how long the experience premium lasts. Second, the first design may be biased due to differences in opportunity costs between familiar and unfamiliar areas. In contrast, the second design is based on the comparison between the previous and new residential areas and is less likely to suffer from these unobserved trip-specific opportunity cost problems. Third, the unfamiliar-area dummy used in the first design may be a noisy proxy for consumers' experience level. A key concern is that I only observe visits to unfamiliar areas when consumers choose to refuel there. In contrast, the second design, which leverages housing relocations, captures consumers' continuous exposure to a new market. As a result, the time elapsed since relocation provides a more accurate measure of experience.

Main Results

Columns (1) and (2) of Table 8 correspond to Equation (3), which replaces the unfamiliar dummy with a set of indicators for the total number of visits n_{ic} to an unfamiliar city. The coefficients β_k are positive and statistically significant at the 1% level across all visit categories. Importantly, the price gap narrows monotonically as the number of visits increases, suggesting that market-specific experience mitigates the initial disadvantage. This finding indicates that accumulated experience leads to more effective search at the individual level. This evidence supports the idea that drivers learn market-specific knowledge. Moreover, the monotonic decline is difficult to reconcile with the leading alternative—that the premium reflects trip-specific opportunity costs or urgency rather than market knowledge. The comparison underlying β_k holds the driver fixed and contrasts distant markets she visits rarely with distant markets she visits more often. Both are equally "away-from-home" refuelings and thus share the same trip characteristics—the distance involved, the associated

[25] Table A5 in Online Appendix B allows both within- and across-market variation using cumulative city visits. The limited within-city evolution supports the concern about the initial state problem.

Table 8: Price Paid and the Number of Visits to Unfamiliar Areas

	Price Paid	Log(Price Paid)
	(1)	(2)
Num Visit = 1	1.9911	0.0156
	(0.0550)	(0.0004)
Num Visit = 2	1.4276	0.0111
	(0.0907)	(0.0007)
Num Visit = 3	1.1865	0.0093
	(0.1336)	(0.0010)
Num Visit = 4	1.0204	0.0080
	(0.1469)	(0.0012)
10 > Num Visit > 4	0.7209	0.0057
	(0.1014)	(0.0008)
Observations	759371	759371
$R^2_{within,adj}$	0.01	0.01
FE: Individual	Yes	Yes
FE: Date	Yes	Yes
FE: City-by-Week	Yes	Yes

Note: This table shows how the price paid varies with the total number of visits to an unfamiliar city. The dependent variable is the price paid in Column (1) and the logarithm of the price paid in Column (2). Each observation is at the car-station-day level. The key independent variables are a set of indicators for the total number of visits n_{ic} to an unfamiliar city, as defined in Section II: "Num Visit = k" equals one if the driver has visited the city k times, and "10 > Num Visit > 4" equals one for five to nine visits. Standard errors are clustered at the car level. $R^2_{within,adj}$ represents the adjusted within R-squared.

value of time, and the leisure or business nature of travel—yet the premium is smaller in the markets she knows better. Because trip type is held fixed and only accumulated market-specific experience varies, the decline isolates the role of market knowledge rather than the opportunity cost of travel.

Figure 1 plots the estimated event-study coefficients γ_k from Equation (4); Table A11 in Online Appendix B reports the corresponding OLS and Sun and Abraham (2021) estimates in tabular form. Both the OLS estimate and the estimate from Sun and Abraham (2021) provide similar results. There are four noteworthy points. First, prior to the relocation, the point estimates are close to zero, suggesting that price levels remain stable relative to the baseline period at $t = -1$. This supports the notion that consumers are in an initial steady state before relocation. Theoretically, Hong and Shum (2006) suggest a one-to-one mapping between the distribution of search costs and market-level price distributions, and the relative value of an individual's search cost compared to others in the same market determines the relative price level. Therefore, once I control for changes in the price distribution, the price level remains stable, as supported by my results.

Second, there is a statistically significant jump in price at $t = 0$, amounting to approximately 0.6% and it is statistically significant. This finding once again rejects the null hypothesis that search behavior is independent of experience. Third, the price gap narrows as consumers accumulate experience. Finally, the price difference eventually converges to zero, reaching a new steady-state level. Specifically, after 1-2 months, small non-zero point estimates persist, yet the effect is not statistically distinguishable from zero. After three months, the point estimates almost revert to 0.

A comparison with the results in Table 8 offers further insights. First, the magnitude of the price jump at $t = 0$ is smaller than the estimated effects reported in Table 8. This is likely because the present analysis focuses on within-market variation, which allows for the exclusion of opportunity cost effects. Moreover, whereas the previous analysis examines temporary visits to unfamiliar areas, the current study focuses on continuous relocation into unfamiliar markets. As a result, consumers may begin the learning process between $t = -1$ and $t = 0$, which could explain the smaller observed effect at the time of relocation. Although the magnitude of the price jump differs, both results provide consistent evidence that accumulated knowledge narrows the observed price gap, pointing to learning in consumer search.

Consumers might face urgency immediately after relocation due to the busy nature of the moving period. As a robustness check, I conduct a heterogeneity analysis based on an urgency proxy. Specifically, I examine whether the event-study estimates change once I interact the post-event dummies from Equation (4) with an indicator for above-median mileage since the last refueling. Table A12 in Online Appendix B reports the results. The coefficient γ_0 becomes smaller once the interaction term is included, suggesting that urgency matters just after the relocation. However, the coefficients γ_1 and γ_2 remain similar to the baseline estimates, indicating that the urgency effect is temporary and does not persist beyond the immediate post-relocation period. More importantly, the overall pattern of the event-study estimates remains consistent with the baseline results, suggesting that the urgency effect does not drive the observed learning pattern.

Besides, as a robustness check, Figure A5 in Online Appendix B re-estimates the event study using two-week periods, rather than calendar months, as the unit of event time; the resulting pattern is broadly similar.

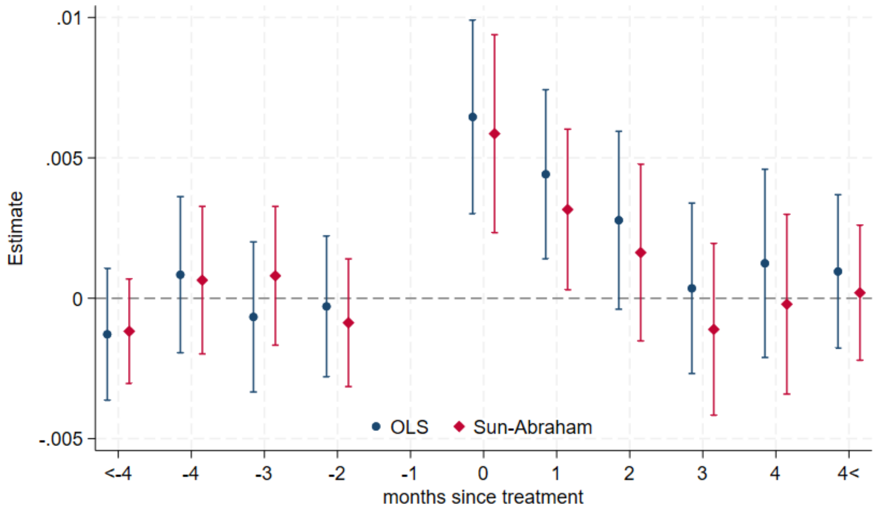


Figure 1: Price Paid and Housing Relocation

Note: This figure presents the results of an event study on housing relocation using OLS and Sun and Abraham (2021). It displays the estimates of γ_k, including point estimates and 90% confidence intervals. The coefficient for the final time period before relocation is normalized to zero. The dependent variable is the logarithm of p_{ijd}. I control for individual fixed effects, city-by-week fixed effects, and day fixed effects. The treatment group consists of movers selected based on the criteria outlined in Section II. As the control group, I use non-movers observed in the same city-by-week as the movers. Observations corresponding to refueling in cities other than the most familiar and the second most familiar cities, as defined in Section II, are excluded. The sample window covers ten months before and ten months after the relocation.

From Choice to Search

The preceding results establish that the prices consumers pay fall as market-specific experience accumulates, but they leave open why this happens. The remainder of this section turns from prices to the underlying search behavior to identify the mechanism. Specifically, I use choice persistence as a behavioral signature that distinguishes stationary search from experience-dependent search and ask how quickly consumers rebuild their consideration set after relocation.

First, I examine the persistence in store choice and use the following metric as the measure of choice persistence. Choice persistence is informative about search because it captures the extent to which past search outcomes continue to shape current choices. In addition, choice persistence helps illuminate the search mechanism by providing evidence from observed choice patterns, rather than from realized prices alone.

The choice persistence rate relative to the r-th month from the time of relocation is defined as:

$$Persist_r = \frac{1}{N_r} \sum_{i \in I_{treatment}} \sum_{d \in T_{i,r}} \mathbf{1}\left(j_{id} = j^{last}_{i,d'}\right) \tag{5}$$

where j_{id} represents the store or brand choice of driver i on day d, and $j_{i,d'}^{last}$ represents their most recent previous choice of store. $I_{treatment}$ denotes the set of drivers who have experienced relocation. $T_{i,r} = \{t \mid m(i,t) = r\}$ denotes the set of all purchases made by driver i in relative month r. Then, $N_r = \sum_{i \in I_{treatment}} |T_{i,r}|$ represents the total number of purchases made by all drivers who have experienced relocation in relative month r.

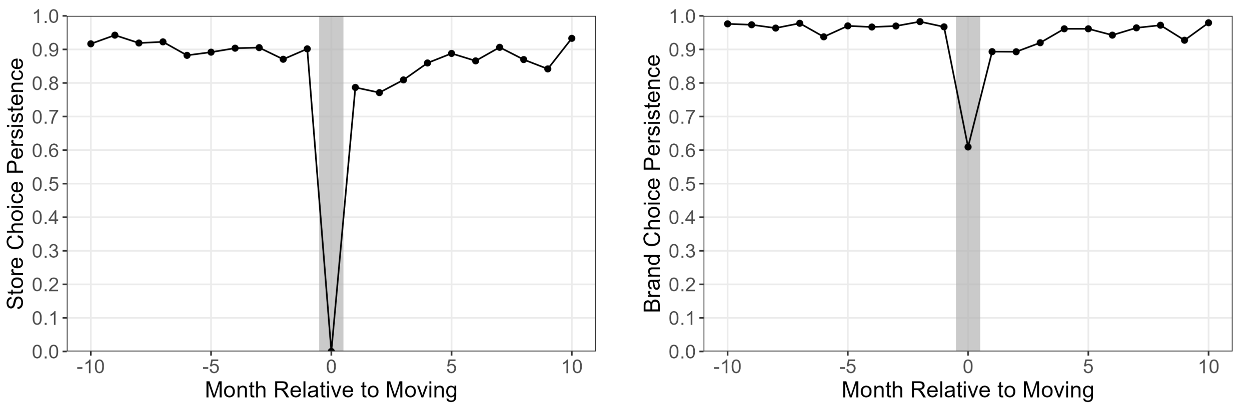
Figure 2a plots store choice persistence around relocation. Three features stand out. First, far from the move, the persistence rate is high—about 0.9—but strictly below one, so even settled drivers occasionally switch stores. Second, it is flat in the months before relocation, consistent with an initial steady state. Third, it drops sharply at the move and then rises gradually, taking about three to four months to return to its pre-move level, after which it is again stable. The drop is partly mechanical: at the move, a driver's most recent store lies in the former market and cannot be chosen again. The informative feature is therefore the speed at which persistence is rebuilt.

To interpret this pattern, it is useful to ask how store choice persistence should evolve within the new market after relocation. The relevant question is whether the driver's effective consideration set is fixed immediately after the move, or instead changes gradually as she searches and learns which stores are worth considering. The hypotheses below distinguish stationary search, in which this adjustment is immediate or absent, from experience-dependent search, in which search costs and beliefs about price distributions evolve over time.

If search is stationary and history-independent, and the market environment is stable over the window, then each driver chooses store j with a fixed probability p_j, where $\sum_{j=1}^{L} p_j = 1$. For purchases whose preceding purchase is also in the new market, store choice persistence equals the time-invariant constant $\sum_j p_j^2$. Persistence is therefore flat once the mechanical first-post-move comparison is excluded.

Even if search is stationary, so search costs and beliefs do not vary with market-specific experience, repeated purchase might create a fixed relationship between a consumer and a store. As in McMillan and Morgan (1988) and Benabou (1993), once a consumer identifies or is matched with a preferred seller, the relationship functions like a fixed contract: the consumer repeatedly returns to that seller rather than re-optimizing over the full set of stores each period. This mechanism can explain a high level of store choice persistence. However, because the relationship is fixed rather than gradually learned, persistence should quickly return to its steady-state level once the mechanical first-post-move comparison is excluded and then remain flat.

If instead the elements that govern the search problem evolve with market-specific experience, in particular, search costs and price beliefs that determine which stores are worth inspecting, then relocation resets the driver to a low-experience state in the new market. As experience accumulates, the consideration set is reconstructed and stabilizes, with fewer new stores being added. Purchases then become concentrated among the stores in this stabilized set, generating a transition path in which persistence rises toward a post-learning steady state.



(a) Store Choice Persistence **(b)** Brand Choice Persistence

Figure 2: Store and Brand Choice Persistence

Note: This figure illustrates the persistence of brand and store choices. The left and right panels show store and brand choice persistence, respectively, before and after housing relocation. The definition of choice persistence is provided in Section IV.

Figure 2b displays brand choice persistence based on the counterpart of Equation (5) with brand choice. The pattern combines a high baseline with a gradual recovery, and each component has a distinct interpretation. First, brand choice persistence is already high, around 0.6, even immediately after relocation. Unlike stores, which are market-specific and must be relearned in a new location, brands are common across markets, so a consumer's brand preference travels with her when she moves. Second, persistence does not stay flat at this baseline: it recovers to around 0.9 one month after relocation and takes about three to four months to converge to the post-learning steady state, a speed quite similar to the recovery of store choice persistence. The fact that brand choice persistence rises gradually rather than jumping immediately to its steady-state level is broadly consistent with the experience-dependent search and suggests that some learning may operate at the brand level as well.

V Additional Findings

Spatial Exploration: I provide descriptive evidence on the geographical patterns of selected stations. Recent literature on spatial exploration shows that consumers decide where to search based on their previous discoveries (Hodgson and Lewis 2023, Morozov et al. 2025). Whereas these studies focus on product space exploration, my analysis centers on geographical space exploration. Specifically, I examine how the distances from both the initial and the most recent refueling stations evolve over time following a housing relocation. Figure A3a plots the distance from the initially chosen refueling station against the number of refueling events following relocation. Three patterns stand out. First, by the 6th to 7th refueling events, consumers gradually increase their distance from the initial station. Second, around the 7th to 10th refueling, the distance peaks and briefly levels off at about 1.2 km. Third, beginning with the 11th refueling, the series shifts to a new regime and stabilizes around 1.0 km. Figure A4a in Appendix Online Appendix B shows the corresponding figure that uses months relative to the move on the x-axis.

Figure A3b plots the distance from the most recently chosen refueling station against the number of refueling events following relocation. Three notable patterns emerge. First, at the beginning (from 1st to around 5th refueling events), the distance from the last refueling station generally increases, peaking above 0.7 km, which suggests active spatial exploration. Second, after this peak, the distance tends to decrease overall, showing a downward trend as the number of refueling events increases. Finally, in the later stages (after about 10th refueling events), the distance fluctuates around 0.4 km. Figure A4b in Appendix Online Appendix B shows the corresponding figure that uses months relative to the move on the x-axis.

VI Conclusion

This article has asked what generates the information frictions consumers face, focusing on their unfamiliarity with the markets in which they buy. Using individual-level refueling records in Japan and within-driver variation in market familiarity from long-distance travel and residential relocation, I find that unfamiliar drivers pay 2.45 yen per liter (1.9%) more than familiar drivers in the same city–week. Isolating the experiential channel, this *experience premium* is 1.52 yen per liter–62% of the overall friction–and it decays as consumers accumulate experience.

These findings carry implications across three dimensions: welfare implications, policy response, and managerial strategy. First, consider the magnitude. For retail gasoline alone, the overall information friction implies excess consumer expenditure on the order of 2.9 billion yen per year, of which the experience premium accounts for roughly 1.8 billion yen. More broadly, because gasoline is only a small share of the goods for which consumers repeatedly choose among retailers, similar frictions are likely to arise in other high-frequency purchase categories in e-commerce or brick-and-mortar retail, so the aggregate cost of unfamiliarity is likely to be much larger.

Second, for policy, what matters is not only the size of the friction but also its source, because experience-based frictions and stable consumer heterogeneity carry different deep parameters, different welfare interpretations, and different responses to policy. If high prices reflected stable heterogeneity in the opportunity cost of time, information policies would have limited bite, because the consumers who pay more would be precisely those who choose not to access available information. Because a majority of the friction I measure instead arises from unfamiliarity that fades with experience, interventions that speed up learning—price-comparison tools and better market information—can improve consumer welfare.

Third, variation in information frictions also has implications for managerial strategy. Although such practices are not yet widespread in Japan, gas stations may be able to design more targeted strategies, such as personalized recommendations or individualized discounts, by exploiting consumer-level information. For example, loyalty-card-linked purchase histories could help firms construct measures of consumers' market familiarity and identify consumers likely to face experience-based search frictions, whereas vehicle- and income-related information could help firms infer heterogeneity in the opportunity cost of time. Consumer-level data may thus allow firms to target consumers along both dimensions.

These findings come with limitations that point to natural directions for future research. Because my data and model do not identify the structural parameters underlying each channel, the analysis stops short of a full decomposition of the mechanisms. Nonetheless, the reduced-form evidence points to the presence of learning (non-stationarity) and provides useful guidance for identifying structural search models. A key next step is to clarify the learning mechanism that operates through repeated purchases—a dimension that is not captured by the stationary model.

A second direction is to endogenize firms' pricing behavior, which my demand-side analysis takes as given. Experience is interesting not only because it changes how consumers search, but also because it changes the demand that firms face, and hence the prices they set. Linking my results to firms' strategic behavior would clarify the broader implications for industrial organization and competition policy. Finally, how information treatments affect search and help mitigate the penalty associated with unfamiliarity is an important open question. Overall, incorporating experience dependence into models of search behavior is a promising avenue for deepening the understanding of the managerial and policy implications.

References

D. A. Ackerberg. Advertising, learning, and consumer choice in experience good markets: an empirical examination. *International Economic Review*, 44(3):1007–1040, 2003.

M. R. Baye, J. Morgan, P. Scholten, et al. Information, search, and price dispersion. *Handbook on economics and information systems*, 1:323–375, 2006.

R. Benabou. Search market equilibrium, bilateral heterogeneity, and repeat purchases. *Journal of Economic Theory*, 60(1):140–158, 1993.

D. P. Byrne and N. de Roos. Consumer search in retail gasoline markets. *The Journal of Industrial Economics*, 65(1):183–193, 2017.

D. P. Byrne and N. De Roos. Start-up search costs. *American Economic Journal: Microeconomics*, 14(2):81–112, 2022.

A. Chandra and M. Tappata. Consumer search and dynamic price dispersion: an application to gasoline markets. *The RAND Journal of Economics*, 42(4):681–704, 2011.

J. Dorsey, A. Langer, and S. McRae. Fueling alternatives: Gas station choice and the implications for electric charging. *American Economic Journal: Economic Policy*, 17(1):362–400, 2025.

A. Düll, H. Karle, S. Martin, and H. Schumacher. Evaluating search cost models: Estimation and prediction. 2025.

Y. Fan, Z. Fan, and Y. Zhou. Dynamic search in a non-stationary search environment: An application to the beijing housing market. Technical report, National Bureau of Economic Research, 2025.

K. Fischer, S. Martin, and P. Schmidt-Dengler. Indirect taxation in consumer search markets: The case of retail fuel. Technical report, Centre for Economic Policy Research, 2024.

R. Friberg, E. M. S. Halseth, F. Steen, and S. A. Ulsaker. Do informed consumers pay less? evidence from a survey with linked grocery purchase data. *NHH Dept. of Economics Discussion Paper*, (15), 2025.

J. B. Gelbach. When do covariates matter? and which ones, and how much? *Journal of Labor Economics*, 34(2):509–543, 2016.

A. Goldman and J. K. Johansson. Determinants of search for lower prices: An empirical assessment of the economics of information theory. *Journal of Consumer Research*, 5(3):176–186, 1978.

C. Hodgson and G. Lewis. You can lead a horse to water: Spatial learning and path dependence in consumer search. Technical report, National Bureau of Economic Research, 2023.

H. Hong and M. Shum. Using price distributions to estimate search costs. *The RAND Journal of Economics*, 37(2):257–275, 2006.

E. Honka, A. Hortaçsu, and M. Wildenbeest. Empirical search and consideration sets. In *Handbook of the Economics of Marketing*, volume 1, pages 193–257. Elsevier, 2019.

P. Jindal and A. Aribarg. The importance of price beliefs in consumer search. *Journal of Marketing Research*, 58(2):321–342, 2021.

C. R. Knittel and S. Tanaka. Fuel economy and the price of gasoline: Evidence from fueling-level micro data. *Journal of Public Economics*, 202:104496, 2021.

S. Koulayev. Search with dirichlet priors: estimation and implications for consumer demand. *Journal of Business & Economic Statistics*, 31(2):226–239, 2013.

S. Lach. Existence and persistence of price dispersion: an empirical analysis. *Review of economics and statistics*, 84(3):433–444, 2002.

M. Lewis. Price dispersion and competition with differentiated sellers. *The Journal of Industrial Economics*, 56(3):654–678, 2008.

M. S. Lewis and H. P. Marvel. When do consumers search? *The Journal of Industrial Economics*, 59(3):457–483, 2011.

S. Martin. Market transparency and consumer searchvidence from the german retail gasoline market. *The RAND Journal of Economics*, 55(4):573–602, 2024.

B. Matsumoto and F. Spence. Price beliefs and experience: Do consumersbeliefs converge to empirical distributions with repeated purchases? *Journal of Economic Behavior & Organization*, 126:243–254, 2016.

J. McMillan and P. Morgan. Price dispersion, price flexibility, and repeated purchasing. *Canadian Journal of Economics*, pages 883–902, 1988.

I. Morozov, S. Malladi, and A. Martinez Marquina. Space exploration. *Available at SSRN*, 2025.

M. Nishida and M. Remer. The determinants and consequences of search cost heterogeneity: Evidence from local gasoline markets. *Journal of Marketing Research*, 55(3):305–320, 2018.

D. Pennerstorfer, P. Schmidt-Dengler, N. Schutz, C. Weiss, and B. Yontcheva. Information and price dispersion: Theory and evidence. *International Economic Review*, 61(2):871–899, 2020.

B. T. Ratchford and N. Srinivasan. An empirical investigation of returns to search. *Marketing science*, 12(1):73–87, 1993.

A. Rhodes and A. Parakhonyak. Dynamic consumer search. *SSRN*, 2024.

M. Rothschild. Searching for the lowest price when the distribution of prices is unknown. *Journal of Political Economy*, 82(4):689–711, 1974.

B. D. L. Santos, A. Hortaçsu, and M. R. Wildenbeest. Search with learning for differentiated products: Evidence from e-commerce. *Journal of Business & Economic Statistics*, 35(4):626–641, 2017.

S. Seiler and F. Pinna. Estimating search benefits from path-tracking data: measurement and determinants. *Marketing Science*, 36(4):565–589, 2017.

A. T. Sorensen. Equilibrium price dispersion in retail markets for prescription drugs. *Journal of Political Economy*, 108(4):833–850, 2000.

G. J. Stigler. The economics of information. *Journal of Political Economy*, 69(3):213–225, 1961.

L. Sun and S. Abraham. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199, 2021.

S. Tanaka. When tax incentives drive illicit behavior: The manipulation of fuel economy in the automobile industry. *Journal of Environmental Economics and Management*, 104:102367, 2020.

R. M. Ursu, T. Erdem, Q. Wang, and Q. Zhang. Prior information and consumer search: Evidence from eye tracking. *Management Science*, 70(12):8685–8708, 2024.

H. R. Varian. A model of sales. *American Economic Review*, 70(4):651–659, 1980.

M. R. Wildenbeest. An empirical model of search with vertically differentiated products. *The RAND Journal of Economics*, 42(4):729–757, 2011.

M. Wu. *Essays on digital economy*. PhD thesis, Massachusetts Institute of Technology, 2017.

X. Wu, M. S. Lewis, and F. A. Wolak. Search with learning in the retail gasoline market. *The RAND Journal of Economics*, 55(2):292–323, 2024.

Appendix A Theoretical Background

Theoretical Motivation

This section develops a simple search model to show how the higher fuel prices that drivers pay in unfamiliar areas can be interpreted as the value of market-specific experience.

I begin with a benchmark model that has no experience dependence, following Seiler and Pinna (2017) and abstracting from vertical differentiation. This is a non-sequential search model in which consumers search over prices rather than over match values.[26,27] Consider a consumer i who chooses a non-negative search intensity $s \in \mathbb{R}^+$ and in return obtains expected utility $g(s, Z_i)$, so that $g(\cdot)$ maps search intensity into expected utility.[28] I impose the two standard assumptions $g_s \geq 0$ and $g_{ss} < 0$: searching more intensively raises expected utility, but at a diminishing rate. The argument Z_i collects the primitives that determine consumer i's return to search, most importantly the price environment she faces. When goods are homogeneous and firms are identical, every station draws prices from the same distribution, $F_j = F$. Net of search costs, the consumer's expected value is

$$EV_i = g(s, Z_i) - c_i s \tag{A1}$$

where c_i represents the search cost per unit of search intensity.

I now introduce experience dependence. To keep the mechanism transparent, I let experience operate through a single channel: the search cost depends on the consumer's experience level n_{it}, where n_{it} denotes how familiar consumer i is with the market at time t.[29,30] This makes the search cost a function of experience, $c_i(n_{it})$.[31] The logic is intuitive: in a familiar market, a driver finds it easier to get around and to locate the information she needs, which lowers the cost of each additional search, so $c_i(n_{it})$ is decreasing in experience.

With this channel, the expected value becomes

$$EV_{it} = g(s, Z_i) - c_i(n_{it})s \tag{A2}$$

The optimal search intensity s^*_{it} solves the first-order condition

$$g_s(s^*_{it}, Z_i) = c_i(n_{it}) \tag{A3}$$

Note that the optimal intensity $s^*_{it} \equiv s^*(c_i(n_{it}), Z_i)$ depends on experience only through the search cost $c_i(n_{it})$. Expected utility at the optimum equals $g(s^*_{it}, Z_i)$. Realized utility differs from this because the search outcome is stochastic; I write $\epsilon_{it} \equiv u_{it} - g(s^*_{it}, Z_i)$ for this deviation.

With homogeneous products, consumers care only about price, so I define utility as the negative

[26] The model is also related to non-parametric search such as Düll et al. (2025).

[27] This non-sequential representation abstracts from within-search information acquisition, often modeled as sequential search (Rothschild 1974).

[28] Following Seiler and Pinna (2017), I drop the subscript i in g. The subsequent discussion remains valid even if g is assumed to depend on i.

[29] More precisely, experience dependence arises because consumers accumulate market-specific knowledge through past search. I assume here that this stock of knowledge is a function of the consumer's experience level.

[30] Note that, although the notation omits a market subscript, the experience variable I analyze is defined at the market level and captures each market's unfamiliarity to drivers. It does not incorporate drivers' overall experience with travel itself—for example, whether someone has seldom traveled at all.

[31] The model is myopic: consumers do not value future experience when choosing s. Wu (2017) use a similar assumption; Rhodes and Parakhonyak (2024), Fan et al. (2025) study dynamic settings.

of the price paid (lower prices are better). The realized price is therefore

$$p_{it} = -g(s^*(c_i(n_{it}), Z_i), Z_i) - \epsilon_{it} \tag{A4}$$

To obtain an estimating equation, I take a first-order Taylor expansion of p_{it} around the steady-state experience level $\hat{n}_i$, defined as the level to which experience converges after sufficiently many purchases ($\lim_{t\to\infty} n_{it} = \hat{n}_i$). A driver reaches this level once she has fully learned a market, for example long after relocating to a new city.

A first-order Taylor expansion of p_{it} around $\hat{n}_i$ then gives

$$p_{it} = -\,g\big(s^*(c_i(\hat{n}_i), Z_i), Z_i\big) + \frac{dp_{it}}{dn_{it}}\bigg|_{n_{it}=\hat{n}_i} (n_{it} - \hat{n}_i) - \epsilon_{it} + O\big((n_{it} - \hat{n}_i)^2\big).$$

Recall that the optimality condition (A3) can be written as $g_s(s^*_{it}, Z_i) - c_i(n_{it}) = 0$. Differentiating this condition with respect to n_{it} gives

$$\frac{d}{dn_{it}}\Big[g_s(s^*_{it}, Z_i) - c_i(n_{it})\Big] = g_{ss}(s^*_{it}, Z_i)\,\frac{\partial s^*_{it}}{\partial c_i}\,c_i'(n_{it}) - c_i'(n_{it}) = 0,$$

which implies $\partial s^*_{it}/\partial c_i = 1/g_{ss}(s^*_{it}, Z_i)$. Therefore, the derivative of the realized price p_{it} with respect to n_{it} is

$$\frac{dp_{it}}{dn_{it}} = -\,g_s\big(s^*_{it}, Z_i\big)\,\frac{\partial s^*_{it}}{\partial c_i}\,c_i'(n_{it}) = -\,g_s\big(s^*_{it}, Z_i\big)\,\frac{c_i'(n_{it})}{g_{ss}\big(s^*_{it}, Z_i\big)}.$$

Denoting the steady-state price by $p^0_{it} \equiv -\,g\big(s^*(c_i(\hat{n}_i), Z_i), Z_i\big) - \epsilon_{it}$ and the higher-order remainder $O\big((n_{it} - \hat{n}_i)^2\big)$ by η_{it}, and substituting $g_s(s^*_{it}, Z_i) = c_i(n_{it})$, I obtain the following estimating equation:

$$p_{it} = p^0_{it} + \underbrace{\frac{c_i(\hat{n}_i)\,c_i'(\hat{n}_i)}{g_{ss}(s^*_{it}, Z_i)}}_{\text{Marginal Effects of Unfamiliarity } (\beta_i)} \cdot \underbrace{(\hat{n}_i - n_{it})}_{\text{Unfamiliarity}} + \eta_{it} \tag{A5}$$

where p^0_{it} is the price the driver would pay at her steady-state experience and η_{it} collects the higher-order terms, which I assume to be small enough to ignore. The term $\hat{n}_i - n_{it}$ measures how far current experience falls short of its steady-state level and therefore captures unfamiliarity. Crucially, this equation no longer involves the search intensity s: once unfamiliarity is observed, β_i can be estimated from the realized price alone, without observing search behavior. The empirical model assumes $E[\beta_i] = \beta$, so I do not require every driver to follow the same learning path, only that the coefficient hold on average. Identification fails if unfamiliarity is correlated with the driver-specific component of β_i. Table 6 reports heterogeneity analyses based on observable characteristics; the interaction terms are null, providing no evidence of systematic heterogeneity in the coefficient along these dimensions.

The coefficient β_i is the marginal effect of unfamiliarity, and it is determined entirely by how the search cost responds to experience. Because experience lowers the search cost ($c_i' < 0$) and search has diminishing returns ($g_{ss} < 0$), the product $c_i(n_{it})\,c_i'(n_{it})/g_{ss}$ is positive. Hence $\beta_i > 0$: the less familiar a driver is with a market, the more she pays.

For simplicity the model lets experience act only through the search cost, but experience could also reshape the price environment Z_i itself, for instance as drivers learn the market-level price distribution (Matsumoto and Spence 2016). Moreover, as in Wu (2017), when stations are not identical the relevant object becomes a vector of station-specific price distributions, $\mathbf{F}_i = (F_{i1}, \ldots, F_{iL})$, where L is the number of gas stations in the market; learning about this environment would let experience operate through $\boldsymbol{Z}_i$ as well. Besides, search intensity might be a vector, such as the set of stations to inspect. However, incorporating these elements would complicate the model.

"Selection" Problem

It is worth explaining why the coefficient in Equation (1) does not consistently estimate the experience premium. Following the notation of Section Appendix A, let $p_{it}(c_i(n_{it}))$ denote the price paid by driver i at time t, written as a function of her search cost $c_i(n_{it})$, which in turn depends on her experience level n_{it} in the market she visits at time t.

The price paid can then be written as

$$p_{it}(c_i(n_{it})) = p^0_{it}(c_i(\hat{n}_i)) + \beta U_{it} + \eta_{it}, \tag{A6}$$

where β is the causal effect of unfamiliarity on the price paid: when a driver's experience in the market she visits falls below her steady-state level, her search cost rises and she pays more. The experience premium is exactly this β. Without individual fixed effects, α_1 in Equation (1) is identified by comparing, within a city-week, the prices paid by unfamiliar drivers with those paid by familiar ones, that is, by comparing different people. But these people differ not only in their unfamiliarity $U_{it} = \hat{n}_i - n_{it}$ but also in person-specific traits. In particular, drivers seen in unfamiliar areas are disproportionately frequent travelers, who tend to have a higher opportunity cost of time and hence a higher baseline search cost. I make this heterogeneity explicit by writing the search cost as $c_i(n_{it}) = c(n_{it}) + \zeta_i$, where $c(\cdot)$ is the common search-cost function from Section Appendix A and ζ_i is a person-specific shifter. Such a driver pays a price determined by the search cost $c(n_{it}) + \zeta_i$ rather than $c(n_{it})$, so that $p_{it}(c(n_{it}) + \zeta_i) > p_{it}(c(n_{it}))$. Because ζ_i is positively correlated with unfamiliarity across drivers, $\mathrm{Cov}\big(p_{it}(c(n_{it}) + \zeta_i) - p_{it}(c(n_{it})),\ \hat{n}_i - n_{it}\big) > 0$, so α_1 is biased upward as an estimate of β and is inconsistent. Adding individual fixed effects, as in Equation (2), removes this bias: each driver is then compared only with herself, so that α_2 identifies β.

Online Appendix

(Not for Publication)

Online Appendix A   Data

Sample Selection

I exclude outlier observations based on the following criteria:

- **Highway Refueling**
  Refueling events occurring on highways are excluded.[1]
- **Price Outliers**
  - Observations with prices more than four times the interquartile range above the 75th percentile or below the 25th percentile.
  - Refueling records where regular gasoline prices are below 75 yen or above 175 yen, or where premium gasoline prices are below 85 yen or above 190 yen.
  - Cases in which the price deviates by more than 15 yen from the state-level half-week listed price.
- **Transaction-Level Outliers**
  - Missing fuel quantity.
  - Total payment exceeding 20,000 yen.
  - Fuel quantity greater than 100 liters.
  - Fuel quantity less than 2 liters.
- **Trip-Level Mileage Outliers**
  - Refueling intervals exceeding 75 days.
  - Travel distance between refueling events exceeding 1,000 km.
  - Travel distance between refueling events less than 20 km.
  - Average daily mileage (distance divided by days between refueling events) exceeding 300 km.
  - Average daily mileage less than 1 km.
- **Vehicle-Level Mileage Outliers**
  - Average daily mileage across all trips for a vehicle exceeding 200 km.

[1] Specifically, stations whose names contain "PA" (Parking Area) or "SA" (Service Area), which typically correspond to highway truck stops, are treated as highway stations.

In the first analysis, to exclude drivers who consistently refuel at gas stations far from their residence—such as those doing so during commutes—I exclude drivers whose median observed distance to gas stations exceeds 50 kilometers. In the second analysis, I do not apply this criterion, as distance information is not used in that analysis. In addition, I exclude samples that lack brand or store information in order to keep the data consistent between the event study and the decomposition analysis.

Other Data Sources

In addition to the E-nenpi data, I use the following supplementary data sources in the analysis:

1. Crude Oil Price: Daily trends in the WTI crude oil spot price.
2. State-level Semiweekly Retail Gasoline Prices: Data from the petroleum price survey conducted by the Agency for Natural Resources and Energy of Japan.[2]
3. Wholesale Gasoline Prices: Also from the petroleum price survey by the Agency for Natural Resources and Energy of Japan.
4. City-level Geographical Area: Data from the "Statistical Reports on the Land Area by Prefectures and Municipalities in Japan," published by the Geospatial Information Authority of Japan.
5. Consumer Price Index (CPI): The April 2019 monthly CPI published by the Statistics Bureau of Japan.
6. Number of commuters: The 2020 Population Census of Japan.
7. Number of tourists: The Japan Travel and Tourism Association's Digital Tourism Statistics Open Data for 2021.

External Validity

The self-selection of respondents raises concerns about external validity, as the data are drawn from drivers who voluntarily participate in the application. To assess the representativeness of the sample, I compare it with a nationally representative dataset of Japanese drivers from official household surveys. These surveys report an average daily vehicle kilometers traveled (VKT) of 23.37 km during the relevant period, which is lower than the average in my dataset. To examine whether this difference in VKT affects the estimated experience premium, Table A1 presents a heterogeneity analysis that splits the sample into the top and bottom 50% of VKT. The results show no substantial difference between the two groups. Moreover, the mean VKT in the bottom-half sample is 23.52 km, which is close to the national average, and the estimated effects remain statistically significant.

[2] I need to account for city-level price distributions, but Japan lacks public station-level posted price data. The large transaction panel nevertheless allows city-week and station fixed effects.

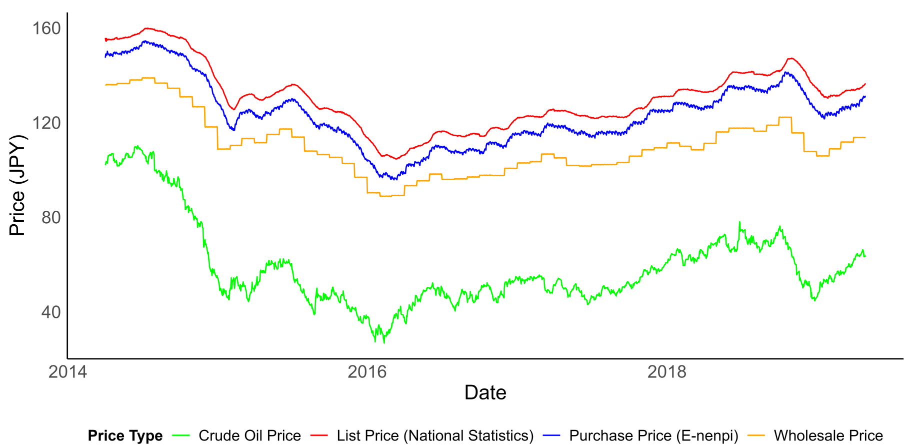
Table A1: Price Paid and Refueling in Unfamiliar Areas: External Validity

	VKT: Top Half	VKT: Bottom Half
Unfamiliar Dummy	1.5312	1.3670
	(0.0272)	(0.0680)
Observations	482460	276882
$R^2_{within,adj}$	0.01	0.00
FE: Individual	Yes	Yes
FE: City-by-Week	Yes	Yes
FE: Day	Yes	Yes

Note: This table reports a heterogeneity analysis assessing external validity. The sample is split into the top and bottom 50% by average daily vehicle kilometers traveled (VKT). The dependent variable is the price paid, and each column estimates Equation (2) with individual, city-by-week, and day fixed effects. Standard errors are clustered at the car level. $R^2_{within,adj}$ represents the adjusted within R-squared.

Gasoline Price Trends

Figure A1: Timeseries Plot of Regular Gasoline Prices



Note: This figure plots weekly time series of regular gasoline prices: the actual retail prices paid from the E-nenpi data and the week-level listed prices from national statistics.

Source of Price Dispersion

I discuss the drivers of price dispersion. To do so, I run a series of simple regressions of prices on wholesale costs and some fixed effects, which is shown in Table A2. Column (1) controls for the daily crude oil price and city-by-week fixed effects, which together explain 89.2% of the price variation. This suggests that price variation is mainly driven by city-level and temporal changes in wholesale costs. Column (2) adds brand fixed effects. Although brand is the main source of vertical differentiation in Japan, the brand fixed effects do not account well for the observed price dispersion.[3] Column (3) introduces 3 km mesh fixed effects to control for area-level differentiation.

[3] One source of brand-level price persistence is brand-affiliated price promotions linked to membership cards. Moreover, centralized price setting is uncommon in Japan; instead, individual store owners set prices based on daily demand and day-to-day changes in marginal costs.

For example, gasoline prices may be higher in tourist areas. These fixed effects again do not adequately explain the price dispersion. Column (4) adds store-level fixed effects. The portion of variation explained by store-level fixed effects corresponds to the spatial price dispersion described in Stigler (1961). For instance, differences in marginal costs across stores could explain store-level price persistence. Although this specification yields better predictive performance than the brand and 3 km mesh fixed effects, it still does not fully account for price variation, and 7.8% of the variation remains, which can be attributed to *"Temporal"* price dispersion. In Column (5), I use the Z-score as a dependent variable and include gasoline station fixed effects in the regression. Variation in the Z-score indicates a shift in a station's position within the price distribution, which is conceptually similar to the price ranking changes discussed in Chandra and Tappata (2011). The result suggests that gasoline station fixed effects explain only 26.3% of the variation in the Z-score, indicating that substantial temporal variation remains.

Table A2: Source of Price Dispersion: Regressions with Cost and Fixed Effects

Dependent Variables:		Price Paid			Z Score
	(1)	(2)	(3)	(4)	(5)
Crude Price	0.0313	0.0325	0.0327	0.0345	
	(0.0037)	(0.0037)	(0.0034)	(0.0032)	
Observations	2,313,563	2,206,558	2,313,563	2,313,563	2,130,686
R^2	0.892	0.894	0.907	0.922	0.263
FE: City-by-Week	Yes	Yes	Yes	Yes	
FE: Brand		Yes			
FE: 3km mesh			Yes		
FE: GS				Yes	Yes

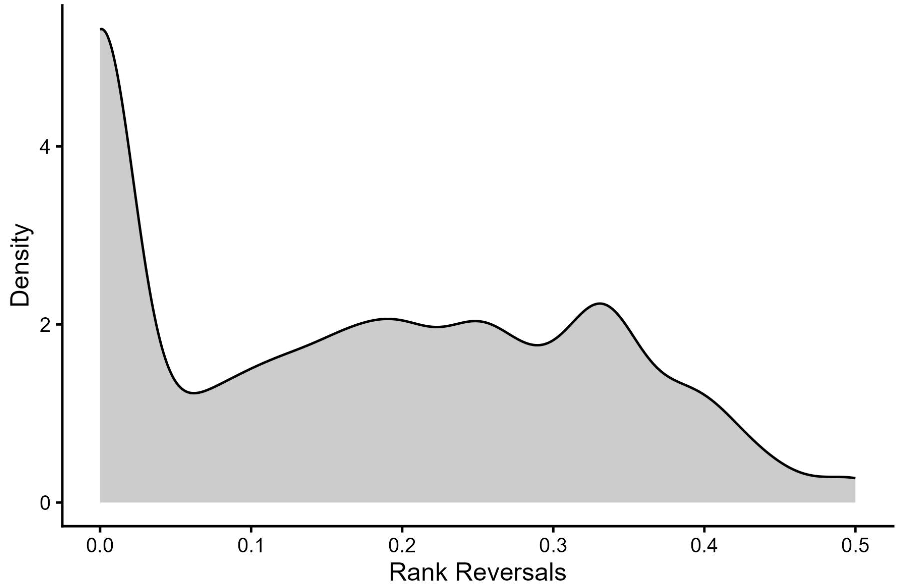
Note: This table reports the results of regressions of the price paid or the Z-score on various control variables. The Z-score is calculated using the mean and the variance at the city-week level based on the data sample. Specifically, the jackknife mean is employed. "Crude Price" indicates the national daily crude oil price. Column (1) includes 298,812 city-by-week fixed effects, Column (2) adds 17 brand fixed effects, Column (3) adds 9,262 fixed effects at the 3 km mesh level, and Column (4) adds 21,616 gasoline station fixed effects. Column (5) uses the Z-score of the price as the dependent variable and includes gasoline station fixed effects.

As further evidence of temporal price dispersion, I replicate the rank-reversal analysis proposed by Chandra and Tappata (2011). I calculate rank-reversal statistics by measuring the percentage of days on which the typically lower-priced firm sets a higher price, showing that nearby gas stations frequently switch which one charges more. Specifically, the statistic is constructed as follows:

$$r_{ij} = \frac{1}{T_{ij}} \sum_{t=1}^{T_{ij}} \mathbb{1}_{\{p_{jt} > p_{it}\}}$$

I construct this statistic for all possible pairs of stations within each 3 km mesh. Figure A2 shows a histogram of the rank reversals for regular gasoline. The results suggest that 50% of the pairs have a nonzero rank-reversal rate. This proportion is higher than that reported by Chandra and Tappata (2011). However, the overall shape of the distribution is similar.

Figure A2: Temporal Price Dispersion (Distribution of Rank Reversals)



Note: This figure illustrates the distribution of rank reversal rates, as defined in Online Appendix A, for regular gasoline. In calculating the rank reversal rate, I consider only pairs with more than two pair-by-day observations. The station pairs are constructed within each 3 km mesh.

Table A3 presents the summary statistics for both regular and premium gasoline. The results show that the average rank-reversal rates for regular and premium gasoline are 0.172 and 0.146, respectively. These figures are comparable to those reported by Chandra and Tappata (2011), who finds rates of 0.149 and 0.123 for the corresponding fuels. The results also indicate that the average price spread between two gas stations is not negligible.

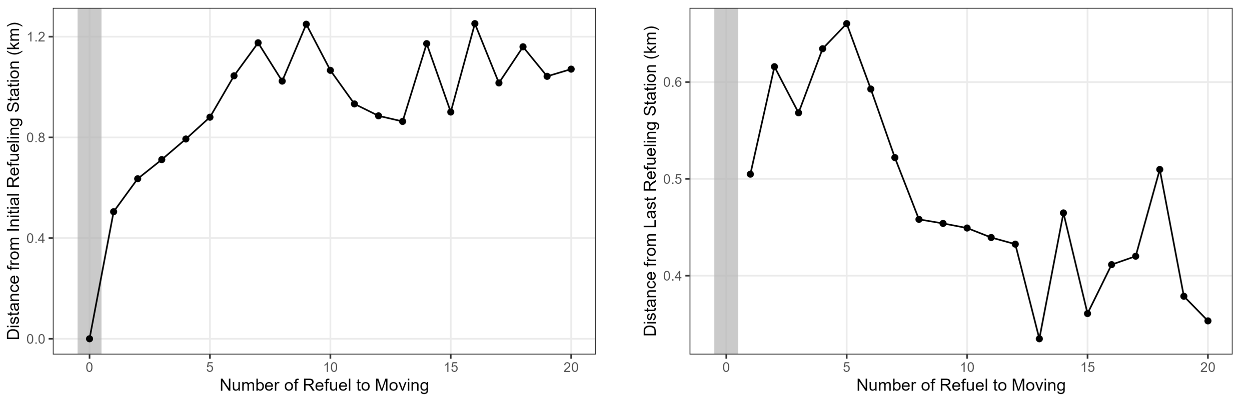
Table A3: Summary Statistics of Rank Reversals

	Regular	Premium
Avg. rank reversal	0.172	0.146
Avg. spread	2.575	2.114
Observations	7098	1877

Note: This table presents summary statistics of rank reversal rates. "Avg. rank reversal" indicates the average rank reversal rate for each fuel type, while "Avg. spread" represents the average price difference between each pair of stations. In calculating the rank reversal rate, I include only pairs with more than two pair-by-day observations. The station pairs are constructed within each 3 km mesh.

Online Appendix B Additional Results

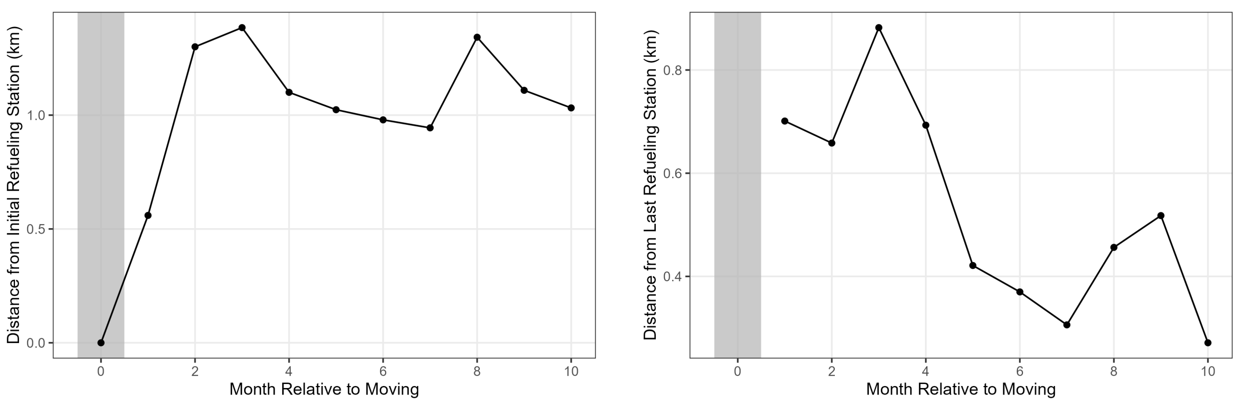
Spatial Exploration



(a) Distance from Initial Refueling Station **(b)** Distance from Last Refueling Station

Figure A3: Spatial Exploration Patterns

Note: This figure illustrates the distances from the initially chosen refueling station and the most recently chosen refueling station, respectively, after housing relocation. The x-axis denotes the number of refueling events following relocation.



(a) Distance from Initial Refueling Station **(b)** Distance from Last Refueling Station

Figure A4: Spatial Exploration Pattern (Relative Month)

Note: This figure illustrates the distances from the initially chosen refueling station and the most recently chosen refueling station, respectively, after housing relocation. The x-axis represents the number of months elapsed relative to the move.

Time Scarcity and Urgency

Table A4 examines whether the experience premium reflects trip-specific opportunity costs or urgency rather than information frictions, because drivers traveling away from home may simply be less willing to search. Column (1) reports the baseline estimate from Equation (2). Column (2) interacts the unfamiliar dummy with an indicator for long holiday periods, defined as stretches of more than three consecutive holidays, when travel is more likely to be leisure-related and thus to carry higher opportunity costs. Column (3) interacts it with an indicator for above-median tourist density, measured by the ratio of tourists to commuters using private cars. Columns (4) and (5) interact it with indicators for above-median mileage since the last refueling and above-median purchase quantity, respectively, capturing differences in urgency during non-routine travel. Across all columns the coefficient on the unfamiliar dummy stays close to the baseline, so the experience premium is not primarily driven by heterogeneity in trip-specific opportunity costs or urgency.

Table A4: Price Paid and Refueling in Unfamiliar Areas: Trip-Specific Opportunity Costs and Urgency

	(1)	(2)	(3)	(4)	(5)
Unfamiliar Dummy	1.5166	1.5466	1.3384	1.5514	1.5467
	(0.0488)	(0.0523)	(0.0614)	(0.0678)	(0.0628)
Unfamiliar Dummy × Long Holiday Period		−0.1358			
		(0.0791)			
High Tourist Density			−0.0745		
			(0.0627)		
Unfamiliar Dummy × High Tourist Density			0.3800		
			(0.0816)		
Long Mileage Since Last Refuel				0.0592	
				(0.0112)	
Unfamiliar Dummy × Long Mileage Since Last Refuel				−0.0630	
				(0.0714)	
Large Purchase Quantity					0.0257
					(0.0127)
Unfamiliar Dummy × Large Purchase Quantity					−0.0545
					(0.0714)
Observations	759371	759371	750196	759371	759371
$R^2_{within,adj}$	0.01	0.01	0.01	0.01	0.01
FE: Individual	Yes	Yes	Yes	Yes	Yes
FE: Date	Yes	Yes	Yes	Yes	Yes
FE: City-by-Week	Yes	Yes	Yes	Yes	Yes

Note: This table investigates the role of trip-specific opportunity costs and urgency in explaining the price premium paid when refueling in unfamiliar areas. Each interaction regressor is the unfamiliar dummy multiplied by a binary indicator, defined as follows. Column (1) reports the baseline estimate. Column (2) interacts the unfamiliar dummy with the Long Holiday Period indicator, which equals one for a refueling that takes place during a stretch of more than three consecutive holidays, when travel is more likely to be leisure-related and thus to carry higher opportunity costs. Column (3) adds a High Tourist Density indicator and its interaction with the unfamiliar dummy; this indicator equals one when the area's tourist density—measured as the ratio of tourists to commuters using private cars—is above the sample median. Columns (4) and (5) augment the baseline with a Long Mileage Since Last Refuel indicator and a Large Purchase Quantity indicator, respectively, along with their interactions with the unfamiliar dummy. The former equals one when the distance traveled since the previous refueling is above the sample median; the latter equals one when the purchased fuel quantity is above the sample median. Both proxy for greater urgency during non-routine travel. All specifications include individual, city-by-week, and day fixed effects. Standard errors are clustered at the car level. $R^2_{within,adj}$ represents adjusted within R^2.

Within City Variations

This analysis builds on the baseline specification in Equation (2). Whereas my main estimate relies on across-market variation, Table A5 allows both within- and across-market variation in experience by using the cumulative number of visits to a city as the key regressor. “N” is the total number of visits a driver makes to the city, and “Cum” is the cumulative visit count at the time of a given refueling. Holding the total number of visits N fixed and increasing the cumulative count traces out within-city experience evolution, whereas variation across N captures across-market differences. The coefficients change little as the cumulative count rises within a given N, so within-city experience accumulation does not show up. This pattern is consistent with the initial-state problem.

Table A5: Price Paid and Refueling in Unfamiliar Areas: Within City Variations

	(Within)
N = 1, Cum=1	0.0151
	(0.0002)
N = 2, Cum=1	0.0105
	(0.0006)
N = 2, Cum=2	0.0106
	(0.0006)
N = 3, Cum=1	0.0102
	(0.0009)
N = 3, Cum=2	0.0083
	(0.0009)
N = 3, Cum=3	0.0077
	(0.0009)
Observations	759371
$R^2_{within,adj}$	0.01
FE: Individual	Yes
FE: City-by-Week	Yes
FE: Day	Yes

Note: This table shows the relationship between the price paid in unfamiliar areas using within-city variation. “N” represents the total number of visits to the city. “Cum” represents the cumulative number of visits to the city. Standard errors are clustered at the car level. $R^2_{within,adj}$ represents the adjusted within R-squared.

Other Control Variables

This analysis builds on the baseline specification in Equation (2), augmenting it with, or substituting in, additional control variables. Column (1) uses city-by-day fixed effects instead of city-by-week fixed effects and the national crude oil price. Although within-day price changes are uncommon in Japan, the differences in refueling times between unfamiliar and familiar drivers could help explain observed price differences. Column (2) includes city-by-day-by-timezone fixed effects to examine whether differences in refueling timing between experienced and inexperienced consumers matter. Column (3) uses national daily crude oil price instead of day fixed effects. Column (4) uses the market-week level average and variance of gasoline prices calculated from the data sample. This is used instead of market-city level fixed effects. Interestingly, the coefficient on price dispersion is negative and statistically significant. As shown in Baye et al. (2006), the expected total cost (price paid) is lower when prices are more dispersed in the sense of a mean-preserving spread. This theoretical prediction is borne out in the data: market-level price variance is negatively related to the price paid by individuals.

Table A6: Price Paid and Refueling in Unfamiliar Areas: Various Control Variables

	(1)	(2)	(3)	(4)
Unfamiliar Dummy	1.4928	1.4427	1.4995	1.8339
	(0.0762)	(0.0556)	(0.0227)	(0.0195)
Crude Price			0.0478	
			(0.0029)	
Average City Price				0.6009
				(0.0013)
Variance City Price				−0.0534
				(0.0023)
Observations	759371	759371	759371	713713
$R^2_{within,adj}$	0.01	0.01	0.01	0.24
FE: Individual	Yes	Yes	Yes	Yes
FE: Day				Yes
FE: City-by-Day-by-Time		Yes		
FE: City-by-Week			Yes	
FE: City-by-Day	Yes			

Note: This table shows the relationship between the price paid at the time of refueling and whether the refueling occurred in an unfamiliar area, with various control variables. The dependent variable is the price paid. Each observation is at the car-station-day level. I categorize refueling times into four time zones: Night (00:00–06:00), Morning (06:00–12:00), Afternoon (12:00–18:00), and Evening (18:00–24:00), and use the city-day-time combination as fixed effects. The average city price is calculated using the jackknife mean. Standard errors are clustered at the car level. $R^2_{within,adj}$ represents the adjusted within R-squared.

Loyalty Program

I control for brand familiarity as a proxy for brand promotion. Specifically, I add a "First Brand Dummy" equal to one if the driver has previously used the refueling brand. Table A7 reports the results, with Column (1) corresponding to Equation (1) and Column (2) to Equation (2). The coefficient on the unfamiliar dummy remains positive and significant after controlling for brand familiarity, so the premium is not merely a loyalty discount.

Table A7: Price Paid and Refueling in Unfamiliar Areas: Loyalty Effect

	(1)	(2)
Unfamiliar Dummy	2.3741	1.4343
	(0.1221)	(0.0499)
First Brand Dummy	0.5171	0.6277
	(0.0775)	(0.0359)
Observations	720365	720365
$R^2_{within,adj}$	0.01	0.01
FE: Individual		Yes
FE: City-by-Week	Yes	Yes
FE: Day	Yes	Yes

Note: This table examines whether the experience premium reflects brand loyalty. The dependent variable is the price paid. "First Brand Dummy" equals one if the driver has previously used the refueling brand, a proxy for brand familiarity. Column (1) corresponds to Equation (1) and Column (2) to Equation (2), which adds individual fixed effects. Standard errors are clustered at the car level. $R^2_{within,adj}$ represents the adjusted within R-squared.

Ideally I would absorb loyalty discounts with station-by-day-by-oil fixed effects, but only 0.2% of observations exhibit within station-by-day-by-oil variation in the unfamiliar dummy, so such fixed effects would absorb nearly all of the identifying variation. Restricting the sample to those observations, Table A8 shows that experience-based price differences do arise within a station on a given day for a given oil type (Column (1)), but that the brand-familiarity proxy absorbs them: once the First Brand Dummy is added, the unfamiliar-dummy coefficient becomes small and statistically insignificant (Column (2)).

Table A8: Within-Station, Same-Day Price Differences by Familiarity

	(1)	(2)
(Intercept)	125.6334	125.5253
	(0.6274)	(0.6250)
Unfamiliar Dummy	1.2195	0.2935
	(0.2386)	(0.3465)
First Brand Dummy		8.1362
		(1.8089)
Observations	1469	1469
R^2_{adj}	0.00	0.02

Note: This table reports OLS estimates of within–station-by-day-by-oil price differences by familiarity. The sample is restricted to the observations that exhibit within–station-by-day-by-oil variation in the "Unfamiliar Dummy." Both columns are estimated by OLS, include station-by-day-by-oil fixed effects, and the dependent variable is the price paid. Column (2) adds the "First Brand Dummy," equal to one if the driver has previously used the refueling brand. Standard errors are clustered at the station-by-day-by-oil level. R^2_{adj} denotes the adjusted R-squared.

Spatial and Vertical Differentiation

I examine whether the premium paid in unfamiliar areas reflects product differentiation rather than imperfect information, separating a spatial and a vertical component. For the spatial component, Table A9 adds 3 km mesh fixed effects to control for persistent price differences across micro-locations within a city: Column (1) augments Equation (1) and Column (2) augments Equation (2). The coefficient on the unfamiliar dummy falls by roughly 20% to 30% but remains large and statistically significant, so within-city location differences account for only part of the premium.

Table A9: Price Paid and Refueling in Unfamiliar Areas: Differences by Refueling Location

	(1)	(2)
Unfamiliar Dummy	1.9085	1.1179
	(0.1138)	(0.0427)
Observations	759371	759371
$R^2_{within,adj}$	0.00	0.00
FE: Individual		Yes
FE: City-by-Week	Yes	Yes
FE: Day	Yes	Yes
FE: 3km mesh	Yes	Yes

Note: This table examines whether the price premium in unfamiliar areas reflects within-city, area-level price differences. The dependent variable is the price paid. Both columns add 3 km mesh fixed effects to control for area-level differentiation: Column (1) corresponds to Equation (1) and Column (2) adds individual fixed effects as in Equation (2). Standard errors are clustered at the car level. $R^2_{within,adj}$ represents the adjusted within R-squared.

For the vertical component, Table A10 adds brand fixed effects to absorb brand-level quality differences, corresponding to the case in which consumers have homogeneous preferences over quality. Column (1) augments Equation (1) with brand fixed effects, and Column (2) does the same for Equation (2). In both columns the coefficient on the unfamiliar dummy is essentially unchanged and stays positive and statistically significant, so brand-level vertical differentiation does not account for the premium.

Table A10: Price Paid and Refueling in Unfamiliar Areas: Controlling for Brand-Level Vertical Differentiation

	(1)	(2)
Unfamiliar Dummy	2.4025	1.4810
	(0.0465)	(0.0228)
Observations	731284	731284
$R^2_{within,adj}$	0.01	0.01
FE: Individual		Yes
FE: Day	Yes	Yes
FE: City-by-Week	Yes	Yes
FE: Brand	Yes	Yes

Note: This table assesses whether the experience premium is driven by brand-level vertical differentiation. The dependent variable is the price paid, and the key independent variable is the Unfamiliar Dummy. Both columns add brand fixed effects to absorb brand-level quality differences, corresponding to homogeneous preferences over quality: Column (1) corresponds to Equation (1) and Column (2) adds individual fixed effects as in Equation (2). Each observation is at the car-station-day level. Standard errors are clustered at the car level. $R^2_{within,adj}$ is the adjusted within R-squared.

Event Study

Table A11 reports the housing-relocation event study estimates plotted in Figure 1 in tabular form, for the OLS (two-way fixed effects) and the Sun and Abraham (2021) estimators. The table lists the point estimate and clustered standard error for every event-time coefficient. The point estimates are close to zero in the pre-relocation periods, jump to about 0.6% at the month of the move ($k = 0$), and decay back toward zero within a few months, mirroring the pattern in the figure.

Table A11: Price Paid and Housing Relocation: Event Study Estimates

	(1) OLS		(2) Sun-Abraham	
	Est.	S.E.	Est.	S.E.
$\gamma_{\leq -5}$	-0.0013	0.0014	-0.0012	0.0011
γ_{-4}	0.0008	0.0017	0.0006	0.0016
γ_{-3}	-0.0007	0.0016	0.0008	0.0015
γ_{-2}	-0.0003	0.0015	-0.0009	0.0014
γ_{0}	0.0065	0.0021	0.0059	0.0021
γ_{1}	0.0044	0.0018	0.0032	0.0017
γ_{2}	0.0028	0.0019	0.0016	0.0019
γ_{3}	0.0004	0.0018	-0.0011	0.0019
γ_{4}	0.0012	0.0020	-0.0002	0.0019
$\gamma_{\geq 5}$	0.0010	0.0017	0.0002	0.0015
Observations	101420		101420	
FE: Individual	Yes		Yes	
FE: Date	Yes		Yes	
FE: City-by-Week	Yes		Yes	

Note: This table reports the housing-relocation event study estimates γ_k from Equation (4) in tabular form, corresponding to Figure 1. For each estimator the point estimate ("Est.") and its standard error ("S.E.") are reported side by side. Column (1) is the ordinary least squares (two-way fixed effects) estimator; Column (2) is the interaction-weighted estimator of Sun and Abraham (2021), which uses never-treated units as controls. The dependent variable is the logarithm of the price paid. The event-time index k measures the number of months relative to housing relocation, with $k = 0$ the month of the move; the coefficient at $k = -1$ is normalized to zero. $\gamma_{\leq -5}$ and $\gamma_{\geq 5}$ bin all periods more than four months before and after the move, respectively. The control group consists of non-movers observed in the same city-by-week as the movers. All specifications include individual, city-by-week, and day fixed effects. Standard errors are clustered at the car level.

The same urgency concern can be addressed within the housing-relocation event study. Table A12 interacts each event-time dummy from Equation (4) with an above-median mileage-since-last-refuel indicator (Column (2)) and an above-median purchase-quantity indicator (Column (3)), both proxies for greater urgency during non-routine travel. The interaction terms are small and not jointly distinguishable from zero, whereas the post-relocation experience premium remains, so the dynamic premium is not primarily driven by heterogeneity in trip-specific urgency.

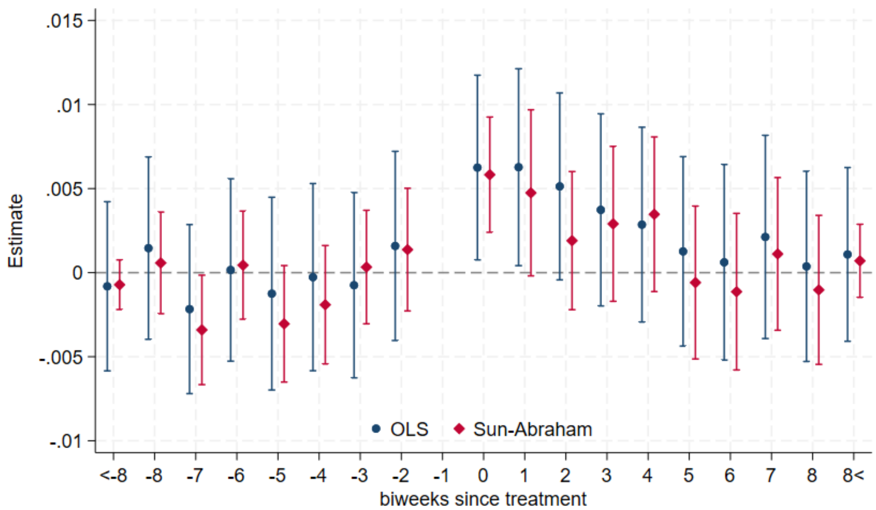
Table A12: Housing Relocation Event Study: Urgency Robustness

	(1) Baseline		(2) Long Mileage		(3) High Quantity	
	Est.	S.E.	Est.	S.E.	Est.	S.E.
$\gamma_{\leq -5}$	−0.0013	0.0014	−0.0013	0.0014	−0.0013	0.0014
γ_{-4}	0.0008	0.0017	0.0008	0.0017	0.0009	0.0017
γ_{-3}	−0.0007	0.0016	−0.0007	0.0016	−0.0006	0.0016
γ_{-2}	−0.0003	0.0015	−0.0003	0.0015	−0.0003	0.0015
γ_{0}	0.0065	0.0021	0.0040	0.0025	0.0037	0.0028
γ_{1}	0.0044	0.0018	0.0038	0.0021	0.0041	0.0021
γ_{2}	0.0028	0.0019	0.0024	0.0022	0.0039	0.0024
γ_{3}	0.0004	0.0018	0.0019	0.0022	0.0034	0.0024
γ_{4}	0.0012	0.0020	0.0003	0.0026	−0.0011	0.0025
$\gamma_{\geq 5}$	0.0010	0.0017	0.0009	0.0019	0.0008	0.0021
Urgency Proxy (main effect)			0.0003	0.0002	−0.0003	0.0002
$\gamma_{0}\times$ Urgency Proxy			0.0039	0.0030	0.0042	0.0031
$\gamma_{1}\times$ Urgency Proxy			0.0012	0.0021	0.0008	0.0023
$\gamma_{2}\times$ Urgency Proxy			0.0007	0.0023	−0.0020	0.0025
$\gamma_{3}\times$ Urgency Proxy			−0.0034	0.0019	−0.0054	0.0023
$\gamma_{4}\times$ Urgency Proxy			0.0019	0.0025	0.0049	0.0026
$\gamma_{\geq 5}\times$ Urgency Proxy			0.0000	0.0015	0.0003	0.0018
Observations	101420		101420		101420	
FE: Individual	Yes		Yes		Yes	
FE: Date	Yes		Yes		Yes	
FE: City-by-Week	Yes		Yes		Yes	

Note: This table examines whether the post-relocation experience premium estimated in the event study of Equation (4) reflects differences in urgency rather than information frictions. For each specification the point estimate ("Est.") and its standard error ("S.E.") are reported side by side. The dependent variable is the logarithm of the price paid, and all specifications are estimated by ordinary least squares (two-way fixed effects). Column (1) reproduces the baseline event study estimates. Column (2) interacts each event-time dummy with the "Long Mileage" indicator, which equals one when the distance traveled since the previous refueling is above the sample median. Column (3) interacts each event-time dummy with the "High Quantity" indicator, which equals one when the purchased fuel quantity is above the sample median. Both indicators proxy for greater urgency during non-routine travel. "Urgency Proxy" denotes the Long Mileage indicator in Column (2) and the High Quantity indicator in Column (3); the interaction rows $\gamma_k\times$ Urgency Proxy report the corresponding moderator for each column. The event-time index k measures the number of months relative to relocation, with $k = 0$ the month of the move and the coefficient at $k = -1$ normalized to zero. The interaction terms are small and not jointly distinguishable from zero, so the premium is not primarily driven by heterogeneity in trip-specific urgency. All specifications include individual, city-by-week, and day fixed effects. Standard errors are clustered at the car level.

As a robustness check on the timing of the experience premium, Figure A5 re-estimates the housing-relocation event study using two-week periods, rather than calendar months, as the unit of event time. The estimated path is broadly similar to the monthly specification in Figure 1, so the results are not sensitive to the choice of time aggregation.

Figure A5: Price Paid and Housing Relocation (Two-Week Periods)



Note: This figure plots OLS and Sun and Abraham (2021) estimates of the event-time coefficients, with point estimates and 90% confidence intervals; the coefficient for the final period before relocation is normalized to zero. The dependent variable is the logarithm of p_{ijd}, and the specification includes individual, city-by-week, and day fixed effects. The treatment group consists of movers selected as in Section II. As the control group, I use non-movers observed in the same city-by-week as the movers. Observations corresponding to refueling in cities other than the most familiar and the second most familiar cities are excluded. The sample window covers ten months before and after the relocation.