Consumer welfare and price controls in gasoline markets

Armando R. Colina

May 18, 2023

Abstract

I estimate the consumer welfare impact of removing a price control policy on retail gasoline in Mexico. I use the features of a highly regulated industry as well as a two-tiered pricing regime in two cities to identify the demand of gasoline for heterogeneous consumers and spatially differentiated goods. I estimate households' price sensitivity and their implicit valuation for convenience. I find that for every peso gained in welfare from increased product availability, consumers lose two pesos due to the increase in prices. The aggregate impact in these cities is a loss of 1.4 billion $^{MXN}/_{year}$, roughly 7% of yearly revenue. High-income households bear the brunt of the price increases since they overwhelmingly consume most of the gasoline. However, as a proportion of income, lower-income households are affected the most. City-year price elasticity is between -0.42 and -0.64, this is 13% to 72% higher than previous estimates for the U.S.

Contents

1 Introduction **4**

2 Institutional background and data **6**

3 The Model **14**

3.1 Aggregation 17

4 Identification Strategy **18**

4.1 Instruments 20

4.1.1 BLP instruments 20

4.1.2 Waldfogel-Fan instruments 21

4.2 Price variation 21

5 Model estimation results **22**

6 Post estimation calculations **26**

6.1 Predicting market shares 28

7 Consumer welfare impacts **30**

7.1 Scenario 1: Consumer welfare change from additional product availability . . 32

7.2 Consumer welfare change from price changes 33

7.2.1 Retail pricing rule 34

7.2.2 Scenario (2): Passive Ministry of Finance 36

7.2.3 Scenario (3): Proactive Ministry of Finance 37

8 Conclusion **42**

9 Glossary **45**

10 Appendix **46**

10.1 Tables and Charts 46

10.2 Estimation procedure 47

10.3 Initial values 49
10.4 Model Alternatives 49
10.5 Alternative model specifications: standard Logit. 50

1 Introduction

This paper measures the consumer welfare effects of liberalizing retail gasoline prices in 2016 in Mexico. After almost eight decades of a federally mandated retail price policy, gasoline retailers were allowed to choose their pricing strategies. This led to some gas station openings in anticipation of price liberalization. However, retail prices increased much more than changes in wholesale prices for the 2016 to 2019 period as entry was curtailed from regulatory backlogs. Higher product availability coupled with higher prices makes the consumer welfare impact ambiguous. To compute the welfare effects, I estimate the demand for gasoline as a spatially differentiated good.

I find that the market level elasticity of demand is -0.42 for Mexicali and -0.64in Tijuana, 80% higher than estimates for the U.S. consumers (Coglianese et al. 2017 and Colina 2023). Mexico is characterized by a wide disparity in gasoline consumption, the top three income deciles consume roughly 60% of gasoline while the bottom three consume 13%. Despite the disparity, low-income households benefit the most from new gas station openings as it increases overall tax transfers. High-income households are the most affected by price increases. Overall, I estimate a loss of welfare of 1,43 billion $^{\text{MXN}}/_{\text{year}}$ for these two markets, 7.1% of the markets' annual revenue.

Demand estimation has its challenges. In most markets, firms simultaneously choose their pricing and product attributes when introducing a product into a market (Borden 1965). This has long been recognized in economic models (Hotelling 1929) and empirical work that does not take this simultaneity into consideration risks of having biased estimates (Berry and Haile 2021). This paper relies on an unusual setting to identify structural parameters. Prior to 2016 the industry had federally mandated price controls and lacked branding differentiation. Therefore, retail stations, while individually owned and operated, were forced to charge the same price despite their location and local competitive pressures. Additionally, retailers operated under a franchise system in which the sole wholesaler was Petróleos Mexicanos (PEMEX), forcing every station to carry the same brand and sell the same quality of products.

I use a uniquely detailed data set of cross-sectional yearly gas station level sales, attributes, and locations for 2015 for the universe of gas stations and census data to estimate the structural demand parameters of the model. Despite a unique and detailed data set, I do not observe quantity data post-2016, however, I observe the pricing and location of all gas stations, incumbents, and entrants. Through a random coefficients model, I estimate demand in the highly regulated period of 2015 and then simulate consumer choices after prices were liberalized and new gas stations entered the market.The random coefficients model is based on the work of Berry et al. 1995 but with a spatial component like Thomadsen 2005 and Davis 2006. This model allows for better substitution patterns across products as it recognizes consumer heterogeneity in their preferences as discussed in Gandhi and Nevo 2021 and in Berry and Haile 2021.

This is the first paper to estimate the city-level retail demand for gasoline as a geographically differentiated good accounting for heterogeneous consumer preferences. I find that the city-year elasticity of demand is between -0.42 and -0.64. Yet, there is substantial heterogeneity in price sensitivity depending on the metropolitan area's demographic characteristics like household income and the number of children: households with higher levels of income and more kids are less price-sensitive to gasoline prices. For example, the median consumer is willing to drive an additional kilometer to save $13^{MXN¢}/_{L}$ ($3\ ^{US¢}/_{gal.}$). Consumers who are less price sensitive need a higher compensation to be indifferent to drive, for example, consumers located in the top 25th percentile of the price-sensitivity distribution would require savings of $21^{MXN¢}/_{L}$ ($5\ ^{US¢}/_{gal.}$) to be indifferent to drive this extra distance. This has important implications for the provision of public transportation (disproportionately used by lower-income individuals) and electric vehicle (EV) charging stations (disproportionately used by high-income individuals in Mexico).

In section 2, I describe the institutional setting and the data that I use. In section 3, I present the model and the structural parameters of interest. Section 4 talks about the identification strategy and section 5 shows the estimation results.

2 Institutional background and data

Despite being the sixth largest consumer of gasoline in the world, Mexico's retail gasoline market was characterized by heavy regulation that forced gas station operators to charge administered prices, sell the same fuels, and carry the same brand regardless of their location. However, after the implementation of Mexico's Energy Reform (MER) several restrictions were sequentially lifted culminating in 2017 when gas station operators could carry their own brands, sell different products, and determine their own prices.

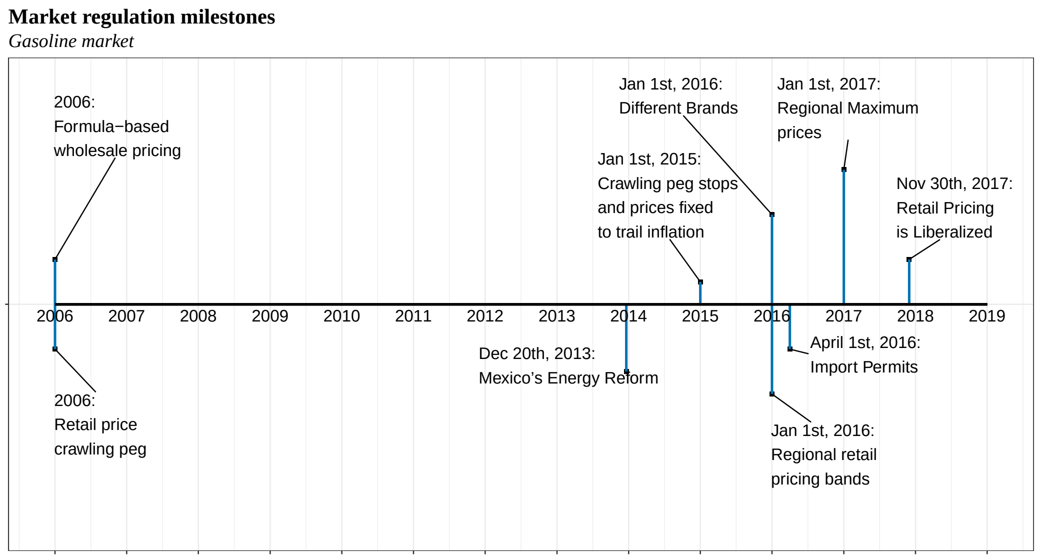


Figure 1: Recent regulatory developments in Mexico's retail gasoline industry

Figure 1 shows a timeline of regulatory changes in the industry's recent history. Gas stations in Mexico operate as franchises that are individually owned and operated. However, between 1938 and 2015, PEMEX, Mexico's national oil company, was enshrined by law as the only gasoline wholesaler, importer, and distributor of gasoline (Diario Oficial de la Federación 1938, Diario Oficial de la Federación 1940, and Diario Oficial de la Federación 1995). Consequently, all gas stations across the country carried the exact same fuels and had the same branding. Additionally, retail and wholesale price decisions were not left either to PEMEX or to the gas station operators. Instead, the industry was under an administered-price regime where prices were determined at the federal level by the Ministry of Finance (MoF) (Secretaría de Hacienda y Crédito Público 2015). This setting contrasts to the setting in most other OECD-member countries in which retailers choose their branding, as well as different pricing and quality schemes (The Organisation for Economic Co-operation and Development- Secretariat 2013).

There were four distinct pricing regimes in the retail gasoline market from 1995 onward.

1. **One country - one price (1995 to Oct 2014)**. During this period the MoF would present the fiscal budget to the Lower House of Congress (Cámara de Diputados). In doing so, it would commit to a pan-national retail price policy for gasoline over the year. Some years this price was held constant throughout the whole year, and some years it followed a steady rise (Diario Oficial de la Federación 1980).

   The difference between the committed price and the international wholesale price would be subsidized or taxed accordingly. By law, retailers had to buy from PEMEX at the national wholesale price (plus taxes/subsidies) and would have to sell at the retail price level targeted by the MoF.

2. **One country - two pricing zones (Sep 2014 to Dec 2015)**. After September 2014, the MoF established two pricing zones (Secretaría de Hacienda y Crédito Público 2015). The first zone consisted of gas stations located 20 km from the U.S.-Mexico border. They were assigned a price based on the prices of the relevant metropolitan area on the U.S. side of the border (see Figure 2). For example, the assigned price for gas stations in Tijuana was set to resemble prices in San Diego, prices in Mexicali were based on the prices in Calexico, and prices in Ciudad Juárez were based on prices in El Paso. Prices were adjusted for inflation throughout the year.

   The second zone consisted of the rest of the country below the 20 km line. Similar to the "One country-one price" price regime, the rest of the gas stations would have to charge the same price regardless if they were located downtown on a busy intersection or on the fringe of the city.

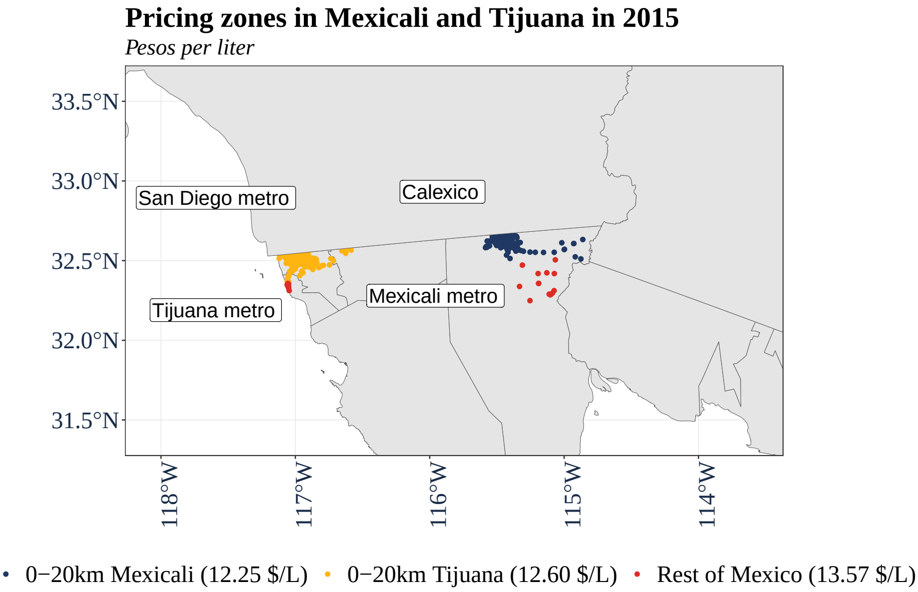


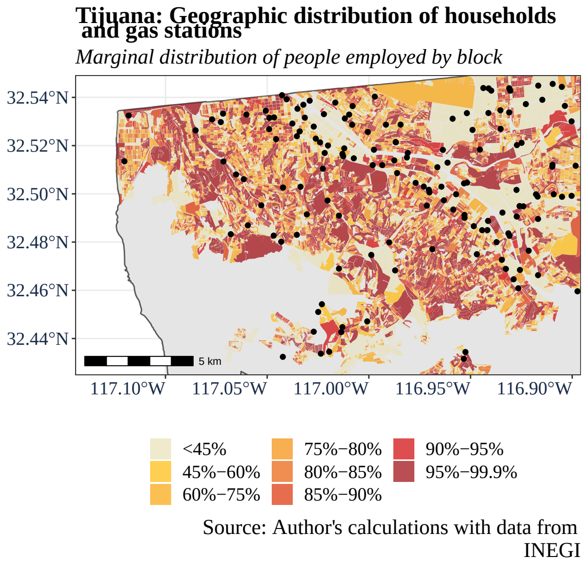
Figure 2: Cross-section of gas station locations in different pricing bands for administered retail gasoline prices

3. **One country - several pricing zones (Jan 2016 to Nov 2017)**. Beginning in 2016 different pricing zones were established, and with them, a pricing band for maximum and minimum prices (Diario Oficial de la Federación 2014 and Comisión Reguladora de Energía 2016). During this period, retailers were allowed to carry non-PEMEX brands, and gasoline imports were liberalized.

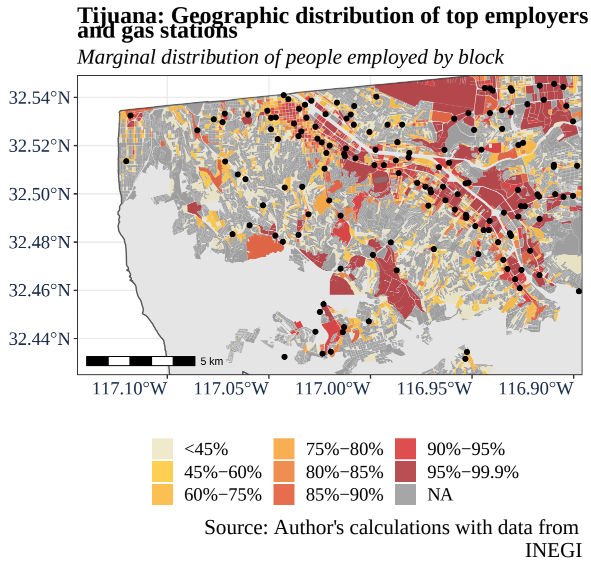
4. **Localized competition (Nov 2017 - currently)**. Finally, from November 2017 to date, pricing bands were sequentially removed until retail prices were completely liberalized across Mexico. Currently, gas station operators are free to determine their pricing strategies (Comisión Reguladora de Energía 2017).

This paper studies retail gasoline demand in 2015 when all gas stations in Mexico carried the same brand, sold the same fuels, and prices were determined at the federal level by the MoF and not by local competitive pressures. However, the existence of different prices across the two pricing zones allows me to identify the price elasticity of demand given the data and the model.

Gasoline is a relatively homogeneous product (Gary et al. 2007), however, for the end



(a) Household density and gas station locations



(b) Employment density and gas station locations

Figure 3: Location of gas stations in comparison to household location and employment location

consumer, it is a spatially differentiated non-durable good since the consumer needs to travel to a given station on a regular basis to fill up their tank, usually on the way to work or shopping (Kitamura and Sperling 1987b). To account for this feature in the retail gasoline market I have the location of the universe of gas stations in Mexico.[1] Additionally, to account for the geographic distribution and demographic characteristics of the population, I use the 2000 and 2010 censuses that provide block-level data, see the bottom entries of Table 1. This data is provided by Instituto Nacional de Estadística y Geografía (INEGI), Mexico's statistics office. I also use the 2014 economic census to identify which gas stations are close to major businesses.

To model consumer heterogeneity, I use the city-level distribution of income and the number of children across households. I use INEGI's 2014 income and expenditure survey. Because income-expenditure surveys are carried out every two years I use the 2014 survey to estimate data in my observation period. See table 2 for a summary of the distribution of income and children per household.

[1]The data is provided by Comisión Reguladora de Energía (CRE), one of Mexico's energy regulators

In 2015 gas stations were not allowed to carry different brands other than PEMEX. Despite this, each gas station offered its customers attributes beyond just gasoline. For example, gas stations could offer convenience store, oil change services, ATMs on-site, online invoicing services, acceptance of gasoline vouchers, etc. These data were manually collected for every gas station in Mexico from "Guía Pemex", PEMEX's official mobile app with the universe of gas stations in Mexico in 2015, their location, and attributes. See Table 1.

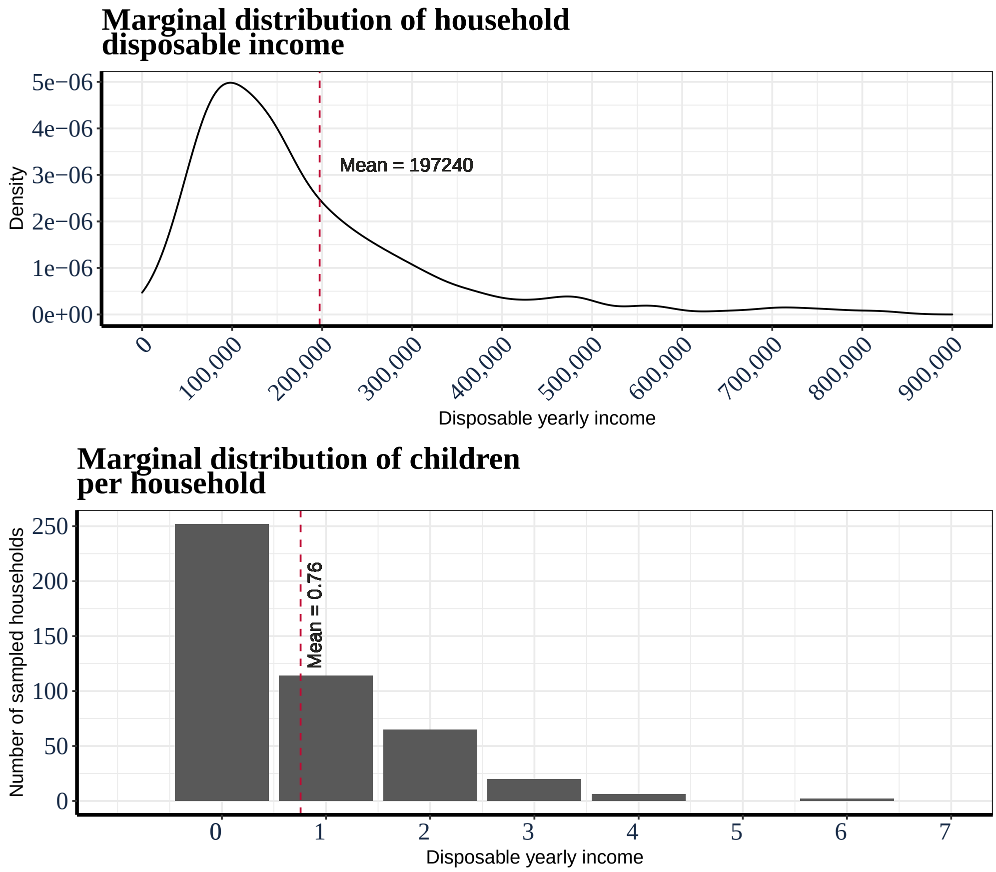


Figure 4: Upper panel: The distribution of households' annual income for 2014. Lower panel: The distribution of the number of children per household in 2014.

INEGI reports monthly retail gasoline prices for different cities, in conjunction with the pricing rules for 2015 and the location of gas stations, I can map the price each gas station was charging in 2015 (See Figure 2).

I have yearly sales for all gas stations in the North-Pacific region of Mexico in 2015. The data is provided by Next-level Consulting, a consulting firm that works alongside PEMEX on permits, regulation, and opening of new stations. Expenditure in public transportation comes from INEGI's 2014 Consumer expenditure survey. The categories in the survey include expenditure on subway and light trains, bus, trolleybus, metrobus, vans, taxis, and intercity bus service, as well as fuels. See Table (2) for further details.

Variable	N	Mean	St. Dev	Min.	Max
Markets	2				
... Mexicali	224	45.6%			
... Tijuana	267	54.4%			
Sales (1,000 MXN/year)	467	37,062	25,443	1,512	199,641
Sales (1,000 L/year)	467	2,952	2,002	120	15,372
Market Shares (%)	467	0.372	0.231	0.011	1.5
Pricing band	491				
... 0-20km from U.S. border	455	92.7%			
... 20+km from U.S. border	36	7.3%			
Avg. distance from consum. (km)	491	3.012	0.646	0.584	4.827
Avg. distance sq. from consum. (km^2)	491	10.825	3.362	0.484	23.313
Convenience Store (Yes/No)	491				
... Yes	108	22%			
ATM on site (Yes/No)	491				
... Yes	6	1.2%			
Offers gas vouchers (Yes/No)	491				
... Yes	14	2.9%			
Close to big business (Yes/No)	491				
... Yes	208	42.4%			
Rack	467				
... TAD Mexicali	221	47.3%			
... TAD Rosarito	246	52.7%			

Table 1: Summary statistics of gas station characteristics

Another important feature to highlight is how the consumption of gasoline is distributed across households of different income levels. In Mexico, the distribution of consumption is very uneven. For example, in the cities of Mexicali and Tijuana, the households in the top three income deciles consume close to 52% of all the liters sold whereas the households in the bottom three deciles consume less than 13% of annual sales. See table 3 for additional information. In addition, within each income decile, there are several households that do

	Yearly income	Yearly expenditure		Household characteristics	
Income decile group	Current income	Public transport	Fuel	Cars owned by household	Children in household
Mexicali					
.. *Decile 1*	59,522	2,206	2,014	0.21	0.59
	(6,362)	(3,877)	(3,354)	(0.43)	(1.09)
.. *Decile 5*	128,329	4,051	7,155	0.71	0.80
	(5,912)	(7,079)	(5,479)	(0.64)	(0.91)
.. *Decile 10*	331,649	1,833	20,228	1.73	0.30
	(39,233)	(4,702)	(9,175)	(1.13)	(0.48)
Tijuana					
.. *Decile 1*	50,644	2,578	1,436	0.18	0.63
	(9,906)	(4,049)	(3,054)	(0.39)	(0.86)
.. *Decile 5*	127,406	6,844	7,307	0.35	0.76
	(5,858)	(8,244)	(7,620)	(0.56)	(0.98)
.. *Decile 10*	327,369	2,985	15,629	0.90	0.73
	(31,558)	(6,283)	(11,824)	(0.79)	(1.00)

Note: I report the average value for households in each income decile.
The standard deviation is reported in parenthesis.

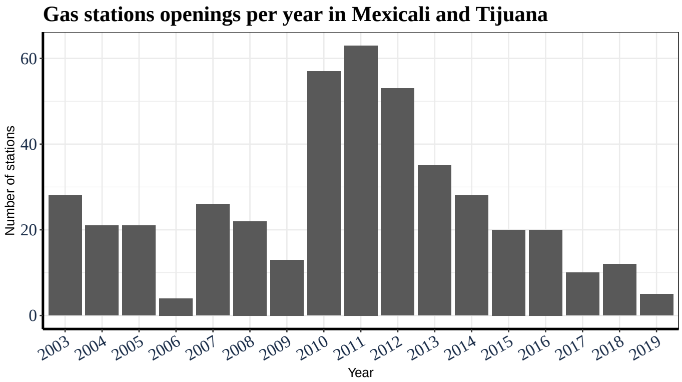
Table 2: Distribution of household characteristics according to income decile

not consume gasoline at all. Nevertheless, these households benefit from the taxation of gasoline consumption as taxes are redistributed in form of government expenditure.

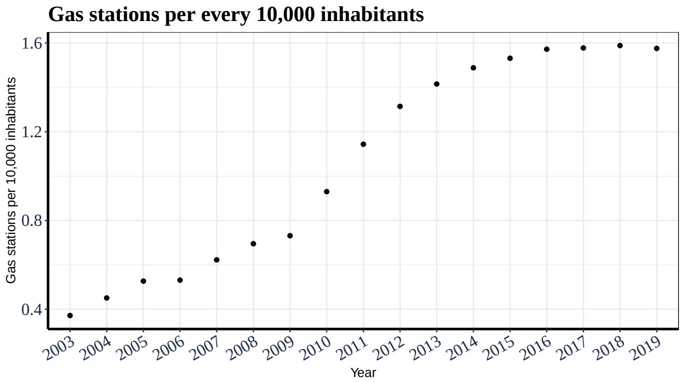
Liters of gasoline consumed per income decile in Mexicali and Tijuana		
Income decile	*Liters (million/year)*	*Percentage*
1	6.9	1.6%
2	20.4	4.8%
3	27.8	6.5%
4	22.4	5.3%
5	43.0	10.1%
6	41.6	9.8%
7	35.7	8.4%
8	57.0	13.4%
9	88.6	20.8%
10	81.6	19.2%

Table 3: Gasoline consumption is very unevenly distributed across households with different levels of income.

A major development in MER was that from 2013 to 2017 changes in regulation followed a pre-established timeline. One of the main changes was allowing Comisión Reguladora de Energía (CRE), one the main energy regulators, to operate as an autonomous entity from the executive branch. However, in 2018, a new President was sworn in and, soon after, political



(a) Prior to the passing of MER, there was an increase in the number of gas station openings. Then the number of openings declined due to regulatory backlogs.



(b) The number of gas stations per 10 thousand inhabitants was 0.37 in 2003 and grew to 1.58 in 2019. The per capita number is low compared to countries without fixed-price regimes.

Figure 5: Gas station entry in Mexicali and Tijuana

pressure began for CRE to roll back permits or to stop the approval of new gas station permits (Wood 2018). In 2019 the new administration announced deep budgetary cuts to CRE for 2019. The cuts translated into 30% less operational budget and downsizing of 60% of the workforce (Arena Pública 2019). This led to numerous articles in the specialized press pointing out the increase of regulatory backlogs, amongst them, the stall in the approval process of new gas station permits (S&P Global Platts 2020, Argus Media 2020, ONEXPO Nacional 2021).

The Mexican fuel markets were left with an unusual mix of policies. Retail prices were liberalized, but entry ended up being constrained due to regulatory backlogs. Figure 5(a) shows that the number of permits approved was substantially reduced in 2017, 2018, and 2019. For Mexicali and Tijuana, the number of stations per 10 thousand inhabitants reached 1.58 in 2019 as shown in figure 5(b). For comparison purposes, the number of gas stations per 10 thousand inhabitants in the U.S. is 4.6, in Canada is 3.1, and in Brazil is 2 (American Petroleum Institute 2023, Canadian Fuels Association 2023, Agência Nacional do Petróleo, Gás Natural e Biocombustíveis 2022).

3 The Model

The model that I use to estimate demand is a random coefficients model following Berry et al. 1995, with a spatial differentiation component like Thomadsen 2005 and Davis 2006. It is a single-address model where consumers are in their home and choose a gas station to visit while having disutility for driving to get to a gas station. I do not impose the assumption that consumers are homogeneous, instead, their sensitivity to price and their idiosyncratic valuation of using a gas station depends on the level of the household income and the number of children in the household. I define a market as a metropolitan area-year unit. Metropolitan areas are defined by Consejo Nacional de Población (CONAPO). CONAPO is a governmental agency in charge of data collection and analysis of demographic phenomena in Mexico. It is comprised of staff from INEGI, SHCP, Mexico's Ministry of Environment, amongst other agencies. CONAPO's analysis is based on commuting patterns and different measures of economic integration. Suppose that in each market
$k = \{Tijuana,\ Mexicali\}$, there are $i = 1 \ldots I_k$ consumers facing a balanced choice set, where they choose among $j = 1 \ldots J_k$ gas stations. The indirect utility of each consumer is comprised of components that are both observed and unobserved to the econometrician. Then, I assume that the indirect utility is given by:

$$u_{ijk} = \gamma_i + \alpha_i(y_{ik} - p_{jk}) + x_{jk}\beta - \lambda_1 d_{ijk} - \lambda_2 d_{ijk}^2 + \xi_{jk} + \varepsilon_{ijk} \tag{1}$$

where γ_i and α_i are consumer-specific parameters and $\beta, \lambda_1, \lambda_2$ are parameters that are common to all consumers.

The variable p_j is the price at the pump for regular gasoline and y_i is the level of income of individual i. The model nests the transport utility specifications by Hotelling 1929 and D'Aspremont et al. 1979 where $d_{ijk} \equiv d(L_{ik}, L_{jk})$ is the Euclidean distance from consumer i's location to the location of gas station j. Several studies like Phibbs and Luft 1995 and Boscoe et al. 2012 show that Euclidean distance is a very close approximation to the distance traveled through the road network, even on short distances (Buczkowska et al. 2019).

Consumers do not always choose locations that are the closest to them because they derive utility from non-distance attributes (Rushton et al. 1967, Eckert 2013, Ellickson and Grieco 2020).The vector of variables x_{jk} includes observed gas station on-site attributes like a convenience store or an ATM, services like oil changes, or payment facilities like the acceptance of gas vouchers.[2]

Consumers are also known to refuel while commuting. For example, Houde 2012 uses origin/destination matrices to simulate, given the location of a consumer, the probability that they will travel through a specific route to each of the possible destinations. Additionally, Kitamura and Sperling 1987a document two salient features of refueling behavior: (1) the vast majority of consumers that refuel have home as an origin or destination, (2) yet other destinations may be visited during the trip (See table 10). To capture this feature, I include an additional variable to x_j, indicating if a gas station is close to a big business. I classify a business as a "big business" if it is among the top 10% of employers in a city. This dummy variable captures the potential demand shock from increased traffic without having to impose an *a priori* substitution structure as in a nested logit. That is $x_{jk} = \{big\_business_{jk}, convstore_{jk}, ATM_{jk}, vouchers_{jk}, oil\_change_{jk}\}$.

In the tradition of McFadden 1978, I assume that some attributes, ξ_{jk}, are unobserved to the econometrician, while the consumers and gas station operators can observe them all. Unobserved product attributes are captured by ξ_{jk} as described by Berry 1994. The variable ε_{ij} is a random utility shock with mean zero from a type-1 extreme value distribution whose draws are independent. For a given $\alpha_i, \beta_i, \xi_{jk}, \lambda_1, \lambda_2$, the probability that consumer i chooses gas station j for each market k is given by:

$$Pr(i,j|k) = \frac{exp(\gamma_i + x_{jk}\beta - \alpha_i p_{jk} - \lambda_1 d_{ijk} - \lambda_2 d_{ijk}^2 + \xi_{jk})}{1 + \sum_{g}^{J_k} exp(\gamma_i + x_{gk}\beta - \alpha_i p_{gk} - \lambda_1 d_{igk} - \lambda_2 d_{igk}^2 + \xi_{gk})} \tag{2}$$

$$\text{for } k = \{Mexicali, Tijuana\}$$

$$\begin{bmatrix} \alpha_i \\ \gamma_i \end{bmatrix} = \begin{bmatrix} \bar{\alpha} \\ \bar{\gamma} \end{bmatrix} + \Pi D_i + \Sigma \upsilon_i \tag{3}$$

[2]In Mexico, employers give gas vouchers as an employment benefit. For employers, it is easier to deduct this payment as a business expenditure.

$$E(\alpha_i) = \bar{\alpha}, \;\; E(\gamma_i) = \bar{\gamma}$$

Similarly to Nevo 2001, the model captures the heterogeneity of tastes in consumers by letting structural taste parameters be distributed according to two main components: non-parametric demographic characteristics (D_i) in each market and a vector of random utility shocks independently drawn from a standard normal (υ_i). These random utility shocks are used to capture any additional unobserved consumer characteristics. The matrix of parameters Π relates demographic draws to individual taste parameters, while Σ is a matrix that relates additional structural shocks to individual taste parameters. I will estimate the elements of these matrices to obtain a distribution of the taste parameters.

The demand system is completed with the introduction of an "outside good" (McFadden 1978). That is, the consumer may decide not to visit any gas stations and instead use public transportation. Without an outside good, a homogeneous increase in price, as in the case of a tax, would not affect the quantities demanded in equilibrium. The indirect utility from the outside good is normalized to 0.

Let the mean utility for all consumers in market k from good j be expressed as

$$\delta_{jk} \equiv \bar{\gamma} - \bar{\alpha} p_{jk} + x_{jk}\beta - \lambda_1 \bar{d_{jk}} - \lambda_2 \bar{d^2_{jk}} + \xi_{jk} \tag{4}$$

where $\bar{d_{jk}} \equiv \int d(L_i, L_j) dP^{\star}(L_i)$ is the weighted average distance between gas station j and its clients, while $dP^{\star}(\cdot)$ is the non-parametric population distribution function of consumer locations in a 5 km radius around j.[3] This cutoff is chosen following reports that, on average, commuters in Mexico City do a 10km commute each day (Guerra 2017, Moovit insights 2022), unfortunately, there is limited information available about commuting patterns in Mexicali and Tijuana.

Let $\psi_{ijk} \equiv d_{ijk} - \bar{d_{jk}}$, and $\psi^2_{ij} \equiv d^2_{ijk} - \bar{d^2_{jk}}$. I can decompose the utility of consuming good j into its average level and an individual utility shock:

[3]For the convenience of notation, I remove income from this equation since it will be differenced-out in the utility function when consumers are comparing among goods.

$$
\begin{aligned}
u_{ijk} &= \gamma_i - \alpha_i p_j + x_{jk}\beta - \lambda_1 d_{ijk} - \lambda_2 d_{ijk}^2 + \xi_{jk} + \varepsilon_{ijk} \\
&= \overbrace{\bar{\gamma} - \bar{\alpha} p_j + x_{jk}\beta - \lambda_1 \bar{d_{jk}} - \lambda_2 \bar{d_{jk}^2} + \xi_{jk}}^{\delta_{jk} \equiv \text{Average utility}} + \\
&\qquad \underbrace{(\gamma_i - \bar{\gamma}) - p_{jk}(\alpha_i - \bar{\alpha}) - \lambda_1 \psi_{ijk} - \lambda_2 \psi_{ijk}^2 + \varepsilon_{ijk}}_{\mu_{ijk} \equiv \text{Individual utility shock}} \\
&= \delta_{jk} + \mu_{ijk}
\end{aligned} \tag{5}
$$

3.1 Aggregation

I can aggregate modeled individual choices to a market level such that they match with the observed market outcomes described in section 2. The market share of gas station j is the weighted probability that consumers in a market choose that gas station.

Let $A_j^k(x, p, \bar{d}, \bar{d^2}, \delta) \equiv \{(D_i, \upsilon_i, \varepsilon_i, \psi_i, \psi_i^2 | u_{ij} \geq u_{ig} \ \forall \ g = 0, 1, \ldots, J_k\}$ be the set of individuals in market k who choose gas station j. The matrices $x, p, \bar{d}, \bar{d^2}$ include the attributes of all gas stations in each market. $dP(\cdot)$ denotes the marginal population distribution function and the third equality in equation (6) follows from an assumption of independence of D, υ and ψ, ψ^2.

Each gas station's market share can be characterized by

$$
\begin{aligned}
\tilde{s_{jk}} &= \int_{A_j^k} Pr(i, j|k) dP(\alpha_i, \beta_i) \\
&= \int_{A_j^k} \frac{exp(\delta_{jk} + \mu_{ijk})}{1 + \sum_g^{J_k} exp(\delta_{gk} + \mu_{ijk})} dP(\alpha_i, \beta_i) \\
&= \int_{A_j^k} \frac{exp(\delta_{jk} + \mu_{ijk})}{1 + \sum_g^{J_k} exp(\delta_{gk} + \mu_{ijk})} dP(\psi, \psi^2, D, \upsilon | \bar{\alpha}, \bar{\gamma}, \beta, \Pi, \Sigma) \\
&= \int_{A_j^k} \frac{exp(\delta_{jk} + \mu_{ijk})}{1 + \sum_g^{J_k} exp(\delta_{gk} + \mu_{ijk})} dP(D, \upsilon | \bar{\alpha}, \bar{\gamma}, \beta, \Pi, \Sigma) dP(\psi, \psi^2 | \bar{\alpha}, \bar{\gamma}, \beta, \Pi, \Sigma) \\
&= s(\delta_{1k} \ldots \delta_{J_k}; \bar{\alpha}, \bar{\gamma}, \beta, \lambda_1, \lambda_2, \Pi, \Sigma)
\end{aligned} \tag{6}
$$

For each market, k, the market share of station j derived from the model, denoted as $\tilde{s_j}$, holds for the J competitors and creates a system of J equations.

$$
\begin{aligned}
\tilde{s_{1k}} &= s(\delta_{1k} \dots \delta_{J_k}; \bar{\alpha}, \bar{\gamma}, \beta, \Pi, \Sigma), \\
&\vdots \\
\tilde{s_{J_k}} &= s(\delta_{1k} \dots \delta_{J_k}; \bar{\alpha}, \bar{\gamma}, \beta, \Pi, \Sigma).
\end{aligned} \tag{7}
$$

To be more precise, note that $\delta_{jk} = \delta(P_{jk}, x_{jk}, \bar{d_j}, \xi_{jk}; \bar{\alpha}, \bar{\gamma}, \beta, \Pi, \Sigma)$. Then,

$$
\begin{aligned}
\tilde{s_{jk}} &= s(\delta_{1k} \dots \delta_{jk} \dots \delta_{J_k}; \bar{\alpha}, \bar{\gamma}, \beta, \Pi, \Sigma), \\
&= s(\delta_{1k}(P_{1k}, x_{1k}, \bar{d_{1k}}, \xi_{1k}) \dots \delta_{jk}(P_{jk}, x_{jk}, \bar{d_{jk}}, \xi_{jk}) \dots \delta_{J_k}(P_{Jk}, x_{Jk}, \bar{d_{Jk}}, \xi_{Jk}); \bar{\alpha}, \bar{\gamma}, \beta, \Pi, \Sigma), \\
&= s(P_{1k}, \dots, P_{Jk}, x_{1k}, \dots, x_{Jk}, \bar{d_{1k}} \dots, \bar{d_{Jk}}, \xi_{1k}, \dots, \xi_{Jk}; \bar{\alpha}, \bar{\gamma}, \beta, \Pi, \Sigma) \ \ \forall j = 1 \dots J_k.
\end{aligned} \tag{8}
$$

From equation (7), note that there are J unobserved mean utilities $(\delta_1, \dots, \delta_J)$ plus the unknown parameters of $\bar{\alpha}, \bar{\gamma}, \lambda_1, \lambda_2, \Pi, \Sigma$, which makes parameter estimation challenging. From equation (8) we observe that each share j depends on its own observed and unobserved attributes, and also on the attributes of its competitors. This result is intuitive as we can think about gas stations being substitutes of one another. Therefore, the attributes of a competing gas station, g, will affect the residual demand that gas station j faces. See section 10.2 and section 4 for further details about how I use instruments to identify demand.

The model has no closed-form solution, therefore I need to use initial values to start the algorithm that minimizes the objective function. I use 360 combinations of parameters as initial guesses and choose the value that yields the smallest value of the objective function. For further details see the Appendix in section 10.3. As a robustness check to initial-value sensitivity I present the results of a logit with a closed-form solution in section 10.5.

4 Identification Strategy

Given the highly regulated environment of Mexico's retail gasoline industry prior to 2015, many of the traditional concerns regarding endogeneity are ameliorated. For example, vertical differentiation in the form of different quality of inputs was nonexistent since every gas station offered the same fuel and all operated on a franchise system which made gas stations have a homogeneous look. Despite the tight regulation, gas station operators were free to choose their location. It is likely that gas station operators may choose a location L_j which is in a high-traffic area. This would mean that they are closer, on average to consumers, i.e. $E(\bar{d_j}\xi_j < 0)$. Figure 3(a) and 3(b) seem to indicate that is the case. To control for this, I use data from the economic census to identify which are the businesses among the top 10% of the largest employers in Tijuana and Mexicali, then I create a dummy variable indicating if a gas station is located within 300 meters from a large employer. Another concern is that the pricing tiers were assigned based on some unobserved factors at the regional level that are systematically driving demand or the cost of logistics from the rack. Racks are nodes where fuel transportation from ports and refineries shifts to distribution. The rack price is considered by the industry as the relevant regional wholesale price (Borenstein et al. 1992 and Borenstein and Bushnell 2005). I either include rack-level fixed effects as a control or I use the distance from each gas station to its corresponding rack as an instrument for unobserved logistics costs. Figure 6 shows a graphical representation of each rack's distribution area.

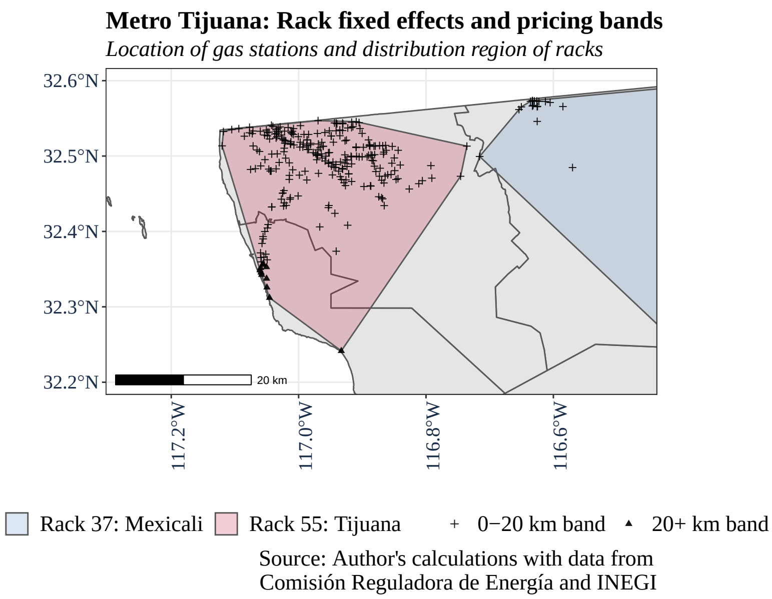


Figure 6: Distribution area for different racks

While all gas stations carried PEMEX's brand, color, and products, gas station operators were free to choose the size of the plot and the number of pumps in their business. Then, it is possible that some observed attributes are correlated with unobserved attributes. For example, gas stations with a larger plot would be more likely to have an on-site convenience store (observed) and more pumps (unobserved). To address this issue, I use different instruments.

4.1 Instruments

The first challenge comes from the necessity to instrument for both prices and quantities. In a demand system of differentiated goods, the demand for one good depends on the demand for all other goods, see equation (6). Therefore, to identify the price response of demand for good j it is essential to keep the demand for all other goods fixed.

This can be done by controlling for the observed product attributes of all other goods, yet ξ_{-jk} remains unaccounted for (see equation (4)). I need $J_k - 1$ instruments to account for the dependency of j's demand with the unobserved $\xi_{1,k}, \ldots, \xi_{j-1,k}, \xi_{j+1,k}, \ldots \xi_{J_k,k}$.

Despite dealing with a setting in which traditional differentiating attributes such as branding, advertising, and quality were closely regulated and homogenized, there might still be a concern for the endogeneity of some attributes with unobserved factors. For example, in this setting, it is likely that the presence of a convenience store or an ATM is not mean independent of other unobserved attributes. That is, I will need an additional $2 \times J_k$ instruments. I impose the common assumption that the rest of the elements of x_k are mean independent to ξ_k.

4.1.1 BLP instruments

The first set of instruments is obtained using a similar rationale to Berry et al. 1995: the demand of good j depends on the demand of all related goods. If attribute a for good g is mean independent to the vector ξ_k, then a linear transformation of this attribute can create exogenous variation in demand for good j. That is, the assumption is that through oligopolistic competition pressures, if the firm selling good g offers attribute a, the firm selling good j will react to that offering independently of its own unobserved attributes ξ_j.[4]

The closest competitors are determined by a procedure similar to Davis 2006 by considering gas stations that fall within a radius of 5km around gas station j. Then, I compute the average attributes they possess. For example, I compute the proportion of competing gas stations that have an ATM on site or that accept gas vouchers.

4.1.2 Waldfogel-Fan instruments

Another set of instruments is obtained using a similar rationale as Waldfogel 2003 and Fan 2013. Consider two distinct but adjacent neighborhoods in a city: a high-income neighborhood and a low-income neighborhood. Suppose that there are two gas stations, each mostly serving one of these neighborhoods but with partially overlapping service areas. Each gas station will offer attributes that cater to the preferences of most of their customers. However, in an effort to attract customers from a competitor, gas stations will also offer attributes that are offered by their competitors.

I use a set of instruments that consider the demographic characteristics of the clients around gas station j's competitors. I consider the same 5km radius as before to determine who the competitors are. I measure the average school years of people living around each gas station and the number of people earning at least five times the minimum wage as a proxy for the the number of high-earners in the area.

4.2 Price variation

Having a fixed price regime has substantial advantages since it eliminates the simultaneity in pricing and attributes decisions. However, this setting has its downsides. For parameter estimation purposes, the main downside of a fixed-price regime is that there is less variety of prices than in a conventional setting. In this case, there are three distinct prices: 12.25,12.60, and 13.57 pesos per liter.

[4]While this assumption might be largely unavoidable, in this setting is less troublesome than in previous research. First, pricing, which normally can be adjusted more freely, is not determined by gas station operators. Second, most attributes are fixed assets which are more difficult to adjust to own-demand shocks. Third, advertisement was almost nonexistent due to a lack of branding differentiation.

Despite less variety, the variability in prices is enough to identify the $\bar{\alpha}$ parameter. For example, a common instrument used to identify price sensitivity is the change in taxes. Li et al. 2014 explain that most tax changes are around $10^{US¢}/_{gal}$. In this setting, the price difference within each market is of $1.32^{MXN}/_{L}$ and $0.97^{MXN}/_{L}$, for Mexicali and Tijuana, respectively. This translates to $31.5^{US¢}/_{gal}$ and $23.1^{US¢}/_{gal}$, respectively.[5]

Additionally, these prices are representative of the prices that Mexican consumers had been facing. That is, they do not correspond to a time where prices were unusually low nor elevated. For example, the left panel of figure 7 shows the distribution of historical retail prices across different cities from January 2011 to right before prices were liberalized. Notice that the "low" administered prices, both for Tijuana and Mexicali, are located close to the 75th percentile. The "high" price is located at the edge of the distribution, but there is a considerable mass in that section. That is, these prices were a common occurrence in both the Mexicali and Tijuana markets. This becomes even starker when we consider the market prices observed after the price regulation was lifted. Then, the "low" prices are located near the median of the distribution whereas the "high" price is located near the 75th percentile.

5 Model estimation results

The estimations of the model parameters can be seen in table 4. The first rows of the table show the parameter estimates and standard errors for the "mean" parameters and then estimates for the parameters for the matrix Π and Σ. Further down the tables, I report the market-level elasticity of demand, savings needed to drive an additional kilometer, and how much being close to a big business increases sales, on average.

RC Model 1 and 3 include rack-level fixed effects, RC Model 2 is identical to RC Model 3, but it does not have rack-level fixed effects. The coefficients in RC Model 2 and 3 seem relatively stable as well as the estimates of market-level elasticity. This could indicate that once the demographic data is included in the model, controlling for local fixed effects is not necessary.

[5] Using an exchange rate of $15.86^{MXN}/_{USD}$, the average rate for 2015.

	RC Model 1		RC Model 2		RC Model 3	
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Parameter Name	*Estimate*	*s.e.*	*Estimate*	*s.e.*	*Estimate*	*s.e.*
Means						
.. Intercept			74.4 ***	0.002		
.. Prices	-3.6 ***	0.051	-3.96 ***	0.053	-3.73 ***	0.188
.. Avg. dist	1.18 ***	0.384	0.88 ***	0.329	1.02 ***	0.427
.. Avg dist sq.	-0.21 ***	0.075	-0.15 ***	0.064	-0.19 **	0.084
.. Big business	0.13	0.077	0.12 *	0.069	0.12	0.076
.. Conv. Store	-0.07	0.068	-0.07	0.069	-0.06	0.070
.. ATM			0.3 ***	0.084	0.3 ***	0.105
.. Accepts vouchers			0.19	0.257	0.19	0.259
.. Sells oil			-0.13	0.163	-0.12	0.169
Σ (cov. Struc.)						
.. Intercept	2.0 ***	0.019	1.984 ***	0.0004	1.98 ***	0.000
.. Intercept & prices	1.83 ***	0.373	1.708 ***	0.046	1.71 ***	0.046
.. Prices	2.4 ***	0.187	2.535 ***	0.059	2.54 ***	0.059
Π (cov demog.)						
.. Intercept-Dist. Shock	0.98 ***	0.034	0.96 ***	0.005	0.99 ***	0.064
.. Intercept-Dist. Sq Shock	0.95 ***	0.037	0.88 ***	0.009	0.94 ***	0.137
.. Interc.-Disp. Income	0.97 ***	0.014	0.98 ***	0.004	0.96 ***	0.021
.. Interc.-Num kids	0.91 ***	0.005	0.9 ***	0.007	0.89 ***	0.032
.. Price-Dist. Shock	0.14	0.155	0.27 *	0.142	0.12	0.547
.. Price-Dist. Shock sq	-0.24	0.181	-0.48 ***	0.040	-0.28 ***	0.091
.. Price-Disp. Income	0.82 ***	0.043	0.9 ***	0.036	0.73 ***	0.031
.. Price-Num children	-0.06	0.157	-0.13	0.105	-0.33	0.472
Rack-level F.E.	Yes		No		Yes	
Post-estimation calc.						
.. ΔP to drike add. 1km	-0.35		-0.24		-0.40	
	Mexicali	*Tijuana*	*Mexicali*	*Tijuana*	*Mexicali*	*Tijuana*
.. Mkt. price elast.	-0.36	-0.7	-0.42	-0.64	-0.36	-0.71
.. Avg. Market share (s)	0.36%	0.43%	0.36%	0.43%	0.36%	0.43%
.. Δs/Δd	-0.004%	-0.005%	-0.022%	-0.027%	-0.004%	-0.005%
.... Change in sales	-1.2%	-1.2%	-6.3%	-6.3%	-1.2%	-1.2%
.. Δs/Δbig business	0.045%	0.055%	0.044%	0.053%	0.046%	0.055%
.... Change in sales	12.6%	12.6%	12.3%	12.2%	12.7%	12.7%

Since the model is not linear, post estimation calculations are evaluated at the mean.
(***) indicate significance at the 2.5% level, (**) at the 5% level, and (*) at the 10% level

Table 4: Estimation results of Random Coefficient models

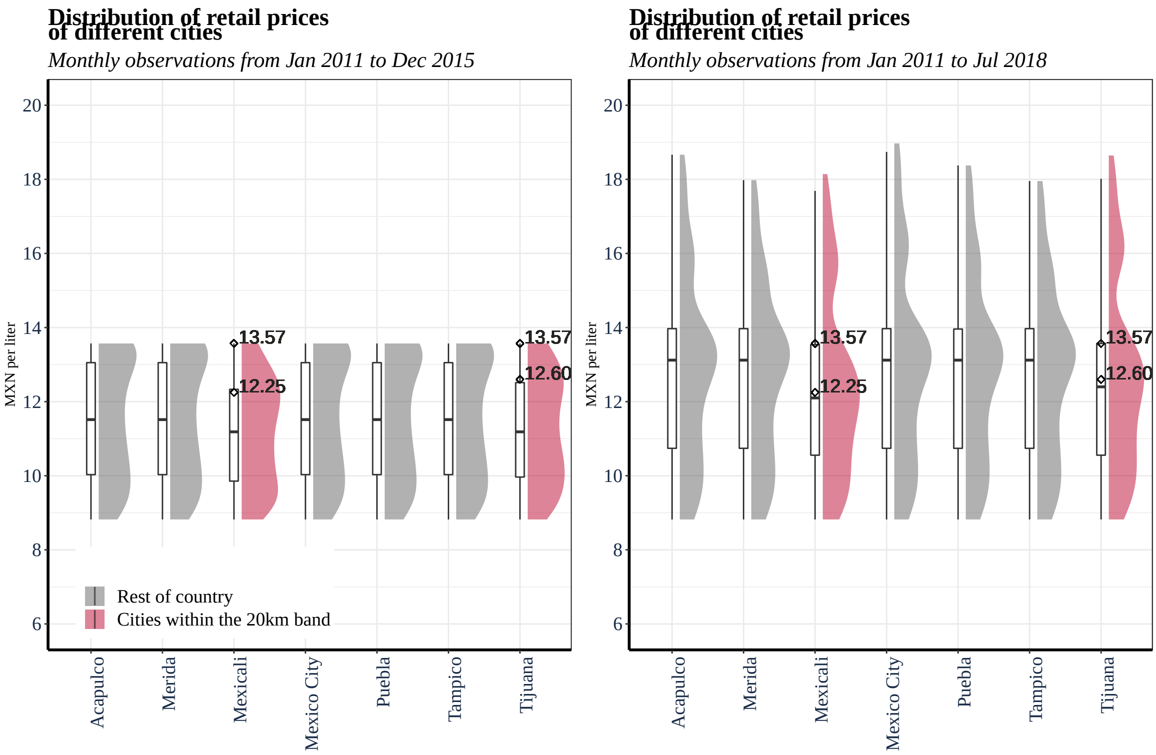


Figure 7: Distribution of historical retail prices before and after MER

My preferred model is RC Model 2. The model estimation confirms that consumers within a city have heterogeneous price sensitivity. This model's estimates show that consumers who have higher disposable income are less price sensitive. This can be seen in the estimate in the row called "Price-Disp. Income" in the section of estimates for the Π matrix. Figure 8 shows the marginal distribution of the α parameter for the city of Tijuana. Similarly, consumers have heterogeneous valuations of refueling at a gas station vs using public transport. For example, households with more children are more likely to prefer to refuel and use their car as opposed to taking a mode of public transportation (See row "Interc.-Num. kids").

These model specifications do not have a closed-form solution. Similar to Ito and Zhang 2020, I try 360 different initial values to start the numerical solution algorithm. Out of these initial values, I report the parameters that yield the smallest value for the objective

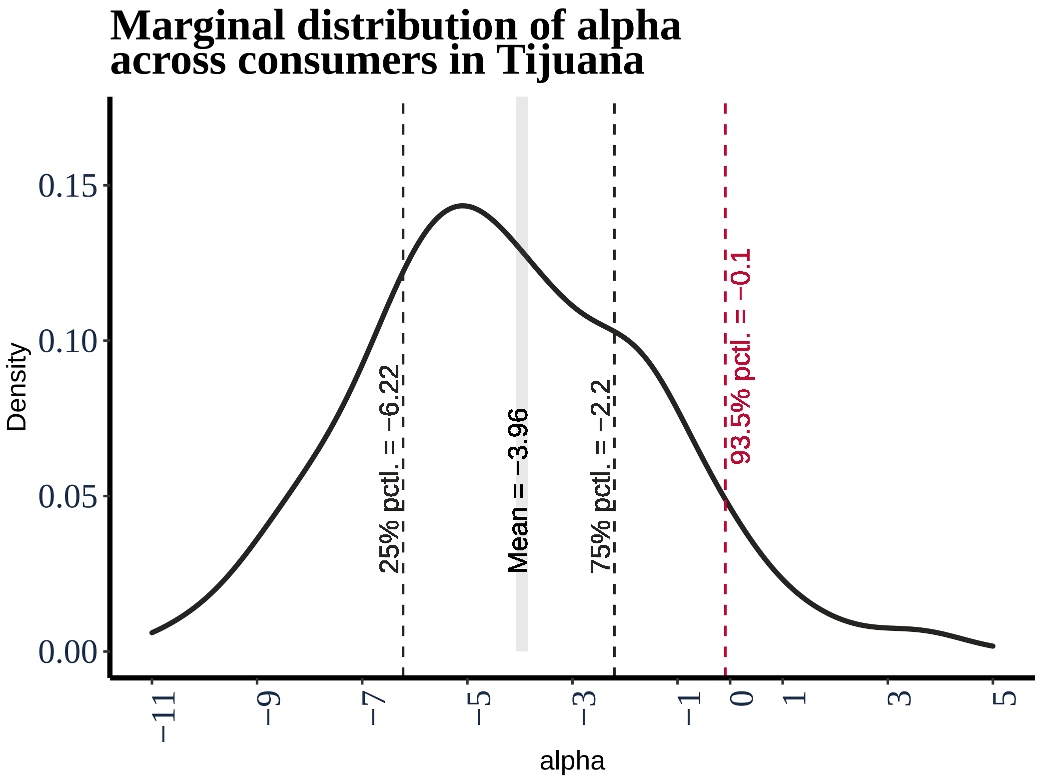


Figure 8: The marginal distribution of α shows consumer heterogeneity in price sensitivity. The mean estimate, $\bar{\alpha}$, is -3.96 with a 95% confidence interval between -4.06 and -3.85 as shown in the gray shaded area. The bottom quartile has a parameter value of -6.22 or less, while the top quartile has a value of -2.2 or greater.

function that is being minimized. Further details are reported in the Appendix in section 10.3.

As a robustness check to the model's sensitivity to initial values, I estimate a similar model to equation (1) using a logit model with a closed-form solution. To do so, I impose the assumption that consumers are homogeneous but still derive utility from all the same attributes as in equation 1. The model is in the Appendix in section 10.5 and written explicitly in equation 15.

The estimates of the logit model are reported in table 11 and are similar to the estimates found by numerical methods in the random coefficients model. In Logit Model 2, the price sensitivity parameter, α is estimated to be -3.06 and is statistically significant. This compares to an estimate of -3.96 for $\bar{\alpha}$ for RC Model 2. Similarly, the estimates Logit Model 2 show an elasticity of demand of -0.56 and -0.73 for Mexicali and Tijuana, respectively. Meanwhile, RC Model 2 estimates elasticities to be -0.42 and -0.64.

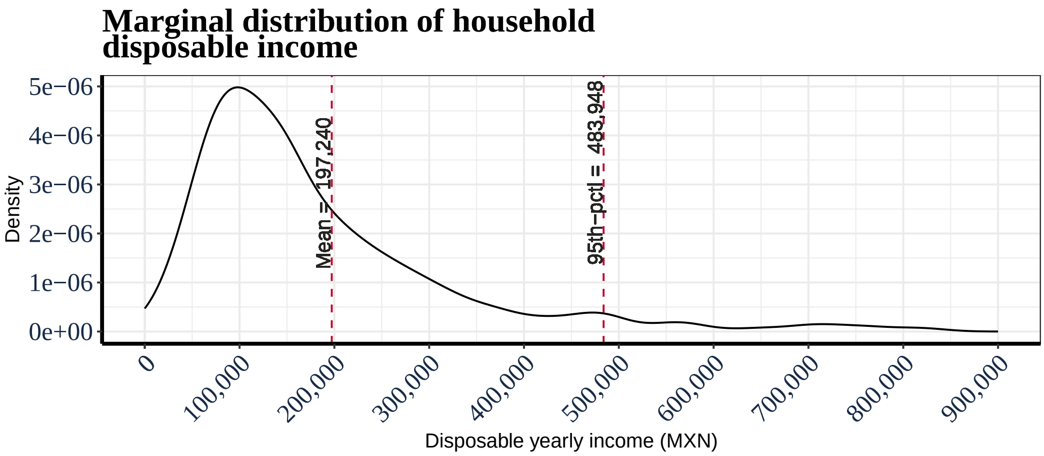
6 Post estimation calculations

Given the estimates of RC Model 2, I can do some post-estimation calculations to learn about the empirical distribution of taste parameters and consumers' willingness to drive. The model estimates a market-level price elasticity of -0.42 and -0.64 for Mexicali and Tijuana, respectively. The estimates vary due to the level of income and number of children in the household, see figure 9. This feature of the model allows for localized estimates and to estimate heterogeneous impacts of price liberalization at the market level.

In Mexico, the average commuting distance is 5 km on a one-way trip. The median consumer would be indifferent to drive an additional kilometer if it meant refueling at a price that is $13^{MXN¢}/_{L}$ ($3.1^{US¢}/_{gal}$) lower. I compute willingness to drive as $\Delta P_i = \frac{-(\lambda_1 + 2\lambda_2 d_{ijk})}{\alpha_i}$. Figure 10 shows the marginal distribution of consumer willingness to drive. The horizontal axis shows the savings that would be needed for a consumer to be indifferent to drive an additional kilometer. The median consumer needs savings of $13^{MXN¢}/_{L}$ ($3.1^{US¢}/_{gal}$), whereas the least sensitive consumers need savings of $21^{MXN¢}/_{L}$ ($5^{US¢}/_{gal}$).

My results show that price sensitivity depends on the level of disposable income. There is a positive correlation with the level of income. That is, higher-income households tend to be less price sensitive than lower-income ones. Therefore, high-income households tend to need higher discounts at the pump to be incentivized to travel an additional kilometer.

While the level of income and price sensitivity are negatively correlated, my results do not show a homothetic relationship. For example, table 5 shows that the median household in income decile 3 has a lower price sensitivity than the median household in decile 4. This estimation result matches the distribution of household consumption of fuel as depicted in table 3. Despite lower income levels, households in income decile 3 consume more liters per year than their counterparts in decile 4.



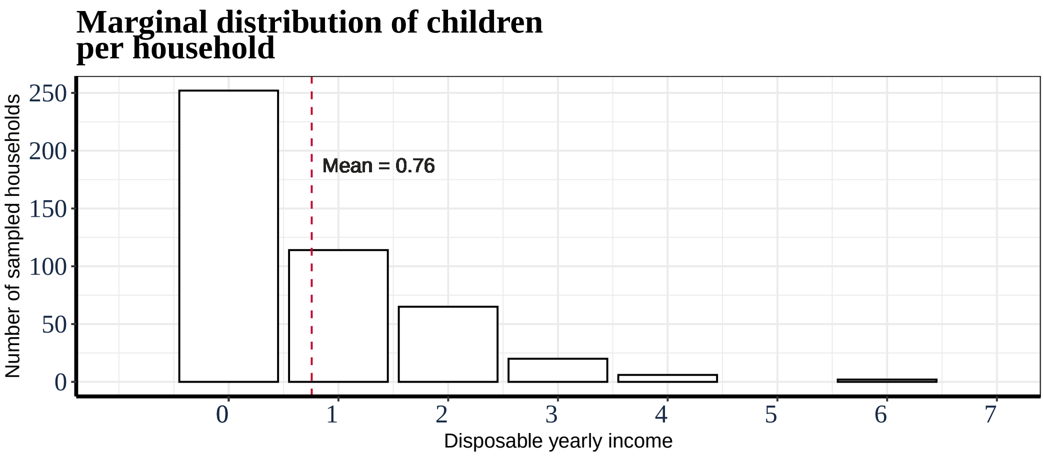


Figure 9: Income distribution is uneven in Mexico: For 2014 the average household in Tijuana had a yearly disposable income of 197,204 MXN while the 95^th percentile earns 483,948 MXN. On average each household has 0.76 children living in them.

Similarly to Seim and Waldfogel 2013, the lack of data on individual purchase decisions prevents me from separating the decision to visit a gas station and the purchase decision. Therefore, I make some assumptions. It would take, on average, two minutes to drive one kilometer in downtown Tijuana. Taking this into consideration and assuming that in each visit a consumer will fill up half their tank of gas (25 L approx.), then we can compute how much consumers in different households value an hour of driving. Unsurprisingly, as income grows the implied value of an hour driven increases as well. However, when compared to high-income households, low-income households value an hour of driving proportionately

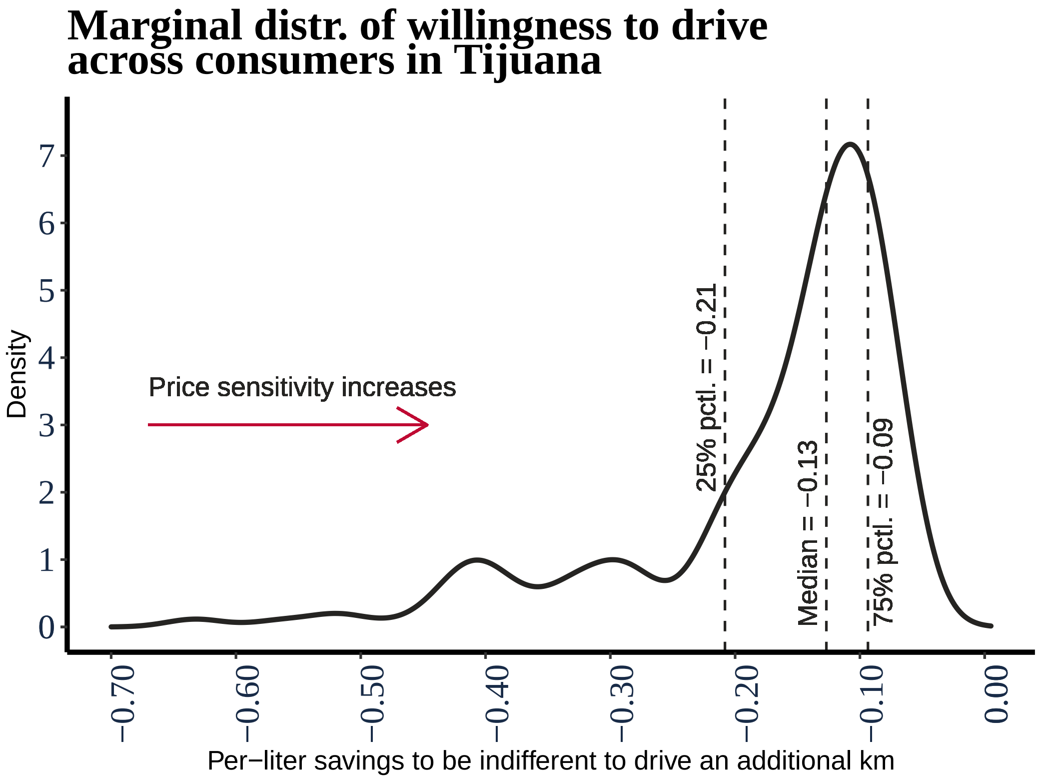
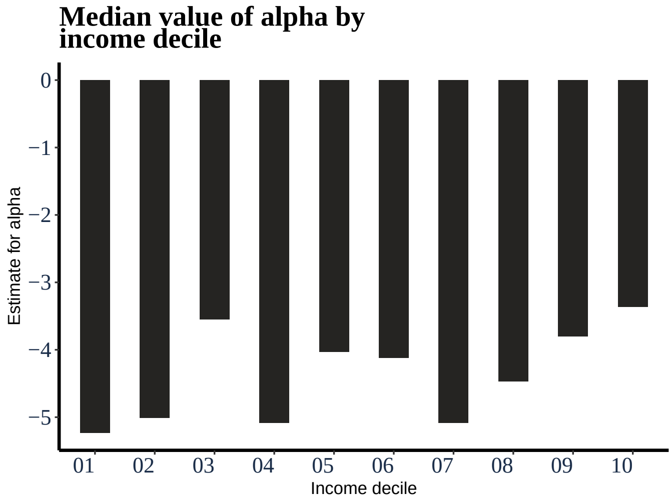


Figure 10: Given the heterogeneity in price sensitivity, the willingness to drive varies across consumers. More price-sensitive consumers will require fewer savings at the pump to be willing to drive an additional kilometer. The median consumer requires to save 13¢/L to be willing to drive one kilometer, while the top 25% more price-sensitive consumers would require savings of 9¢/L or less.

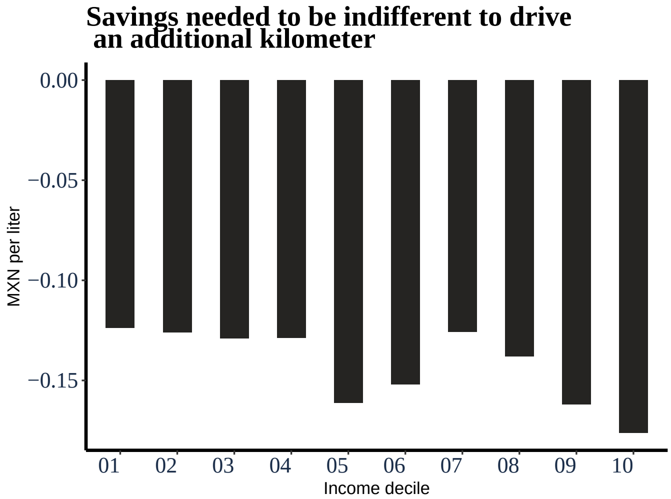
much more than their median hourly earnings. The top decile of households, who have incomes comparable to the median U.S. household, value their driving time as 65% of their hourly earnings which is consistent with Dorsey et al. 2022 who estimate that customers from the U.S. value their driving time at 89% of their wage. See table 5 for further details.

6.1 Predicting market shares

To estimate demand for retail gasoline, I use a highly detailed data set of the universe of gas stations in Mexicali and Tijuana in 2015 which includes their location, sales volume, attributes, and average price charged throughout that year. For the following years, I can observe the location, attributes, and pricing of the universe of gas stations, incumbents, and



(a) Income and price sensitivity

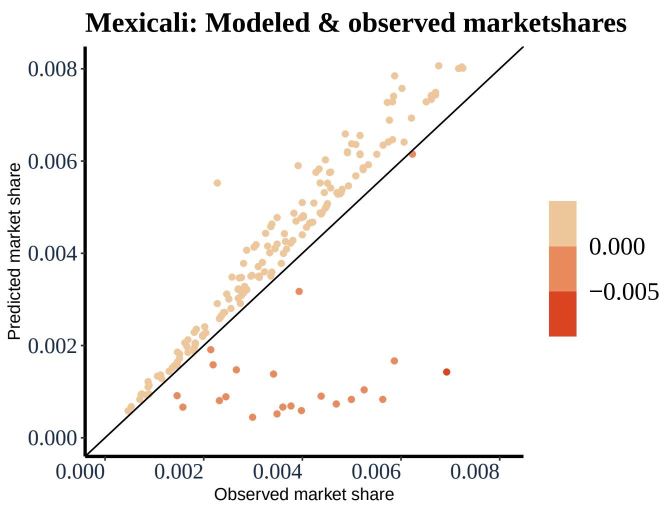


(b) Income and savings to drive. Savings needed are computed as $\Delta P = \frac{-(\lambda_1+2\lambda_2 d_{ijk})}{\alpha_i}$

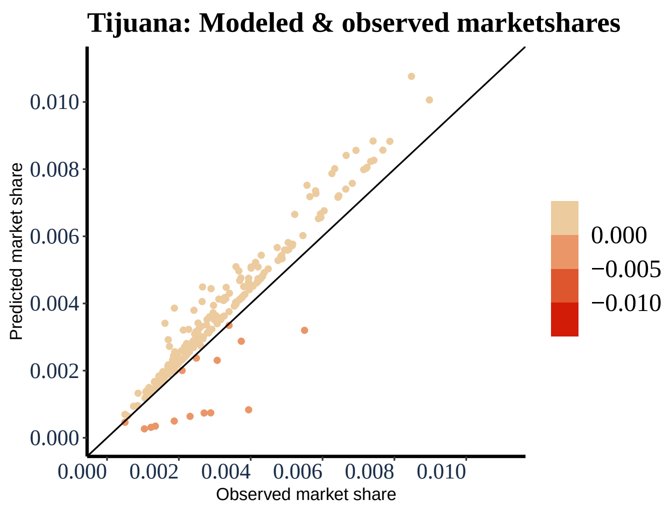
Figure 11: Gas station entry in Mexicali and Tijuana

Income decile	**Earnings** (MXN/hr.)	**Parameter:** α	**Savings to drive** (MXN/km)	**Implied value of driving** (MXN/hr.)
1	21.3	-5.23	-0.124	92.9
2	32.3	-5.01	-0.126	94.4
3	39.3	-3.55	-0.129	96.8
4	46.7	-5.09	-0.129	96.6
5	55.9	-4.03	-0.161	121
6	64.6	-4.12	-0.152	114
7	77.5	-5.09	-0.126	94.3
8	96.1	-4.47	-0.138	104
9	128	-3.8	-0.162	121
10	204	-3.36	-0.176	132
Corr. with hourly earnings		0.615	-0.786	0.784

Table 5: Summary of estimation results by income decile



(a)



(b)

Figure 12: The model tends to slightly overestimate some gas station sales.

entrants alike. However, I cannot observe the level of sales.

As discussed in section 10.4, a random coefficient model would be the best choice to predict changes in the choice set as it doesn't impose a priori substitution patterns. To test how well the model is performing, I follow a three-step procedure: first, I estimate the model's parameters. The estimates are shown in section 5. Second, I compute the values of ξ for all gas stations in Mexicali and Tijuana (see figure 13). Third, I simulate the consumer choices by taking the parameter estimates and ξ as given and then I input the different product and demographic characteristics into the model. The prediction errors are displayed in Figures 12(a) and 12(b).

Overall, the model seems to be performing relatively well at predicting market shares given the observable attributes of the gas stations.

7 Consumer welfare impacts

Price liberalization brought substantial gas station openings, initially, yet it also led to an increase in prices. However, the entirety of the price change cannot be attributed to MER

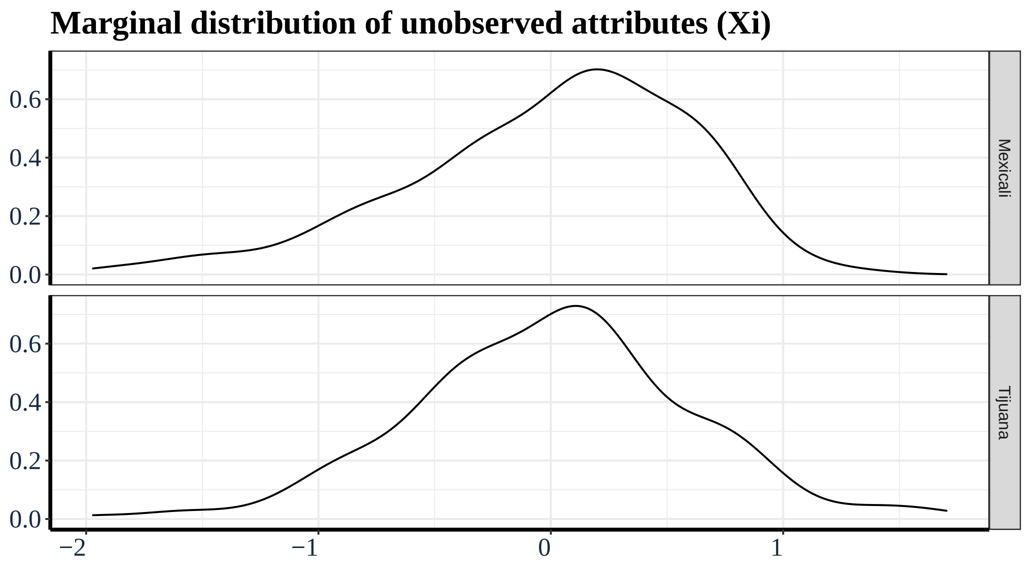


Figure 13: Distribution of unobserved attributes ξ_k

since international wholesale gasoline prices increased during the 2016-2019 period and the MoF would have adjusted taxes accordingly. Therefore, the main challenge is to define the correct counterfactual scenario to compare the observed outcomes to. Fortunately, the MoF followed a pricing formula for the retail prices as described in equation (9).

While the ultimate objective is to compute the welfare change from enacting MER, I start by computing, separately, the welfare change that households experienced from 2015 to 2019 from increased product availability and from price changes. In each setting, given the prices and product availability, I compute the taxes collected from the sales of fuel. Given that there is no explicit information about the use of these taxes, I assume that what is collected is redistributed evenly across households. This is an arbitrary, yet necessary assumption. While ENIGH provides data on the amount of government transfers received, the vast majority of government expenditure goes to other services like education and healthcare (Castelletti 2013). Estimating how expenditure in these services is consumed by households, conditional on their income level, requires additional assumptions. Future work will test different distribution assumptions. In the same vein as Petrin 2002, I use compensating variation to compute the consumer welfare gains.

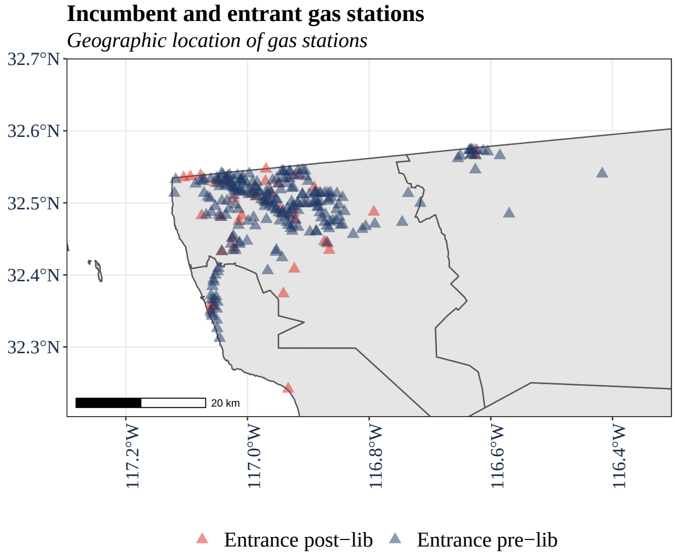


Figure 14: Location of gas stations in Tijuana

7.1 Scenario 1: Consumer welfare change from additional product availability

While it is generally agreed upon that consumers benefit from additional products, Dixit and Stiglitz 1977 and Mankiw and Whinston 1986 show that, in theory, consumers might not benefit much if the new products are not sufficiently differentiated. I measure how much consumers benefited from additional gas stations.

Few empirical papers have tested if consumers actually benefit from the availability of more products. Some notable exceptions include Berry and Waldfogel 1999, Petrin 2002, Seim and Waldfogel 2013, and Berry et al. 2016 who examined radio stations, minivans, and liquor stores. No one has studied the benefits of increased availability in retail gasoline markets.

To isolate the benefit of additional gas stations, I keep prices constant at the 2015 levels and simulate consumer choices in a counterfactual market with the gas stations that opened between 2016 and 2019. I refer to these stations as entrants. I can observe the attributes and location of entrant gas stations. Depending on their location relative to the U.S. border, I assign the price they would have been charging given the administered price regime in 2015.

In this setting, the first challenge comes from data limitations. Since I cannot observe gas-station-level sales post-liberalization, I cannot estimate ξ for entrants. However, I have the non-parametric distribution of ξ for incumbent gas stations. See figure 13 for the distribution in Mexicali and Tijuana. I take random draws with no replacement from the corresponding non-parametric distribution and assign them to each incumbent gas station depending on its market location.

I then compute the individual choice probabilities of each gas station for all consumers in each market, given the new choice set. Figure 14 shows the location of entrant and incumbent gas stations in the metropolitan area of Tijuana. Table 6 summarises the number and types of gas stations in the new choice set.

Market	**Station type**	**Number**
Metro Mexicali	Incumbent	214
	Entrant	10
Metro Tijuana	Incumbent	235
	Entrant	32

Table 6: Gas stations by market

Higher product availability increases overall consumption as transportation costs are reduced for some households. The increase in consumption translates into an increase in the taxable base of roughly 8.6% which increases the transfers to all households. High-income households, who are less price sensitive and consequently less willing to drive to refuel, benefit from the increased convenience. While the consumption of gasoline by low-income households is relatively small, they benefit from increased transfers.

The gain from increased product availability results in an aggregate increase in consumer welfare of 1,459,453,575 MXN/year. The gasoline industry in Mexicali and Tijuana has a yearly revenue of 20,225,820,000 MXN (approx. 1.27 billion USD). Then, the entry of 42 gas stations increased household welfare by roughly 7.2% of the annual revenue.

7.2 Consumer welfare change from price changes

Between 2016, when price liberalization started, and 2019 we observe an increase in retail prices. However, the entirety of the retail price change cannot be attributed to the price

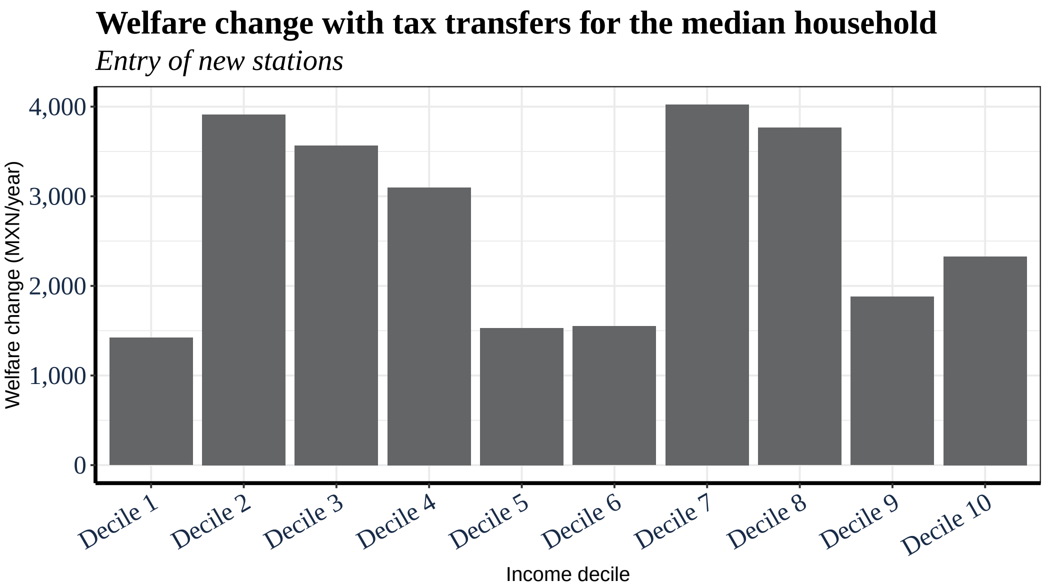


Figure 15: The median household in the first decile benefits in 1,422$^{MXN}/_{year}$, mostly derived from increased transfers from taxes. The median household in the tenth decile benefits in 2,328$^{MXN}/_{year}$, mostly from the convenience of increased station availability. Overall, households are better off by 1,459,453,575 $^{MXN}/_{year}$ (91,963,048 $^{\text{USD}}/_{\text{year}}$) due to additional product availability.

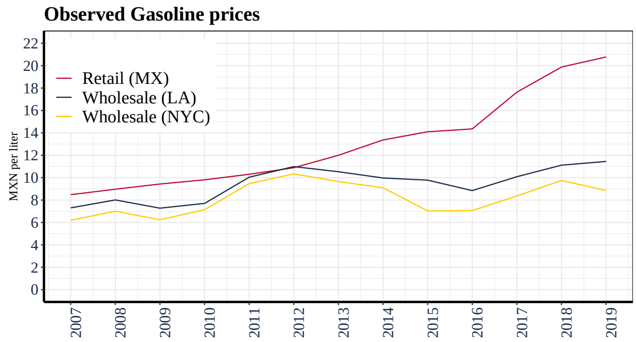
liberalization since international wholesale prices also increased during this time period and Mexico imports the vast majority of the gasoline it consumes. Then, there is the challenge of determining the right benchmark that would have been observed in 2019 had the policy not been removed.

In figure 16(a) we can observe that wholesale prices in New York City and Los Angeles follow a similar trajectory with an increase of 1.82$^{MXN}/_{L}$ and 1.67$^{MXN}/_{L}$ throughout the 2016-2019 period, respectively.[6] Mexican retail prices are also increasing during this period but disproportionately more at 6.67$^{MXN}/_{L}$ on average. Figure 16(b) shows a cross-section of average observed retail prices at every gas station in Tijuana throughout 2019.

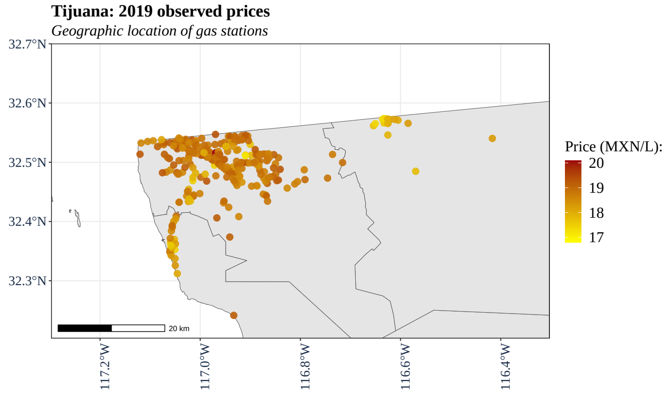
7.2.1 Retail pricing rule

Prior to MER, the MoF would follow a pricing formula as depicted in equation (9) (Instituto Belisario Domínguez. Senado de la República 2015 and Secretaría de Energía 2017). The

[6]New York City prices are taken from the prices of futures contracts with delivery in New York Harbor. Los Angeles prices are taken from CARBOB spot prices measured in Los Angeles.



(a) The observed increase in retail prices is disproportionately higher than wholesale price increases in some of the main benchmark prices



(b) Cross-section of retail prices in Tijuana

Figure 16: Gas station entry in Mexicali and Tijuana

variable $P_{ret,t}$ is the objective retail price for time period t; $P_{whol,t}$ is a reference price for wholesale gasoline; $TranspC_t$ are estimated transport costs from the rack to the gas station; $RetMargin_t$ is the "profit margin" left to the retailer; $ProdLoss_t$ are the estimated losses from product transportation and handling; Tax_t^{fossil} are taxes on fossil fuels; $Tax_t^{IEPS-Fed}$ is a federal variable tax. IEPS stands for *Impuesto Especial sobre Producción y Servicios* and it is a family of taxes applied to gasoline, tobacco, alcohol, and sodas. VAT_t is a value-added tax of 16%, and $Tax_t^{IEPS-State}$ is a state-level tax similar to the federal tax, $Tax_t^{IEPS-Fed}$.

$$\begin{aligned} P_{ret,t} =& (P_{whol,t} + TranspC_t + RetMargin_t + ProdLoss_t + \\ & Tax_t^{fossil} + Tax_t^{IEPS-Fed}) \times VAT_t + Tax_t^{IEPS-State} \end{aligned} \tag{9}$$

Through this formula, the MoF would target a retail price, $P_{ret,t}$, for a period t and would adjust $Tax_t^{IEPS-Fed}$ to partially compensate for changes in $P_{whol,t}$. In Figure 17 we can observe the different prices and components in the years from 2006 to 2014 $Tax_t^{IEPS-Fed}$ even become a subsidy in periods of increased wholesale prices.

In the following, section I will compute welfare changes based on different benchmark scenarios of retail pricing policies. I'll begin with the observed price change between 2015 and 2019 to gauge the size of the impact prices have on consumer welfare and then compare it to the welfare gains from gas station entry. Then, I'll simulate possible pricing policies that the MoF could have followed had it kept its price-control policy and adjusted Federal IEPS taxes to the changes observed on international wholesale prices.

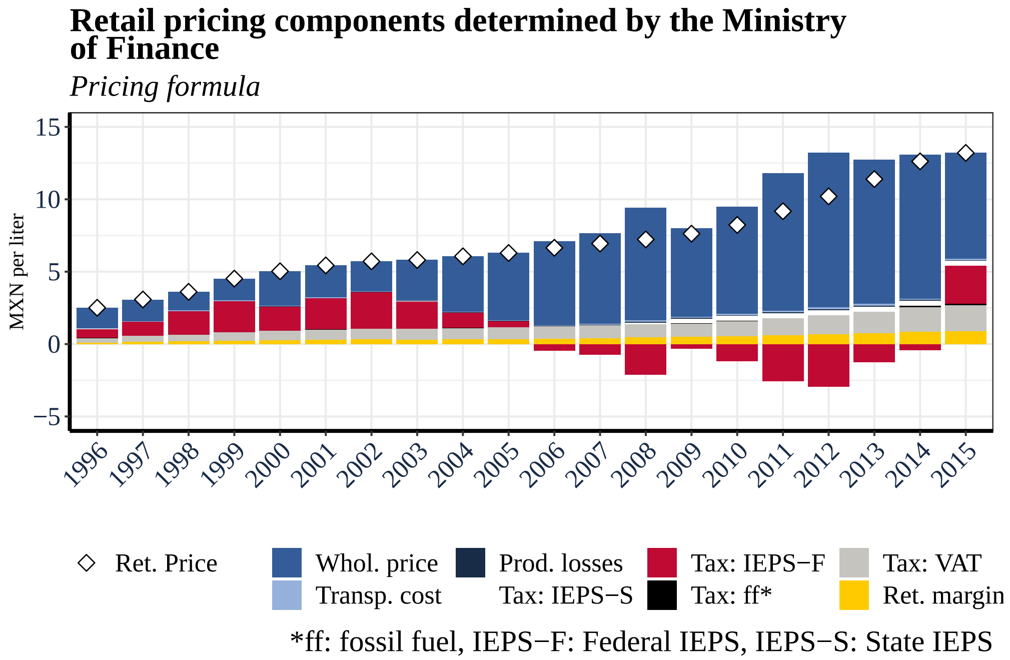


Figure 17: The MoF followed a formula for retail prices. The Federal IEPS tax served as the adjusting factor in the pricing formula. For example, between 2006 and 2014 the Federal IEPS transitioned from a tax to a subsidy as international wholesale prices increased substantially with respect to the target retail price.

7.2.2 Scenario (2): Passive Ministry of Finance

The purpose of this scenario is to gauge the magnitude that price increases have on consumer welfare. This counterfactual scenario assumes that the price control policy had been kept in place and that the MoF let a complete pass-through of the change in international wholesale prices onto retail prices. The change for wholesale L.A. gasoline, the relevant benchmark for Mexicali and Tijuana, was of 1.67 MXN/L from 2015 to 2019.

To isolate the effect of price on welfare, I keep the choice set as the one consumers faced in 2015, that is, I assume there were no entrants. I also keep the number of households constant to isolate the effect of price on the quantity consumed. I call this scenario the

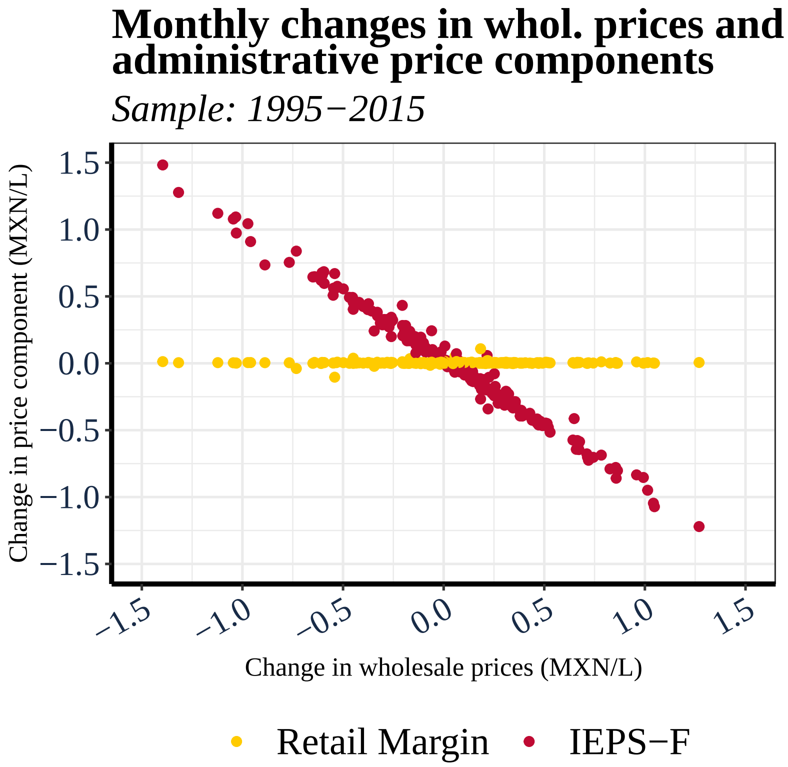


Figure 18: The MoF would adjust the Federal IEPS tax to counter changes in international wholesale prices to stabilize prices. Other formula components, like retail margins, saw little to no change in responding to wholesale prices.

"Passive MoF" as the MoF would not have lowered taxes in response to a price increase.

The increase in retail prices translates into a welfare loss of 1,354,021,316 MXN/year (6.7% of the industry's annual revenue). The brunt of the loss is mostly evenly distributed. High-income households are affected by higher prices at the pump, which in turn reduces the taxable base by 11.15% which results in lower transfers to lower-income households. See figure 19 for the impact distribution across households.

7.2.3 Scenario (3): Proactive Ministry of Finance

Figure 17 shows how the MoF adjusted the level of the Federal IEPS in response to international wholesale prices before retail price liberalization. The MoF would reduce (increase) the Federal IEPS tax in response to increases (reductions) in wholesale price, but it wasn't

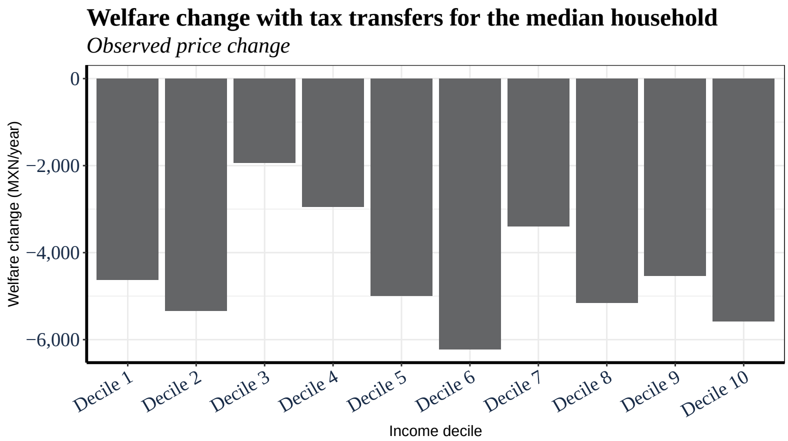


Figure 19: All consumers are negatively affected by the increase in prices. The median household in the first decile is worse off by 4,623 MXN/year, while the median household in the tenth decile is worse off by 5,576MXN/year.

Variance-Covariance Matrix from 1995-2016 (Changes)			
	Wholesale price	**Federal Tx (IEPS)**	**Ret. Margin**
Wholesale price	0.2110	-0.2062	0.00032
Federal Tx (IEPS)	-0.2062	0.2040	-0.0003
Ret. Margin	0.0003	-0.0003	0.0001

Table 7: Variance-covariance matrix of contemporaneous changes in the retail pricing formula

always a one-to-one change. The relevant counterfactual scenario involves a continuation of this policy, however, it is uncertain how aggressively the MoF would have adjusted taxes in the presence of the observed changes in wholesale prices throughout 2016 and 2019.

I perform simulations of possible pricing paths the MoF could have followed using Montecarlo simulations. These counterfactual pricing paths will serve as the prices in the benchmark scenarios. I use the variance-covariance matrix in table 7 to make random draws out of a multivariate normal.

The results of the simulation exercise are presented in figure 20. It is interesting to note that as time has passed from when prices were liberalized, the observed average yearly price lies further away in the distribution of simulated prices. For example: had the administered price regime been kept, the median simulated price would have been 13.78$^{MXN}/_{L}$ for 2017, while the average observed price is 17.63$^{MXN}/_{L}$. This is beyond 17$^{MXN}/_{L}$, the average simulated price that is two standard deviations away from the mean.

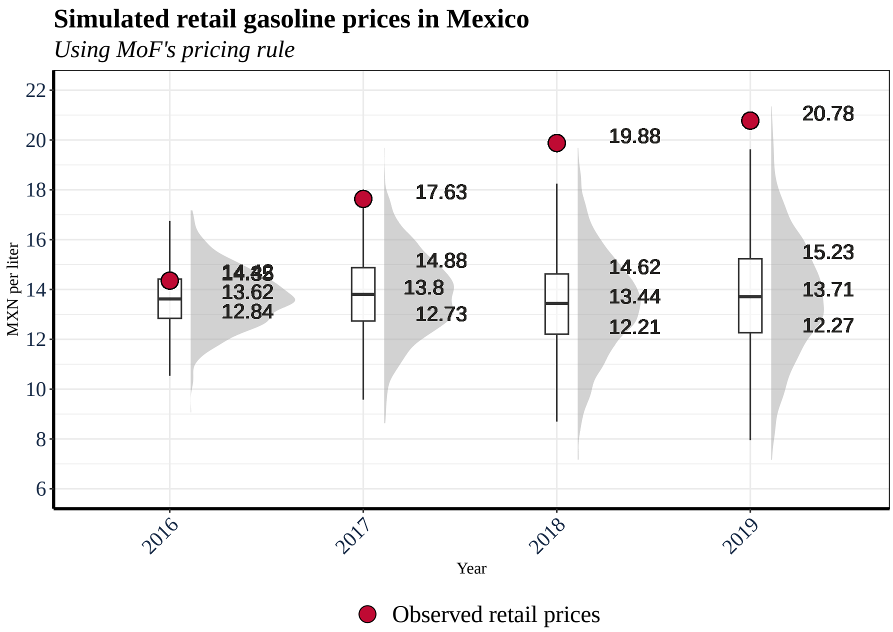
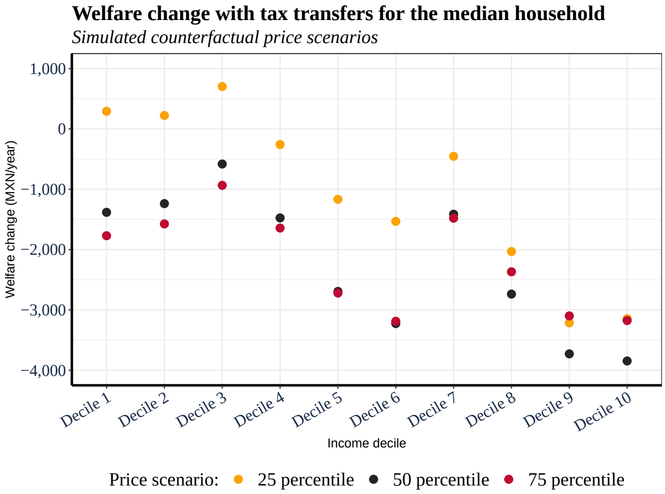


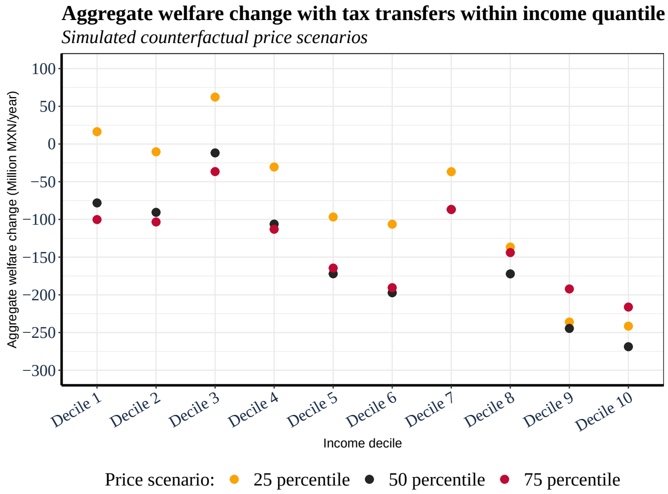
Figure 20: Possible pricing paths in counterfactual pricing policy. Each year shows the marginal distribution of retail prices given the observed wholesale price changes. Next to each distribution is a traditional box and whiskers plot showing the values of the 25^{th}, 50^{th}, and 75^{th} percentiles in black numbers. In red, I plot the average observed price during each year.

Had the price-control policy remained in place the MoF could have followed different tax policies. For ease of exposition, I will focus on three scenarios. These three outcomes are characterized by retail prices in the 25^{th}, 50^{th}, and 75^{th} percentile of the distribution of simulated prices. Retail prices in the 25^{th} percentile result in a "low tax" scenario, while retail prices in the 75^{th} percentile result in a "high tax" scenario. I refer to the scenario with prices in the 50^{th} percentile as the "base case" scenario.

Based on these three main outcomes, the policy implementation had an adverse effect on consumer welfare. However, the magnitude of the impact varies. Table 8 summarises the welfare outcomes that use these three counterfactual scenarios as a starting point to compare to the 2019 outcome. In the 2019 outcome, retail prices are the highest at 20.78MXN/L. This is due, in some part, to an average per-liter tax of 4.6MXN/L, however, most of the difference with respect to wholesale prices cannot be explained by high taxes alone. High retail prices result in 1,001.35 million liters sold per year and a tax collection of 4,646 million MXN per



(a)



(b)

Figure 21: Welfare change under different counterfactual simulated scenarios for the median household (left) and for the aggregate income decile (right).

year.

Figure 21(b) shows the welfare change for all the households within one decile. There are two things that stand out: in all the scenarios there is an aggregate welfare loss for most of the income deciles. Second, generally speaking, the higher the income the higher the loss. As households become more wealthy they consume more liters of gasoline and the brunt of higher prices overshadows the benefits of transfers.

In a "low tax" counterfactual scenario, retail prices would have been relatively low at 12.27MXN/L and, despite having the largest amount of liters sold, tax collection would have been relatively low at 4,327.48 million MXN/year given a low average per-liter tax of 2.8MXN/L. The retail price change between this scenario and the 2019 observed outcome is the largest. The brunt of this effect is carried by high-income households (Decile 9 and 10) who consume the most gasoline. Low-income households are affected much less as they consume much less gasoline and benefit from marginally higher tax transfers. See the yellow dots in figures 21(a) and 21(b).

In a "high tax" scenario the converse is true. Counterfactual retail prices would have been high, to begin with at 15.23MXN/L, so high-income households would have been affected

Outcomes of simulated tax policy given the retail pricing rule						
Scenarios	**Percentile**	**Prices**	**Liters sold** (mill./year)	**Taxes** (mill. MXN/year)	**Aggr. welfare change** (mill. MXN/year)	**Share of revenue** **(annual)**
Counterfactual	25th	12.27	1,501.22	4,327.48	-816.36	4.0%
	50th	13.71	1,363.48	5,062.45	-1,428.47	7.1%
	75th	15.23	1,286.76	5,274.66	-1,347.43	6.7%
2019		20.78	1,001.35	4,646.12	-	-

Table 8: Following the MoF's pricing rule and its historical reaction to wholesale price shocks, I simulate different pricing paths and their implications for consumption, taxation, and welfare. I report outcomes at the 25^{th}, 50^{th}, and 75^{th} of the pricing distribution.

much less due to a lower price increase. In this case, however, low-income households bear the brunt of the policy as the change in collected taxes would have been the highest going from 5,274 million $^{\text{MXN}}/_{\text{year}}$ to 4,646.12 million $^{\text{MXN}}/_{\text{year}}$.

The "base case" scenario has the highest level of consumer welfare among the three analyzed scenarios. Therefore the change going from this scenario to the 2019 outcome is the largest. See the blue dots in figures 21(a) and 21(b). In the "base case" scenario there is somewhat of a balance between the amounts of taxes redistributed and the prices paid at the pump, see table 9 for further details. For example, for deciles 8, 9, and 10, the median household experiences lower levels of utility from consumption. At the same time, these households receive a higher level of utility from the tax transfers in Pricing Scenario 50% than in Pricing Scenario 25%.

Starting from the "base case" scenario every income decile is worse off in aggregate. Going from this scenario to the 2019 outcome results in going to a high-price setting with relatively low taxes. The top three deciles are the most affected in this scenario since they overwhelmingly consume more gasoline

In the last column of table 8 we observe that, on aggregate, the price liberalization policy has led consumers to an outcome in which they are worse off. The main culprit is that retail prices have increased much more than the increase in taxes and wholesale prices combined. This has had two adverse effects on welfare: first, higher-income households, who are the customers who derive the most utility out of the consumption of gasoline in this market, end up paying higher prices. Second, due to higher prices, aggregate consumption falls and overall tax collection decreases. Even though I assumed that tax collection is distributed evenly, it is lower-income households that benefit the most since they tend to be more price sensitive and more sensitive to income levels due to a higher magnitude of α_i, see table 5

Utility level for the median household within each decile according to source												
Income decile	**Pricing Scenario 25%**			**Pricing Scenario 50%**			**Pricing Scenario 75%**			**2019 Outcome**		
	Consumption	*Transfers*	***Total***	*Consumption*	*Transfers*	***Total***	*Consumption*	*Transfers*	***Total***	*Consumption*	*Transfers*	***Total***
1	3,583	50,248	**53,831**	3,174	58,781	**61,955**	2,782	61,245	**64,027**	2,185	57,048	**59,233**
2	16,351	49,943	**66,294**	15,627	58,425	**74,052**	15,082	60,874	**75,956**	14,403	51,958	**66,361**
3	15,121	35,606	**50,727**	14,771	41,653	**56,424**	14,359	43,399	**57,758**	14,343	45,746	**60,089**
4	17,292	39,194	**56,486**	16,526	45,850	**62,376**	15,725	47,772	**63,497**	14,623	41,681	**56,304**
5	21,544	56,511	**78,055**	19,986	66,109	**86,095**	18,398	68,880	**87,278**	16,093	59,119	**75,212**
6	18,763	49,737	**68,500**	16,480	58,184	**74,664**	14,212	60,623	**74,835**	11,244	46,524	**57,768**
7	22,435	40,870	**63,305**	20,656	47,812	**68,468**	19,052	49,816	**68,868**	16,835	34,803	**51,638**
8	20,933	54,756	**75,689**	17,188	64,056	**81,244**	14,030	66,741	**80,771**	11,047	46,470	**57,517**
9	45,408	43,426	**88,834**	42,303	50,801	**93,104**	39,527	52,931	**92,458**	36,179	40,208	**76,387**
10	42,503	46,117	**88,620**	38,190	53,950	**92,140**	34,102	56,211	**90,313**	30,488	37,563	**68,051**

Table 9: Increasing taxes can be beneficial for some income deciles as the increase in transfers more than compensates the loss for paying more at the pump. For example, increasing taxes and going from Pricing scenario 25% to Pricing scenario 50% increases the level of utility for the median household in Decile 1. However, increasing taxes too much will eventually affect households. See the median household in Decile 10 going from Scenario 50% to 75%.

My preferred choice for a counterfactual scenario is Pricing Scenario 50%. In this scenario taxes collected are higher than in the 2019 outcome, therefore consumers still experience a decrease in transfers jointly with an increase in prices. In this scenario, consumers are worse off by 1,428.47 million MXN per year in aggregate.

Almost every decile is worse off, yet the ones more affected by the policy, in absolute levels, are households in the eighth, ninth, and tenth income deciles who overwhelmingly consume more gasoline. Households in deciles 1 and 2 are affected by a decrease in transfers.

However, welfare loss, as a percentage of income, is the highest for low and middle-income households. Figure 22 shows that the welfare loss is 3.07% and 2.2% for the median household in deciles 1 and 2, respectively. Households in deciles 5 and 6 lose 2.8% and 2.88% of their annual income. Finally, the median household in the 10th decile is worse off by 1.5% of their annual income. These results suggest this policy has had regressive effects on welfare.

8 Conclusion

After more than eighty years of price control policies, the Federal Government allowed gas stations to choose their own pricing strategy. However, due to regulatory backlogs,

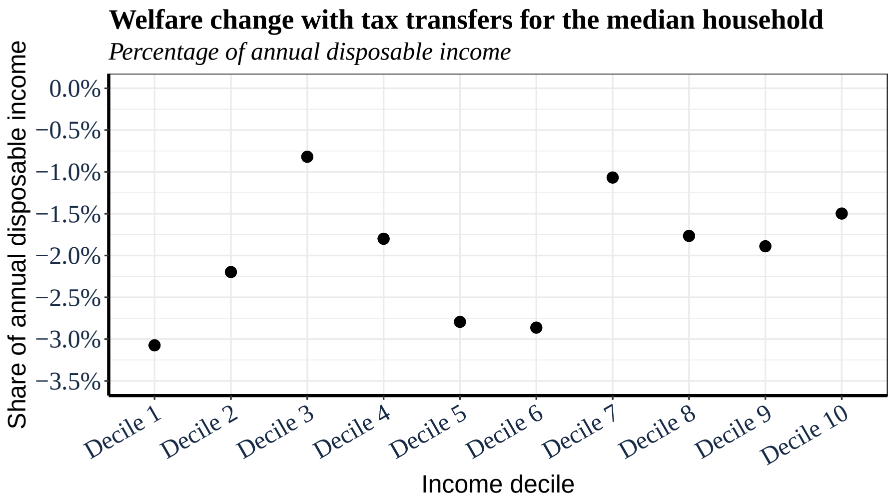


Figure 22: As a proportion of their income, households in the bottom deciles and middle-class households are affected the most.

the number of gas stations that entered the market was constrained. This led gas station operators to increase their prices more than the increase in wholesale prices during that time period.

This is the first paper to estimate the retail elasticity of demand for gasoline as a spatially differentiated product. I find that market-level price elasticity is -0.42 and -0.64 for Mexicali and Tijuana, respectively. Additionally, I find that a consumer's price sensitivity depends on their level of income and that low-income households are up to 55% more price sensitive than high-income households. I also find that consumers dislike driving and that the median household would need 13 MXN¢/L in savings at the pump to be indifferent to driving an additional kilometer.

To calculate the welfare change, I define a counterfactual scenario of how prices would have looked like had the price control policy remained in place. I first leverage the MoF's formulaic pricing rule to estimate the variance-covariance matrix between taxes and wholesale prices. Then, given the observed wholesale monthly price changes, I simulated 1,000 possible price paths across four years, yielding a total of 48,000 simulated prices. For 2017 onward, the observed retail price is higher than the 99th percentile of simulated prices.

My preferred counterfactual is the scenario of taxes that yields retail prices in the 50th percentile of the distribution of simulated prices. I compare the welfare change from this scenario to the observed 2019 outcome. I find that, in the aggregate, households are worse off from this policy by 1,428 million $^{\text{MXN}}/_{\text{year}}$ (7.1%of the markets' annual revenue).

In absolute levels, households in the eighth and ninth income deciles are affected the most as they overwhelmingly consume more gasoline than lower-income households. However, lower-income households are adversely affected as well. The increase in retail prices leads to a decrease in liters sold and a reduction in taxes collected. This, in turn, turns into lower government spending which households benefit from.

Overall, this policy has had a regressive impact on household welfare. The median household in the bottom decile is worse off by 3.07% of their annual income, while the median household in the top income decile is worse off by 1.5%.

The observed increase in retail prices cannot be explained only by the increase in wholesale prices and the change in taxes. It is plausible that regulatory backlogs resulting in a lower number of gas station openings and the ability of gas stations to choose their own pricing have given station operators the ability to exercise local market power.

9 Glossary

Acronyms

CONAPO Consejo Nacional de Población. 14

CRE Comisión Reguladora de Energía. 9, 12, 13

INEGI Instituto Nacional de Estadística y Geografía. 9–11, 14

MER Mexico's Energy Reform. 6, 12, 13, 24, 30, 31, 34

MoF Ministry of Finance. 7, 8, 31, 34–39, 41, 43

PEMEX Petróleos Mexicanos. 4, 6, 7, 10, 11

SHCP Secreatría de Hacienda y Crédito Público. 14

10 Appendix

10.1 Tables and Charts

In table 10 we can observe the results of a survey done of customers who were refueling at gas stations in different locations in California. Surveyed customers were asked where they were coming from and where they were headed.

Activities before and after refueling						
	Activity at destination					
Activity at origin	Home	Work	Business	Shopping	Social	**Total**
Home	7.0%	8.1%	3.1%	12.5%	12.0%	**42.7%**
Work	10.5%	2.6%	1.6%	2.5%	0.9%	**18.1%**
Business	1.6%	0.9%	1.1%	0.5%	0.3%	**4.4%**
Shopping	12.1%	0.9%	0.7%	5.6%	2.1%	**21.4%**
Social	7.9%	0.6%	0.1%	1.8%	3.2%	**13.6%**
Total	**39.1%**	**13.1%**	**6.6%**	**22.9%**	**18.5%**	**100.0%**

Totals may not sum to 100% due to rounding, Table 2 from Kitamura and Sperling (1987a).

Table 10: The vast majority of customers have "home" either as an origin or destination. However, a sizeable percentage refuel on their way to work or as part of a shopping trip.

10.2 Estimation procedure

Modeling demand for gasoline as a differentiated product that is part of a demand system poses several estimation challenges as described in Section 3.1. Let $\hat{\theta} \equiv \{\bar{\alpha}, \bar{\gamma}, \beta, \lambda_1, \lambda_2, \Pi, \Sigma\}$, from equation (6) it is evident that demand for gas station j depends on the characteristics of all other gas stations $(\delta_1(x_1, \xi_1; \hat{\theta}) \ldots \delta_J(x_{jk}, \xi_{jk}; \hat{\theta})$. Additionally, I am not imposing the assumption of a single representative agent. Instead, I am assuming heterogeneous agents whose mean level of utility is determined by the parameters $\bar{\alpha}, \bar{\gamma}, \beta, \lambda_1, \lambda_2$ interacted with the product's characteristics. Parameters $\bar{\alpha}, \bar{\gamma}, \lambda_1, \lambda_2$ are scalars while β is a vector of length 6. In this model, what makes agents heterogeneous are deviations from the mean level of utility, these taste parameters are present in the matrices Π, Σ where $dim(\Pi) = 2 \times 4$ and $dim(\Sigma) = 2 \times 2$. The matrix Σ is symmetric, therefore has only 3 unknown parameters. Overall, there are J_k equations per market and $J_k + 10 + 8 + 3 = J_k + 21$ unknowns.

Suppose there are $m = 1 \ldots \mathcal{M}_k$ instruments, Z_m ,for each of the j products in market t such that

$$E(Z_{mk}\xi_k) = 0 \quad m = 1 \ldots \mathcal{M}_k \;\; for \; k = \{Mexicali, Tijuana\} \tag{10}$$

The sample moment condition is

$$\begin{aligned} m_{m,J_k}(\bar{\alpha}, \bar{\gamma}, \beta, \lambda_1, \lambda_2) &\equiv \frac{1}{J_k} \sum_{j=1}^{J_k} \xi_{jk} Z_{mjk} \\ &= \frac{1}{J_k} \sum_{j=1}^{J_k} (\hat{\delta}_{jk} - \bar{\gamma} - x_{jk}\beta + \bar{\alpha} P_{jk} + \lambda_1 \bar{d_{jk}} + \lambda_2 \bar{d^2_{jk}}) Z_{mjk} \end{aligned} \tag{11}$$

To estimate the model, I follow the estimating procedure of Conlon and Gortmaker 2020. Broadly speaking, it is a three-step procedure:

- **Step 1**: Take some values of $\hat{\theta} = \hat{\theta}_0$ as given and solve for estimates $\tilde{\delta}_1 \ldots \tilde{\delta}_J$ using the system of J equations generated by equation (6).

$$
\begin{aligned}
\tilde{s_1} &= s(\tilde{\delta}_1 \ldots \tilde{\delta}_J; \bar{\alpha}, \bar{\gamma}, \beta, \Pi, \Sigma) \\
&\vdots \\
\tilde{s_J} &= s(\tilde{\delta}_1 \ldots \tilde{\delta}_J; \bar{\alpha}, \bar{\gamma}, \beta, \Pi, \Sigma)
\end{aligned} \tag{12}
$$

- **Step 2**: Given the estimates of the mean utilities from Step 1, use a set of instruments Z_{mk} to compute $Q(\hat{\theta})$:

$$
Q(\theta) \equiv \left[G_{J_k}(\bar{\alpha}, \bar{\beta}, \lambda_1, \lambda_2)\right]' W_{J_k} \left[G_{J_k}(\bar{\gamma}, \beta, \bar{\alpha}, \lambda_1, \lambda_2)\right] \tag{13}
$$

where

$$
G_{J_k}(\theta; \tilde{\delta}, x_k, P_k, \bar{d_k}, \bar{d_k^2}) = \begin{bmatrix} \frac{1}{J_k} \sum_{j=1}^{J_k} (\tilde{\delta}_{jk}(\hat{\theta_0}) - \bar{\gamma} - x_{jk}\beta + \bar{\alpha} P_{jk} + \lambda_1 \bar{d_{jk}} + \lambda_2 \bar{d_{jk}^2}) \times Z_{1jk} \\ \vdots \\ \frac{1}{J_k} \sum_{j=1}^{J_k} (\tilde{\delta}_{jk}(\hat{\theta_0}) - \bar{\gamma} - x_{jk}\beta + \bar{\alpha} P_{jk} + \lambda_1 \bar{d_{jk}} + \lambda_2 \bar{d_{jk}^2}) \times Z_{\mathcal{M}jk} \end{bmatrix} \tag{14}
$$

and where W_j is a weighting matrix with $dim(W_j) = \mathcal{M} \times \mathcal{M}$.

- **Step 3**: In the next estimation round, r, choose $\hat{\theta}(r)$ and repeat Step 1 and Step 2. Find $\hat{\theta(\cdot)}$ that minimizes equation (14).

More details on the estimation procedure are discussed at length in Conlon and Gortmaker 2020, Nevo 2000 and Berry et al. 1995.

A particular challenging point in the estimation procedure arises because each consumer faces a different distance to a gas station j. That is, even if two customers were to visit the same gas station and be exposed to the same prices or services, $d(L_i, L_j) \neq d(L_{i'}, L_j)$ $i \neq i'$, where $d(L_i, L_j) \equiv d_{ij}$ be the Euclidean distance between consumer i's home and gas station j. Following Davis 2006 As described in section 3, specifically in equation (5).

10.3 Initial values

As discussed in section 3, the model does not have a closed-form solution. To solve the model, an initial guess is required to start the algorithm that looks for the minimum of the objective function.

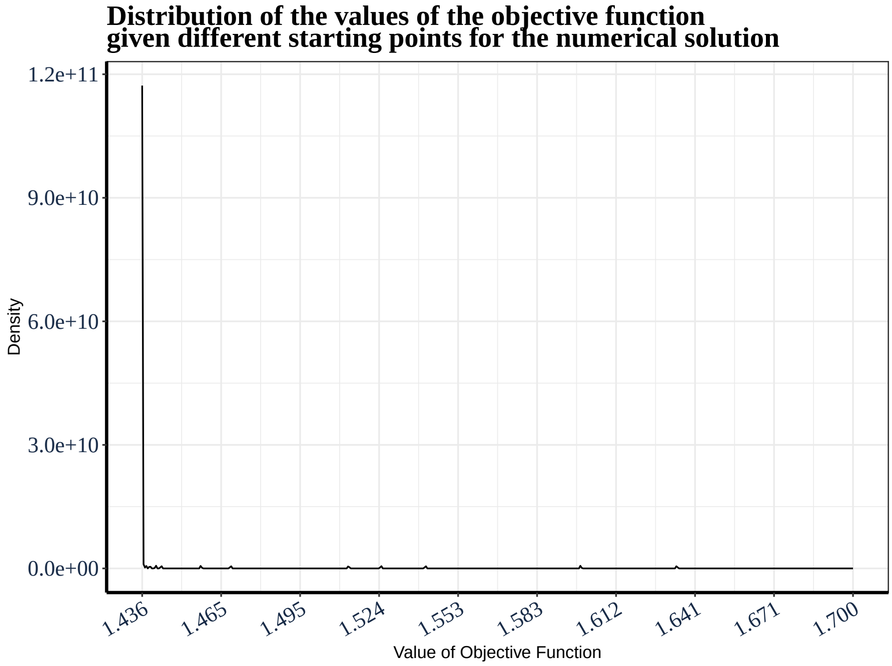


Figure 23: Out of 360 initial values this is the distribution of the values of the objective function once the numerical solution is found.

10.4 Model Alternatives

Despite being challenging to estimate, a random coefficients model is ideal for this setting. For example, there is no need to observe individual choices to estimate individual taste parameters. Similarly, by observing the distribution of demographic characteristics in the Mexicali and Tijuana markets, I can estimate taste parameters for heterogeneous agents.

This setting is also advantageous for the estimation of this model's parameters. Mexico's gasoline industry, as described in section 2, provides a unique setting to estimate gasoline as a geographically differentiated good since most traditional confounders like perceived product quality and branding are not simultaneously determined with price by the gas station operator. However, as Berry and Reiss 2007 point out, economists rarely have all the data to construct precise measurements of all factors affecting demand in a market, yet we rely on other data to draw inferences about demand.

In this setting, in particular, there is no gas-station-level sales data post-2015, yet there are gas-station-level prices. In addition, to estimate the change in welfare it is necessary to simulate a counterfactual environment had the policy not been implemented. The counterfactual environment follows a similar logic as the one presented in Petrin 2002. The lack of gas-station-level data post-2015 creates the necessity to use a model that does not impose *a priori* substitution patterns across goods and the random coefficients model satisfies that requirement.

There are other alternatives, though. The logit and multinomial logit models present the *independence of irrelevant alternatives* (IIA) issue that has been described by Ray 1973 and Berry 1994. The nested logit model ameliorates the IIA issue with the cross-substitution across goods, however, the substitution patterns are derived from a priori segmentation into discrete categories. For example, Verboven 1996 estimates the demand for cars by grouping them into six categories ranging from mini to luxury and sports cars; then subcategories are divided into domestic and foreign; finally, within each subcategory, the products are separated into their characteristics like horsepower and weight. However, the sequence through the researcher established the categories will have an effect on estimates since the IIA still exists within the lower nest (Nevo 2000).

In this setting's retail gasoline demand, there are no brands to discretize or segment the market. Furthermore, the distance a consumer would travel to the pump is a central feature of the behavior characterizing the consumption of this product, making a random coefficients model ideal for this setting.

10.5 Alternative model specifications: standard Logit.

In table 11, I report the estimates of three different model specifications. One of the advantages of the standard logit model is that it has a closed-form solution, but the implicit assumption is that consumers are homogeneous.

In the simplest model, Logit Model 1, consumers only have utility over prices and distance:

$$u_{jk} = \gamma + \alpha y - \alpha p_{jk} - \lambda_1 d_{jk} - \lambda_2 d_{jk}^2 + \varepsilon_{jk}$$

Columns 1 and 2 in table 11 show the model's estimates and its standard errors. The estimates show intuitive results, consumers dislike driving to a gas station and dislike paying for gasoline. However, there is a concern for omitted-variable bias since the model is not accounting for demand shifters as amenities offered on-site. Abnormally high estimates for the elasticity of demand at -4.66 for Tijuana and -4.15 for Mexicali could indicate this is the case.

	Logit Model 1		**Logit Model 2**		**Logit Model 3**	
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Parameter Name	*Estimate*	*s.e.*	*Estimate*	*s.e.*	*Estimate*	*s.e.*
Param. Estim.						
.. Intercept	17.93 ***	2.478	-0.24 ***	0.009		
.. Prices	-1.62 ***	0.181	-3.06 ***	0.116	0.23	0.164
.. Avg. dist	-0.54 ***	0.162	0.00	0.000	0.00	0.000
.. Avg dist sq.			0.00	0.099	0.00	0.011
.. Big business			0.04 ***	0.005	0.12	0.080
.. Conv. Store			-0.02 ***	0.001	-0.07	0.068
.. ATM			-0.01	0.000	0.01	0.050
.. Accepts vouchers			0.00	0.000	0.02	0.010
.. Sells oil			-0.01	0.001	0.03	0.094
Rack-level F.E.	No		No		Yes	
Post-estimation calc.						
.. ΔP to drike add. 1km	-0.33		-0.32		0.14	
	Mexicali	*Tijuana*	*Mexicali*	*Tijuana*	*Mexicali*	*Tijuana*
.. Mkt. price elast.	-4.15	-4.66	-0.56	-0.73	0.34	0.46
.. Avg. Market share (s)	0.36%	0.44%	0.36%	0.43%	0.36%	0.43%
.. Δs/Δd	-0.004%	-0.005%	-0.003%	-0.004%	0.000%	0.000%
.... Change in sales	-1.1%	-1.1%	-0.9%	-0.9%	0.0%	0.0%
.. Δs/Δbig business			-0.1%	-0.1%	0.054%	0.066%
.... Change in sales			-20.8%	-20.8%	15.2%	15.2%

Since the model is not linear, post-estimation calculations are evaluated at the mean.
(***) indicate significance at the 2.5% level, (**) at the 5% level, and (*) at the 10% level

Table 11: Estimation results of Logit models

Columns 3 and 4 in table 11 show the estimates for Logit Model 2. This specification controls for gas-station-level attributes.

$$u_{jk} = \gamma + \alpha y + x_{jk}\beta - \alpha p_{jk} - \lambda_1 d_{jk} - \lambda_2 d_{jk}^2 + \varepsilon_{jk} \tag{15}$$

The estimated parameters are statistically different from zero and considerably reduce the elasticity estimates. This supports the intuition that consumers value gas station-level amenities. However, a potential source of endogeneity could come from self-selection in which gas stations in affluent regions offer different attributes. To control for this, Logit Model 3 includes rack-level fixed effects, however, as is common with these types of controls, they absorb all price variation (Berry and Haile 2021).

References

Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (2022). Anuário Estatpistico Brasileiro do Petróleo, Gás Natural e Biocombustíveis 2022. Technical report, Agência Nacional do Petróleo, Gás Natural e Biocombustíveis, Rio de Janeiro.

American Petroleum Institute (2023). How many service stations are there in the United States?

Arena Pública (2019, January). Recortes y renuncias dejan a la Comisión Reguladora de Energía en los huesos. *Arena Pública*.

Argus Media (2020, October). Mexico's retail fuel permits still stalled.

Berry, S., A. Eizenberg, and J. Waldfogel (2016). Optimal product variety in radio markets. *RAND Journal of Economics 47*(3), 463–497.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile Prices in Market Equilibrium. *Econometrica 63*(4), 841–890.

Berry, S. and P. Reiss (2007). Chapter 29. In *Handbook of Industrial Organization*, Volume 3, pp. 1845–1886. Issue: 06 ISSN: 1573448X.

Berry, S. T. (1994). Estimating Discrete-Choice Models of Product Differentiation. *The RAND Journal of Economics 25*(2), 242–262.

Berry, S. T. and P. A. Haile (2021, September). Foundations of Demand Estimation.

Berry, S. T. and J. Waldfogel (1999). Free Entry and Social Inefficiency in Radio Broadcasting. *RAND Journal of Economics 30*(3), 397–420.

Borden, N. H. (1965). The Concept of the Marketing Mix. *Science in Marketing*, 389–390.

Borenstein, S. and J. Bushnell (2005, May). Retail Policies and Competition in the Gasoline Industry. Technical report, UC Berkeley: Center for the Study of Energy Markets, Berkeley.

Borenstein, S., A. C. Cameron, and R. Gilbert (1992). Do Gasoline Prices Respond Asymmetrically to Crude Oil Price Changes? *The Quarterly Journal of Economics 112*(1), 305–339.

Boscoe, F. P., K. A. Henry, and M. S. Zdeb (2012, May). A Nationwide Comparison of Driving Distance Versus Straight-Line Distance to Hospitals. *Professional Geographer 64*(2), 188–196.

Buczkowska, S., N. Coulombel, and M. de Lapparent (2019, December). A comparison of Euclidean Distance, Travel Times, and Network Distances in Location Choice Mixture Models. *Networks and Spatial Economics 19*(4), 1215–1248. Publisher: Springer.

Canadian Fuels Association (2023). Fuel retailing.

Castelletti, B. (2013, July). How redistributive is fiscal policy in Latin America? The case of Chile and Mexico. Technical Report Working paper 318, OECD, Paris.

Coglianese, J., L. W. Davis, L. Kilian, and J. H. Stock (2017). Anticipation, tax avoidance, and the price elasticity of gasoline demand. *Journal of Applied Econometrics 15*(January 2016), 1–15.

Colina, A. R. (2023, April). Consumer welfare and price controls in gasoline markets.

Comisión Reguladora de Energía (2016). Acuerdo que establece el cronograma de flexibilización de precios de gasolinas y diésel previsto en el artículo Transitorio Décimo Segundo de la Ley de Ingresos de la Federación para el ejercicio fiscal de 2017. Publisher: Comisión Reguladora de Energía Publication Title: Diario Oficial de la Federación Place: Ciudad de México.

Comisión Reguladora de Energía (2017, November). A partir del 30 de noviembre habrá libre comercio en los mercados de gasolinas y diésel en todo el país.

Conlon, C. and J. Gortmaker (2020). Best practices for differentiated products demand estimation with PyBLP. *RAND Journal of Economics 51*(4), 1108–1161.

D'Aspremont, C., J. Jaskold Gabszewicz, and J. Thisse (1979). On Hotelling 's " Stability in Competition ". *Econometrica 47*(5), 1145–1150. ISBN: 0012-9682.

Davis, P. (2006). Spatial competition in retail markets : movie theaters. *RAND Journal of Economics 37*(4), 964–982.

Diario Oficial de la Federación (1938, March). Decreto que expropia a favor del patrimonio de la Nación, los bienes mueble e inmuebles pertenecientes a las compañías petroleras que se negaron a acatar el laudo de 18 de diciembre de 1937 del Grupo No.7 de la Junta Federal de Conciliación y Arbitraje. Publisher: Ejecutivo Federal Place: Distrito Federal.

Diario Oficial de la Federación (1940, November). Decreto que adiciona el párrafo sexto del artículo 27 constitucional.-(Petróleo). Publisher: Poder Ejecutivo Place: Distrito Federal.

Diario Oficial de la Federación (1980). Ley del impuesto especial sobre producción y servicios. Publisher: Poder Ejecutivo.

Diario Oficial de la Federación (1995, March). Decreto por el cual se declara reformado el cuarto párrafo del artículo 28 de la Constitución Política de los Estados Unidos Mexicanos. Publisher: Ejecutivo Federal Place: Distrito Federal.

Diario Oficial de la Federación (2014, August). Decreto por el que se expide la Ley de Hidrocarburos y se reforman diversas disposiciones de la Ley de Inversión Extranjera. Publisher: Ejecutivo Nacional Place: Distrito Federal.

Dixit, A. K. and J. E. Stiglitz (1977). Monopolistic Competition and Optimum Product Diversity. *The American Economic Review 67*(3), 297–308.

Dorsey, J., A. Langer, and S. McRae (2022, March). Fueling Alternatives: Gas Station Choice and the Implications for Electric Charging. *NBER Working Paper Series 29831.*

Eckert, A. (2013, February). Empirical studies of gasoline retailing: A guide to the literature. *Journal of Economic Surveys 27*(1), 140–166.

Ellickson, P. B. and P. L. E. Grieco (2020). Measuring competition in spatial retail. *The RAND Journal of Economics 51*(1), 189–232.

Fan, Y. (2013). Ownership consolidation and product characteristics: A study of the US daily newspaper market. *American Economic Review 103*(5), 1598–1628.

Gandhi, A. and A. Nevo (2021). Empirical Models of Demand and Supply in Differentiated Products Industries. Technical report, National Bureau of Economic Research. Volume: 4 Issue: 1.

Gary, J. H., G. E. Handwerk, and M. J. Kaiser (2007). Industry Characteristics. In *Petroleum Refining. Technology and Economics*, pp. 18–19. Boca Raton, FL: Boca Raton : CRC Press.

Guerra, E. (2017, August). Does where you live affect how much you spend on transit? The link between urban form and household transit expenditures in Mexico City. *Journal of Transport and Land Use 10*(1).

Hotelling, H. (1929). Stability in Competition. *The Economic Journal 39*(153), 41–57.

Houde, J.-F. (2012). Spatial Differentiation and Vertical Mergers in Retail Markets for Gasoline. *American Economic Review 102*(5), 2147–2182.

Instituto Belisario Domínguez. Senado de la República (2015, November). Aspectos relevantes para la determinación del precio de la gasolina en 2016 y 2017.

Ito, K. and S. Zhang (2020). Willingness to pay for clean air: Evidence from air purifier markets in China. *Journal of Political Economy 128*(5), 1627–1672.

Kitamura, R. and D. Sperling (1987a, May). Refueling behaivour of automobile drivers. *Transportation Research Part A: General 21*(3), 235–245.

Kitamura, R. and D. Sperling (1987b). Refueling Behavior of Automobile Drivers. *Transportation Research Part A: General 21*(3), 235–245.

Li, S., J. Linn, and E. Muehlegger (2014). Gasoline taxes and consumer behavior. *American Economic Journal: Economic Policy 6*(4), 302–342.

Mankiw, N. G. and M. D. Whinston (1986). Free Entry and Social Inefficiency. *The RAND Journal of Economics 17*(1), 48.

McFadden, D. (1978). Modeling the choice of residential location in spatial interaction theory and planning models. In *Spatial Interaction Theory and Planning Models*, pp. 75–96. Amsterdam.

Moovit insights (2022, June). Moovit Public Transit Index.

Nevo, A. (2000). A practitioner's guide to estimation of random-coefficients logit models of demand. *Journal of Economics and Management Strategy 9*(4), 513–548.

Nevo, A. (2001). Measuring Market Power in the Ready-to-Eat Cereal Industry. *Econometrica 69*(2), 307–342.

ONEXPO Nacional (2021, January). Retrasos en permisos generan incertidumbre jurídica en el sector gasolinero, dice Onexpo.

Petrin, A. (2002). Quantifying the benefits of new products: The case of the minivan. *Journal of Political Economy 110*(4), 705–729.

Phibbs, C. S. and H. Luft (1995). Correlation of travel time on roads versus straight line distance. *Medical Care Research and Review 52*(4), 532–542.

Ray, P. (1973, September). Independence of Irrelevant Alternatives. *Econometrica 41*(5), 987–991.

Rushton, G., R. G. Golledge, and W. Clark (1967). Formulation and test of a normative model for the spatial allocation of grocery expenditures by a dispersed population. *Annals of the Association of American Geographers 57*(2), 389–400.

Secretaría de Energía (2017). Sistema de Información Energética.

Secretaría de Hacienda y Crédito Público (2015). Acuerdo por el que se da a conocer la banda de precios máxima de las gasolinas y el diésel para 2016 y otras medidas que se indican. Publisher: Poder Ejecutivo Place: Distrito Federal.

Seim, B. K. and J. Waldfogel (2013). Monopoly and Economic Efficiency: Evidence from the Pennsylvania Liquor Control Board's Entry Decisions. *American Economic Review 103*(2), 831–862.

S&P Global Platts (2020, February). Mexico energy project permit delays seen as blocking growth.

The Organisation for Economic Co-operation and Development- Secretariat (2013, November). Competition in Road Fuel 2013. Technical report, The Organisation for Economic Co-operation and Development, Paris.

Thomadsen, R. (2005). The Effect of Ownership Structure on Prices in Geographically Differentiated Industries. *The RAND Journal of Economics 36*(4), 908–929.

Verboven, F. (1996). International Price Discrimination in the European Car Market. *Source: The RAND Journal of Economics 27*(2), 240–268.

Waldfogel, J. (2003). Preference Externalities: An Empirical Study of Who Benefits Whom in Differentiated-Product Markets. *Source: The RAND Journal of Economics 34*(3), 557–568.

Wood, D. (2018, October). La nueva reforma energética de México. Technical report, The Wilson Center's Mexico Institute.