Shell shocked

Elliott, Rob; Sun, Puyang; Zhu, Tong

DOI:
10.1016/j.eneco.2020.104690

Citation for published version (Harvard):
Elliott, R, Sun, P & Zhu, T 2020, 'Shell shocked: the impact of foreign entry on the gasoline retail market in China', *Energy Economics*, vol. 86, 104690, pp. 1-16. https://doi.org/10.1016/j.eneco.2020.104690

Shell shocked: The impact of foreign entry on the gasoline retail market in China

Robert Elliott[1], Puyang Sun[2], and Tong Zhu*[3]

[1] *University of Birmingham, UK*

[2] *Birmingham-Nankai Joint Research Centre, UK; Centre for Transnationals Studies of Nankai University, China; School of Economics, Nankai University, China*

[3] *The Economic and Social Research Institute, Ireland; Trinity College Dublin, Ireland*

Abstract

Since joining the WTO in 2001 restrictions on foreign entry into China's energy sector have been steadily reduced. We investigate the impact of Royal Dutch Shell's entry on the pricing behavior of three varieties of gasoline in the retail market of China. Using a difference in difference pairwise estimator we show that a year after entry, the average absolute price differential of gasoline between two cities increased by around 1.4 percent, before falling the following year. In other words, Shell's entry caused prices to diverge but only for a short period of time. The largest price effect was found for highly refined fuels in Western cities. Similar results are found when we examine the effect of entry of the top four foreign retailers. Policy implications are discussed.

JEL: L130, L810, E310

Keywords: Royal Dutch Shell; price dispersion; gasoline; energy

*Corresponding author. The Economic and Social Research Institute, Whitaker Square, Sir John Rogerson's Quay, Grand Canal Dock, Dublin 2, Ireland; tong.zhu@esri.ie.

1 Introduction

As one of the most rapidly growing consumer markets in the world, China has received considerable attention from multinational retailers. China's retail sales in 2018 reached 38.10 trillion RMB (US$ 5.46 trillion in 2018 prices), 9.0 percent higher than in 2017.[1] As a result, China's growing consumer market presents an enormous opportunity for both domestic and foreign firms. One particular difficulty faced by foreign firms has been a lack of market access to certain sectors which have been subject to strict government regulation on the degree to which foreign investment is permitted. These restrictions were particularly prevalent in the retail gasoline sector which is considered to be one of the most profitable in China. As the world's second largest oil consumer, and since 2010 the world's largest energy consumer, China's demand for conventional gasoline has grown rapidly and is predicted to continue to grow further as automobile ownership expands.

Prior to China's entry into the World Trade Organization (WTO) very little foreign investment was permitted in the energy sector. However, following a series of concessions agreed by China as part of the WTO negotiations, entry barriers in the gasoline retail sector were gradually eliminated. China's strategy for the energy sector is based on the premise that the economy and consumers would benefit from greater market-oriented competition, enhanced brand awareness and improvements in non-fuel service quality. The most significant policy change came in December 2004 when the *Measures for the Administration on Foreign Investment in Commercial Sector* was issued by the Ministry of Commerce. This new ruling permitted, for the first time, foreign-funded retail enterprises to establish gasoline retail service stations and allowed foreign firms to buy and sell refined oil without geographical restrictions (Ministry of Commerce, PR China). The 2004 deregulation was followed in December 2006 by a similar measure that opened up the wholesale market for refined oil to foreign firms after which China's energy market was considered to be fully open.

The contribution of this paper is to be one of the first to examine the impact of China's open energy policy on the retail gasoline market. Our research follows in the tradition of Basker (2005) who examines Wal-Mart's entry into local markets on local prices in the US and Holmes et al. (2013) who investigate the determinants of gasoline price integration in the US. More specifically, we provide an insight into how the entry of a large multinational company, in this case Royal Dutch Shell (hereafter Shell), affected retail energy prices and the pattern of energy price convergence within

[1] Retail sales of oil and refined products increased by 13.3%, registered the second highest percentage increase after daily necessities (National Bureau of Statistics of China, 2019).

China. Our data are derived from the combination of two unique datasets. The first is the location and opening year of China's Shell branded service stations opened either by Shell alone or as part of a joint venture. The scale of Shell's activity coupled with their early entry into China makes their experience an interesting quasi-experiment by which to investigate various aspects of energy price dynamics in China. Given the potential impact of the entry of other foreign retailers, we also collected the location and opening information for British Petroleum (BP), Total S.A. and Exxon who also made positive but less significant inroads into China's retail market during the same period. The second is a dataset of service station-level fuel prices for 2001 to 2012 obtained from the China Price Monitoring Centre (CPMC) which is a branch of the National Development and Reform Commission (NDRC). To take into account heterogeneity in the quality across varieties of fuel we consider three different grades of gasoline: super premium (Gas#97); premium (Gas#93); and standard (Gas#90).[2]

Our methodological approach is to combine a pairwise approach with a difference-in-differences (DID) estimator to identify the causal relationship between Shell's entry into a local market and the subsequent effect on absolute price differentials between cities. Following Pesaran et al. (2009) and Holmes et al. (2013) we first calculate the absolute price differential between gasoline prices for any two cities which have similar propensities to be populated by Shell in a given year. We then take a DID approach and compare the absolute price differential between two cities where one city has experienced Shell's entry (the treated) and the absolute price differential between two cities where neither has a Shell station (the control), and compare the price differentials before and after the entry year. Besides Shell's entry, we also account for the impact of distance between cities and a range of economic and demographic factors. Considering differentials in the macroeconomic data helps to alleviate concerns of endogeneity as the entry decision of international retailers is unlikely to be based on the economic development gap between any two cities. The pairwise approach is relatively new but is commonly used in studies that investigate price behaviour (see e.g. Pesaran et al., 2009; Yazgan and Yilmazkuday, 2011; Holmes et al., 2011, 2013). Through taking the differential in gasoline prices between any two pairs of cities, the pairwise approach is reasonably robust to cross-section dependence (Pesaran et al., 2009). For example, since within our time period there was considerable energy market deregulation leading to a decentralization of prices, taking differences enables us to eliminate the impact of common shocks that effected all cities in the same year. In addition, taking the difference between any two cities, instead of the difference relative to the group mean or a baseline city, enables us to avoid any undesirable shocks from a specific year or the baseline city. Finally, the pairwise approach allows us to investigate the impact of

[2]Descriptions of each gasoline type are provided in the data section of the paper.

distance on price dispersion.

To briefly summarize our results, we find that Shell's entry into China's retail gasoline market had a statistically significant effect on gasoline prices across China. Specifically, our results suggest that the entry of Shell into a city increased the average dispersion of prices in China between 1.3 and 1.5 percent in the year of entry although, crucially this positive impact disappears over time. Of the three different fuel types, we find a larger effect for super premium gasoline (Gas#97), that historically had higher profit margins both locally and nationally. To explain the dynamics in price dispersion we then look at how Shell's entry impacts prices taking spatial variation into account. The greatest price impact was found in China's less developed western cities which, due to limited competition in the past, would have been exposed to greater competitive pressures following Shell's entry, certainly compared to the already relatively more mature markets in eastern cities.

The remainder of the paper is organized as follows. Section 2 provides a review of the literature. Section 3 describes the data. Our empirical approach is presented in Section 4 and our empirical results are provided in Section 5. Section 6 investigates potential mechanisms that are driving the entry effect. The final section concludes.

2 Literature Review

Price dispersion is an important indicator of market efficiency. The last thirty years has seen considerable growth in research examining price dispersion among related products or customers with different preferences. In a monopolistic competition setting, assuming firms can distinguish between customers, firms will price discriminate based on demand elasticities with the result that monopolistically set prices tend to be higher than those under perfect competition. As a result, firms with pricing power set prices above marginal costs. In an imperfectly competitive setting, prices should gradually converge as the entry of new firms leads to greater competition. With new entry (and no exit), price differences should disappear and the average market price should return to the marginal cost level. This brief literature review discusses the various mechanisms by which Shell's entry into China could affect absolute price levels for gasoline and levels of price dispersion within and across cities.

There are numerous explanations why price dispersion can persist across firms within an industry. One sector that has been analyzed in detail is the airline industry. For example, Borenstein and Rose (1994) examined the US airline industry in 1986 and surprisingly found evidence of a positive relationship between price dispersion and competition. They highlighted two main sources of price differences. The first is that variation in prices is driven by the cost-base of firms (airlines serving different passengers in different locations could have a different cost base depending on where the airline is geographically located). Second, prices could differ due to price discrimination usually as a result of differences in market structure, consumer preferences and product characteristics. More recently and in line with traditional microeconomic theory, a second study of the US airline industry by Gerardi and Shapiro (2009) that coincided with the entry of the low-cost carriers finds a significant negative relationship between competition and price differences.[3]

A smaller body of research examines the competition effect of the entry of multinational retailers into local markets including Chen (2003), Basker (2005), Dukes et al. (2006), Hausman and Leibtag (2007) and Gielens et al. (2008). As one of the largest players in international retailing, Wal-Mart often serves as a case study. For example, Basker (2005) studies the entry effect of Wal-Mart by combining quarterly price data for a series of commonly bought products after the opening of Wal-Mart stores across the US and finds that after a store is opened there are significant price reductions for several daily consumables such as toothpaste and shampoo. The magnitude of the reduction tended to be in the range of 1.5 to 3 percent in the short term and up to four times higher in the long run. Drawing upon studies on the Wal-Mart effect, two mechanisms are highlighted. First, Wal-Mart's highly efficient supply chain reduces operating costs with Walmart's superior logistics, infrastructure and distribution networks regarded as important components of their retail strategy. In addition, Wal-Mart's high sales volume means that it is generally regarded as a principal buyer so that it is able to obtain concessions from manufacturers including preferential wholesale prices (Dukes et al., 2006). An indirect effect of Walmart's purchasing power is that suppliers also appear to lower the prices they charge to other retailers in the same locale who stock similar goods (Chen, 2003). A second mechanism is through an increased competition effect in local markets. Hausman and Leibtag (2007) find a price effect from new competition when they examine customer benefits after the entry and expansion of Wal-Mart with the downward pricing pressure on incumbents being greater, the geographically closer they are to a new store. Wal-Mart's entry also had the effect of reducing wholesale prices. Gielens et al. (2008) argue that incumbents need to consider specialized and exclusive selling strategies or increase service augmentation to remain competitive given Wal-Mart's emphasis on low prices.

[3] One reason given by Borenstein and Rose (1994) to explain their positive findings was the widespread use of frequent flier points (FFPs) such that new entrants affect the less-frequent customers market but have little effect on frequent fliers due to high levels of brand loyalty.

Studies of the gasoline retail market are more limited and have tended to focus on developed countries, such as the US and Canada. For example, Eckert and West (2004) investigate the difference in pricing behavior between Ottawa and Vancouver and argue that the substantial differences in pricing are likely to be caused by differences in market structure and market conduct. The authors point out that tacit collusion and coordination behavior contributed to short-run price stickiness and long-run price uniformity in Vancouver whereas in Ottawa the entry of maverick independent gasoline retailers eroded existing collusive behavior as new competitors sacrificed short run profits and undercut rivals to gain market share. A study of independent retailers in Southern California by Hastings (2004) following the acquisition of the Thrifty Oil Company in 1997, finds that retail prices rose once independent gasoline stations were converted into fully vertically integrated service stations. Their quasi-natural experiment predicts a 5 cent increase in the average retail price following the loss of one independent gasoline station. The ability of branded gasoline stations to exploit brand loyalty was also given as an explanation for an increase in prices as market share increases. Using station-level gasoline price data across the San Diego, San Francisco, Phoenix, and Tucson areas, Barron et al. (2004) find that an increase in the number of competitors in a market consistently decreases both price levels and price dispersion.

Based on Hastings and Gilbert (2005) study of the US West Coast gasoline market, vertical integration is argued to be the primary reason for regional wholesale price differentials. In the market where integrated refiners have a large market share, they have a positive incentive to raise the unbranded wholesale price for its retailing rivals. In a study of retail and wholesale gasoline prices in the US, Zimmerman (2012) finds that hypermarkets with a gasoline retail service exerted considerable downward pressure on nearby retail gasoline prices. Employing a pairwise approach, Holmes et al. (2013) find cointegration in regional gasoline prices in the US. Their results suggest that the distance between states plays an important role in determining the speed that prices converge. Yilmazkuday and Yilmazkuday (2016) list four sets of determinants of price dispersion in the retail gasoline sector: competitor density and station concentration levels; station characteristics (whether they have convenience stores or repair services); local demographic and economic conditions; and brand or contractual agreements.

A second strand of the literature examines the asymmetric movement of gasoline retailing prices. Maskin and Tirole (1988) first described the saw-toothed pricing pattern and developed the Edgeworth price cycle theory. Under a duopoly with alternating moves, competitors selling homogenous goods tend to undercut each other gradually until they approach the wholesale price. Following a war of attrition, one competitor relents and retail prices rebound rapidly before starting the next round of price cuts.

Empirical evidence for a price cycle effect, based on high frequency retail prices, has been found in mature gasoline markets worldwide such as in Canada (Noel, 2007a,b, 2009; Eckert, 2002, 2003; Eckert and West, 2004; Byrne and de Roos, 2017), the US (Lewis, 2009; Lewis and Noel, 2011) and Australia (Bloch and Wills-Johnson, 2010; Byrne, 2012). Independent (non-branded) small firms have the greatest incentive to begin undercutting their larger rivals while large firms normally take the leadership in restoring prices (Eckert, 2003; Byrne et al., 2013). The Edgeworth cycle revolves faster, with a greater magnitude with less asymmetry in markets with more independent competitors (Noel, 2007a).[4]

Finally, a small number of studies have examined the behaviour of energy prices in China. Ma et al. (2009) use 10-day energy spot price data for 35 major cities to investigate the extent of price convergence across China graphically and parametrically. The data reveal considerable price variation across the 31 mainland provinces for four major fuels (coal, electricity, gasoline and diesel) with price differences driven, to a large extent, by different energy reserve levels and transportation costs. As a consequence, cities close to gasoline production areas are predicted to enjoy lower retail prices. While gasoline and diesel prices behave fairly consistently, spatial price dispersion is still observed, implying that the energy market for China remains segmented. Ma and Oxley (2012) confirm the result that energy prices in certain regional energy markets in China have converged but that this tends to be within groups of regions rather than across the whole country. Testing for club convergence and clustering, the evidence shows that several regional clusters for gasoline exist simultaneously but that they are becoming generally geographically connected as energy markets have become more open.

3 Data

3.1 Royal Dutch Shell in China

We begin with a brief review of the retail gasoline sector in China. Despite the complicated licensing process, multinational oil giants such as Shell, BP and Exxon have managed to access China's energy markets and have expanded their networks to include not just downstream retail gasoline service stations but also upstream refinery and fuel supply businesses. The reasons we focus on Shell are two-fold. First, Shell has

[4] See Eckert (2013) for a detailed summary of empirical studies that examine different aspects of gasoline retailing.

the largest retail service network of the big four foreign companies (Shell has 862 stations in 48 cities compared to BP's 291 stations in 26 cities, Total's 202 stations in 41 cities, and Exxon's 14 stations across 10 cities). Second, Shell's official website provides detailed information on the location of each retail gasoline station, while information for the other three companies is not so readily available. Although we focus primarily on Shell, we also consider the entry of Shell and the other top three foreign companies to capture the entry effect of international firms more generally.[5]

From the late 1990s onwards Shell began to build jointly branded gasoline stations with local partners across a number of Chinese cities. The first Shell station was established in Tianjin in 1997, with 25% of the equity held by Tianjin Nongken State-owned Company. Shell proceeded to develop a retail network that now covers the majority of the metropolitan areas in the northern, southern, central and southeastern provinces. Shell's primary strategy was to find a local partner and then initiate a joint company within a given geographical area (which tended to be at the province level). By the end of 2018 Shell operated approximately 1,400 retail gasoline retail stations in China. The second stage of development was to build new service stations and to take over existing unbranded service stations whilst continuing to seek new partnerships. At the current time, all of China's major national oil companies have collaborative agreements with Shell, including China National Petroleum Corporation (CNPC), China Petroleum & Chemical Corporation (Sinopec), China National Offshore Oil Corporation (CNOOC) and Shaanxi Yanchang Group Company (YCG). These four companies are also the only ones that own petroleum exploration licenses for China.

The service station data for this paper are available from the Shell website.[6] Figure 1 provides a geographical overview of Shell's gasoline service station footprint for 2013. The larger the symbol the higher the number of service stations in the city. The largest service station networks are in Shaanxi, Sichuan and Hebei provinces. For example, Yanchang and Shell Petroleum Company own more than 290 service stations in Shaanxi and Yanchang and Shell (Sichuan) Petroleum Company own 153 stations. In total Shell has a presence in 9 of China's 31 provinces.[7] Table 1 complements Figure 1 by providing a summary of Shell's service station network also for 2013 including

[5]We managed to collect the location and opening information for BP, Total and Exxon from online sources although we are less confident in the accuracy of our data. Although we always included stations from the other three companies in the paper, thanks to the comment from two anonymous referees we now include the results in the main body of the paper.

[6]See http://www.shell.com.cn/en/aboutshell/our-business-tpkg/china.html for details.

[7]Shell has different partners in different regions for both upstream and downstream operations. According to the official website, there are five joint companies for the retailing business in China. They are Beijing Shell Oil Co. covering Beijing and Chengde city, Shell North China Petroleum Group Co. covering Tianjin, Hebei and Shandong, Yanchang and Shell (Guangdong) Petroleum Co. covering Guangdong, Yanchang and Shell (Sichuan) Petroleum Co. covering Sichuan, Yanchang and Shell Petroleum Co. covering Shaanxi and Shanxi.

a number of other variables of interest. There is no obvious relationship between the location of Shell's gasoline stations and the production of gasoline per million people or per million vehicles. Figures 2 and 3 show the number of retail gasoline service stations per million people and per million vehicles for each of the 9 provinces with a Shell retail presence. Figure 4 shows the relationship between the gasoline production/surplus level and Shell's entry decision. A surplus in this context means that the province is a net exporter of gasoline products and a negative value implies that the province is a net importer of gasoline. The raw data suggests that Shell does not have an obvious tendency to locate its retail network in relatively high fuel production areas such as Shaanxi, Hebei and Shandong as the majority of provinces that Shell has chosen to locate are net fuel importing provinces such as Sichuan, Tianjin, Shanxi, Chongqing, Guangdong and Beijing. Table A1 in the appendix lists each province ranked by gasoline production as well as consumption and surplus/deficit values.

[Figure 1, 2, 3 and 4 about here]

[Table 1 about here]

Within the 9 provinces listed in Table 1, Shell has a presence in 48 prefecture-level or above cities.[8] However, we only have gasoline price data for 11 out of these 48 cities.[9] In order to have both pre-Shell and post-Shell price informatio, we only include cities where the first entry year is between 2001 and 2012. This means excluding Tianjin (first entry year 1997) and Jinan (first entry year 2013). Our Shell-presence group is therefore 9 cities across 6 provinces with a first year of entry between 2001 and 2012. The nine cities are Beijing, Taiyuan, Guangzhou, Shenzhen, Chongqing, Chengdu, Zigong, Xian and Baoji. The dates for Shell's first year of entry and the number of stations are listed in Table 2. Our non-Shell group includes cities which have never previously had a Shell gasoline station. Tianjin (first entry year 1997) and Jinan (first entry year 2013) are excluded from both groups.

[Table 2 about here]

3.2 Gasoline Prices in China

We now provide a brief overview of the gasoline pricing mechanism in China. To prepare for entry into the World Trade Organization (WTO), China adopted the world

[8] According to China.org.cn "*Prefectural-level cities are large and medium-size cities not including sub-provincial level cities. Normally, they are cities with a non-farming population of more than a quarter of a million.*" Source: http://www.china.org.cn/english/Political/28842.htm.

[9] Our gasoline price dataset includes prices for 62 cities in total.

price of oil in June 1998. The pricing reform scheme issued by the State Development Planning Commission (SDPC) (renamed as National Development and Reform Commission in March 2003) proposed an adjustment to the monthly prices of crude oil and refined products based on Singapore market prices although the adjustment frequency did not become monthly until June 2000. China's previous pricing mechanism was a two-tier system which had remained unchanged since the early 1980s. In 2001 the Chinese government went further and linked domestic prices to international prices based on prices from the Singapore, Rotterdam, and New York markets with weights given by 6:3:1 respectively. At the same time crude oil prices were set by the two largest vertically-integrated oil companies instead of the SDPC. However, the government continued to provide guidelines and would make adjustments when the gap between domestic and international prices became too large. A new formula for setting baseline prices based on prices from the Brent, Dubai and Minas markets was introduced in January 2007. More recently there has been a narrowing of the interval and an increased frequency for price adjustments to be made. The latest refined products pricing scheme issued on March of 2013 allows the mutually supplied crude oil prices to be set jointly by Sinopec and PetroChina based on local refining conditions. Crude oil produced by CNOOC and other producers are set by the companies themselves. In the retail market, gasoline service station retailers can set their own prices as long as they are below a maximum price set by the NDRC for individual provinces and some major cities.

Generally speaking, each province has its own price cap. For example, the NDRC announced a decrease of 540 yuan/ton for gasoline retail price on November 30th 2018. Shaanxi province then announced a decrease of 0.41 yuan/liter for gasoline, with a maximum price of 7.16 yuan/liter for the middle and north of the province and 7.24 for the south of the province. At the same time, Guangdong province set a maximum price of 6.86 yuan/liter for the province as a whole for the same product. A common perception of the energy market in China is that it is highly regulated and as such domestic fuel prices are underpriced relative to the world price which in turn encourages greater domestic consumption with implications for congestion and air pollution (Hang and Tu, 2007; Tan and Wolak, 2009). Drawing lessons from a widespread power shortage in 2008, Wang et al. (2009) argue that the centrally planned policy-oriented pricing mechanism encouraged excessive growth in heavy industry, and hence was inconsistent with a strategy to promote energy-conservation and sustainable development.

This paper analyzes the pricing behavior of three different grades of fuel. The current industrial classification for automotive gasoline products in China is GB 17930 "Automotive Gasoline". Within this classification there are three type of gasoline that we label Gas#90, Gas#93 and Gas#97. The number associated with each product represents the octane rate. For example, Gas#97 gasoline contains 97% iso-octane and 3% n-heptane. The gasoline octane number is designed to be compatible with the compression ratio of the engine. If an engine with a high compression ratio is used with a low octane gasoline the engine tends to knock, which can cause problems such as piston splintering and piston ring breakage. If an engine with a low compression ratio is used with a high octane gasoline, the ignition timing may be altered, which can cause incremental sedimentation inside the engine cylinder thus shortening engine life (Sinopec Corp. 2014). The compression ratio is an important structural parameter for the engine. Generally speaking, a high compression ratio indicates high thermal efficiency giving the automobile more power, greater acceleration and a higher maximum speed. Hence, high performance cars need higher quality gasoline (a higher octane percentage).

The initial gasoline price data is a panel dataset collected by the China Price Monitoring Centre (CPMC) on every 5th, 15th and 25th day of each month. From 2001 to 2012, spot prices were obtained from several (normally 2-5) permanent anonymous services stations for a number of different gasoline varieties. Since there is no brand information for the stations from which the price information is taken, we simply take the arithmetic mean between monitoring stations as the market average price for that city. For our analysis, we aggregate spot price data to an annual average level for the following reasons. First, the gasoline pricing regime and the price cap mechanism during the study period have reduced the speed of price adjustment significantly. Figure A1 in the Appendix presents the spot price trend from the first anonymous station in each of our sampled cities. For most cities the price changes 2 to 5 times a year even though the price information is collected every 10 days.[10] Using higher frequency data would also result in a large number of repeated records and may lead to biased estimations of Shell's entry effect. Second, due to a lack of the precise date of Shell's entry into a city, using annual price data enables us to match the entry as well as other economic and demographic variables on a yearly basis. Finally, many retail sectors show a strong seasonality and using annual data avoids undesired seasonal fluctuations.

[10]A similar example can be found in Figures 6 and 7 of Ma et al. (2009)

4 Identification Strategy

Our methodological approach is to use a difference-in-differences (DID) pairwise estimator to help us identity the causal relationship between price dispersion across cities and the entry of Shell into a local market. We first match each of the Shell-presence cities with several non-Shell station cities using propensity score matching (PSM) techniques. The matching procedure includes a number of macroeconomic control variables to address self-selection concerns related to the difference in development levels between cities.[11] Second, following Pesaran et al. (2009) and Holmes et al. (2013), we take the absolute gasoline price differentials between any two cities in a group including one Shell-presence city and its corresponding matched non-Sell cities. Since there was considerable deregulation in China's energy market during our period of study, taking differences removes the effect of common shocks that impacted all cities in a given year. By taking differences relative to another city or cities instead of simply taking the group mean also allows us to examine the impact of distance on price dispersion. Finally, we adopt the DID method on the bivariate price differentials. The coefficient on the DID interaction term measures the entry effect of Shell on the price dispersion in a given year in China. In this sense, a DID model on differential data is a triple difference model.

Our PSM approach matches on a range of city-level characteristics which are GDP per capita, population density, average wages, passenger traffic by highway and the price for Gas#90.[12] The results of our initial matching are presented in Table 3. In Table 4 we present the results from post-match balancing tests which assess the accuracy of our PSM matching method for individual variables as well as the overall covariate balance. The results show that the balancing assumption is satisfied both individually

[11]The usual DID approach selects the treatment and control groups based on a different application of a policy across, for example, geographical regions where a region is allocated to either a treatment group (policy applied) or a control group (no policy applied). However, if the regions selected to receive intervention were picked because of certain regional characteristics, the treated and control groups may have different pre-policy trends. Since conducting a natural experiment in economic research is rare, finding an efficient counterfactual for the treated units of analysis becomes a precondition for the construction of the DID framework. PSM allows the researcher to generate matched controls for each treated unit of analysis with post-matching balancing tests used to ensure that for a given PSM, the DID approach is appropriate.

[12]See Table A2 in the appendix for definitions of our matching variables. All monetary values are deflated in 2005 prices using a province level Customer Price Index (CPI). Estimation of propensity scores using a logit estimation are presented in Table A3. Overall coefficients were significant and the pseudo-R square reveals that the equation explains 21.9% of the variation and indicate the potential benefit of matching the sample according to propensity scores (Baser, 2006). The Gas#90 price is included in the matching as a benchmark of local gasoline price level and is insignificant in both the logit regression and post matching tests. Although the local gasoline price is not an important consideration for Shell to make the entry decision, results from the logit regression suggest that Shell's presence tends to be associated with low gasoline prices.

and overall.

[Table 3 about here]

[Table 4 about here]

After matching we adopt a pairwise approach and calculate the extent of any absolute price differences between a city with a Shell presence and its corresponding matched controls. The pairwise approach was developed in Pesaran (2007) to test for output and growth convergence and Pesaran et al. (2009) to test the Law of One Price across 50 countries for the period 1957 to 2001. Holmes et al. (2013) also employ a pairwise approach to examine gasoline market integration in the US and find strong evidence of gasoline price convergence.[13] For cities i and j in group g, the price differential of each of our gasoline fuel types z in time t is defined as:

$$Dispersion_{z,ij,t} = |log(Price_{z,i,t}) - log(Price_{z,j,t})| \tag{1}$$

For convenience we use pair p to denote a combination of city i and j.

$$Dispersion_{z,p,t} = |log(Price_{z,i,t}) - log(Price_{z,j,t})| \qquad p = 1, 2 \ldots \sum_{g=1}^{9} C_{G_g}^{2} \tag{2}$$

group g contains one treated city and $(G_g - 1)$ control cities which implies $\sum_{g=1}^{9} C_{(G_g-1)}^{1} = \sum_{g=1}^{9}(G_g - 1)$ treated pairs and $\sum_{g=1}^{9} C_{(G_g-1)}^{2}$ control pairs.[14]

Before estimating the causal relationship between price dispersion and foreign entry, we characterize some general patterns of price dispersion that we find in the data. Following Fan and Wei (2006) and Holmes et al. (2013) we examine fluctuations in retail gasoline prices taking into account the distance between cities and a time trend estimate by:

$$Dispersion_{z,p,t} = \gamma_0 + \gamma_1 Trend_t + \gamma_2 Distance_p + \gamma_3 Distance_p^2 + \varepsilon_t \tag{3}$$

Since one of our key testable hypotheses is whether price dispersion changed after Shell's entry, for each of our pair of cities, the treatment effect after s years of entry is

[13] To be precise, the pairwise approach used in our paper is similar to the the pairwise method mentioned in Pesaran (2007), Pesaran et al. (2009) and Holmes et al. (2013). In these three papers a variety of unit root tests are conducted on all possible $N(N-1)/2$ bivariate differentials between any pairs of N countries or states.

[14] For example, there are seven cities in the first group $G_1 = 7$. Taking a combination of any two cities among those seven cities leads to $C_{G_1}^{2} = C_7^2 = 21$ distinct pairs of cities, among which $C_{G_1-1}^{1} = C_6^1 = 6$ pairs are the treated and $C_{G_1-1}^{2} = C_6^2 = 15$ pairs are the controls.

measured as:

$$Dispersion^1_{z,p,t+s} - Dispersion^0_{z,p,t+s}$$

where the superscript indicates whether a pair of cities is being treated ($Treatment = 0,1$) and the subscript $t+s$ implies $s\,(s = 0,1,2)$ years after the entry date t. However, the second term in the equation is unobservable. Based on the results of our matching procedure, each treated city has several matched controls which have similar economic and demographic characteristics. Therefore, the price dispersion between two valid control cities within one group are regarded as the appropriate substitution for the unobservable term. As a result, the average treatment effect on the treated (ATT) given by:

$$\begin{aligned} ATT &= E[Dispersion^1_{z,p,t+s} - Dispersion^0_{z,p,t+s} | Treatment_{p,t} = 1] \\ &= E[Dispersion^1_{z,p,t+s} | Treatment_{p,t} = 1] - E[Dispersion^0_{z,p,t+s} | Treatment_{p,t} = 1] \\ &= E[Dispersion^1_{z,p,t+s} | Treatment_{p,t} = 1] - E[Dispersion^0_{z,p,t+s} | Treatment_{p,t} = 0] \\ & s = 0,1,2 \qquad p = 1,2 \ldots \sum_{g=1}^{9} C^2_{G_g} \end{aligned}$$

Hence, our pairwise DID estimator, based on the matched sample, is given by:

$$ATT = \frac{1}{\sum_{g=1}^{9}(G_g - 1)} \sum [\Delta_s Dispersion^1_{z,p} - \Delta_s Dispersion^0_{z,q}]$$

where $\Delta_s Dispersion^1_{z,p}$ is the difference in absolute price differential after s years of the entry for the treated pair p, and $\Delta_s Dispersion^0_{z,q}$ is the difference in absolute price differential after s years of entry for a control pair q.

Disentangling the different sources of price dispersion is important if we are to identify the causal relationship between Shell's entry and the subsequent behavior in gasoline prices. Based on Vita (2000) and Zimmerman (2012), a vector of covariates are taken into account in the measurement of price dispersion including income per capital, population density, average wage per worker, population at the end of the year and highway passenger traffic. According to Borenstein and Rose (1994), price dispersion may arise from two sources, namely cost variations and discriminatory pricing. However, it is difficult to separate cost-based dispersion from discrimination-based dispersion because of market heterogeneity. For example, on one hand, management costs and transaction costs per sale could be lower in high population density cities than in smaller cities due to economies of scale and implies that retailers can set lower prices in these locations to acquire greater market share, even by compensating for losses in smaller cities. On the other hand, the demand elasticity for fuel may be lower in larger cities which provides a greater incentive to price discriminate. Hence, we control for the relationship between price dispersion and factors that might indicate a basis either for cost variation or discrimination within a DID framework. Considering the possibility of a delayed impact of Shell's entry and the fact that the first Shell station might only have a minor effect on the local market, we include lags and the corresponding interaction terms for Shell's entry decision. We therefore estimate a standard and an extended model given by.

Standard framework

$$\begin{aligned} Dispersion_{z,p,t} &= \alpha + \beta_1 Treatment_p + \beta_2 Post_{p0} + \beta_3 Treatment_p \times Post_{p0} \\ &+ \boldsymbol{\Delta_p Control \cdot \gamma} + \sum_p \gamma_p Pair_p + \sum_t \delta_t Year_t + \varepsilon_{z,p,t} \end{aligned} \tag{4}$$

Extended framework

$$\begin{aligned} Dispersion_{z,p,t} &= \alpha + \beta_1 Treatment_p + \beta_2 Post_{p0} + \beta_3 Treatment_p \times Post_{p0} \\ &+ \beta_4 Post_{p1} + \beta_5 Treatment_p \times Post_{p1} + \beta_6 Post_{p2} + \beta_7 Treatment_p \times Post_{p2} \\ &+ \boldsymbol{\Delta_p Control \cdot \gamma} + \sum_p \gamma_p Pair_p + \sum_t \delta_t Year_t + \varepsilon_{z,p,t} \end{aligned} \tag{5}$$

Subscript z represents our three gasoline products Gas#90, Gas#93 and Gas#97 described in previous section. Subscript p indicates each combination of two cities. $Treatment_p$ is a binary variable indicating whether one city in the pair has been or will be entered by Shell. $Post_{p0}$, $Post_{p1}$ and $Post_{p2}$ are dummy variables representing the entry year, one year and two years after entry respectively. $\boldsymbol{\Delta_p Control}$ represents the covariate matrix including the differences in income per capital, population density, average wage of workers, population at the end of year and passenger numbers travelling by highway. Our measure of the absolute differences in the independent variables is calculated using the same order between cities as we used to generate our price dispersion variable, which were randomly assigned. $Pair_p$ and $Year_t$ represent pair-specific fixed effects and time fixed effects respectively and $\varepsilon_{z,p,t}$ is the error term. β_3 in the standard framework indicates the static, or aggregated, impact of Shell's entry on price dispersion, while β_3, β_5 and β_7 in the extended framework captures the dynamic impact across multiple time periods. Table 5 presents a summary of our differenced variables and shows, for example, that the average price dispersion for the Gas#90 retail price between any two matched cities in our sample is approximately 4% and is slightly lower than those for Gas#93 and Gas#97.

[Table 5 about here]

Finally, we include data from the top four international retailers' entry in China combined with our gasoline price dataset to examine the broader impact of foreign entry on gasoline price convergence. A city is labelled as a treatment city if any one of the four retailers opened a gasoline retail station in that city. The treatment date was set as the year of the first retailer's entry. The location and entry date information are collected manually on Baidu Map and local newspaper/news website. Table A4 in the appendix provides a summary of the network distributions of the four retailers in China.

5 Results

Table 6 presents the results for a simple OLS regression examining the impacts of a linear trend and distance on price dispersion. The spatial distance between pairwise cities is calculated using the great circle formula according to information on the latitude and longitude of the centroid of the cities. The immediate observation is the importance of distance. The further two cities are away from each other, the greater the absolute price difference. In columns (4) (5) and (6) where we also include squared distance terms, we find a positive impact of distance on price dispersion for all of our three products but at a decreasing rate.[15] Our finding is in line with Holmes et al. (2013) who find an asymmetric relationship between the distance and the speed at which gasoline prices converge to the long-run equilibrium. The negative and significant time trend for our period of analysis also suggests a general reduction in price dispersion for Gas#93 and Gas#97 but a significant and positive effect on price dispersion for Gas#90. In the following analysis distance is considered as a pair-specific effect and year-specific effect accounts for the linear trend. Other controls are also included.

[Table 6 about here]

Table 7 presents our difference-in-difference estimation results for the determinants of retail price dispersion for the standard framework in equation (4). Columns (1), (2) and (3) present the results from a fixed-effect regression for each of our three gasoline products using the matched sample concerning Shell's entry into a city. Columns (4),

[15] The negative estimated coefficients on $Log(Distance)^2$ suggest that after the turning point distance is negatively associated with price dispersion. In our sample a large majority of the observations are to the left of the turning point. For example, the turning point based on estimates in Column (4) of Table 6 is at around 4,447 km (=exp(0.0690/(2*0.0041))). The distance between two cities ranges from 137 km to 7,011 km, with 89 percent of the observations having a distance below 4,447 km.

(5) and (6) present the fixed-effect regression results using an alternative sample that captures the entry of all four foreign retailers (Shell, BP, Total and Exxon).[16] Year specific and individual pair effects are included in all regressions. Coefficients are estimated with robust standard errors clustered at individual pair level. The coefficient on the interaction term $Treatment \times Post_0$ is our main variable of interest, and captures the static impact of foreign entry on the retail price dispersion controlling for a number of additional social economic factors.

The results in Table 7 reveal no significant impact of Shell's entry on gasoline price dispersion during the study period. The estimated coefficients of the difference-in-difference interaction term tend to have very large p values, suggesting the impact of foreign entry on price dispersion is not statistically different to zero. This finding is confirmed across three types of gasoline products, and across two samples concerning Shell's entry and all retailers' entry, respectively. Turning to the other covariates, population density and average wage per employee tend to play important roles in determining gasoline price dispersion, with consistently signed coefficients across all specifications that are significant for Gas#93 and Gas#97. Of the remaining explanatory variables, population and highway passenger traffic numbers are demand shifters, exerting negative impacts on price dispersion of Gas#90. GDP per capita tends to play a mixed role and could be regarded as both a cost and a demand shifter. In our case, GDP per capita is significant and positive determinant of price dispersion for Gas#90 but negative determinant for Gas#93 in the sample regarding Shell's entry.

[Table 7 about here]

Due to the possibility of a delayed effect of Shell's entry our extended framework explores the lagged impact of foreign entry on price dispersion. Table 8 presents our difference-in-difference estimation results based on equation (5). The results in Columns (1), (2) and (3) reveal a positive and then negative impact of Shell's entry on the degree of price dispersion for highly refined products as indicated by the $Treatment \times Post_0$ and $Treatment \times Post_1$ variables. The degree of price dispersion between Shell populated cities and non-Shell populated cities increases by approximately 1.3 to 1.5 percentage points before falling by 1.5 to 1.8 percentage points in the following period. The finding is confirmed by the results using the sample concerning the entry of top four foreign retailers in Columns (4), (5) and (6). The degree of price dispersion between a city populated by a foreign retailer and a city without a foreign retailer increases by approximately 1.15 percentage points on average, before falling by 0.53 to

[16]On one hand, using the alternative sample that considers the entry of all foreign retailers strengthens our confidence in the findings regarding the impact of Shell's entry. On the other hand, it is important to be aware of possible measurement errors in the collection of data for Total, Exxon and BP.

1.32 percentage points in the following period.

[Table 8 about here]

When we compare across gasoline types it appears that a foreign retailer entry into a city tends to have a larger impact on the price dispersion of highly refined gasoline rather than the more regular gasoline products. The coefficients on the interaction term for Gas#93 and Gas#97 are similar in both the range and the statistical significance, while the elasticity for Gas#90 tends to be lower and insignificant except for some lagged periods after entry. There are two possible explanations. First, from the producer perspective, highly refined gasoline tends to be more profitable (had previously higher margins). The market for high performance fuels is also likely to be more attractive to foreign firms where they may also have a technological advantage coupled with higher advertising and marketing budgets. For the 862 Shell service stations over the whole of China, more than 95 percent of the stations sell highly refined gasoline. In contrast, only a small number of Shell service stations in Beijing, Shandong and Guangdong sell the regular gasoline product. Furthermore, unbalanced entry implies that the two large government-owned vertically-integrated oil companies, Sinopec and PetroChina, have greater monopoly power in the production and retail of regular gasoline. With a large market share occupied by incumbents, the average price level is less likely to respond to or be influenced by the entry of international retailers such as Shell and such an effect is confirmed by our results.

Second, from the consumer perspective, the cross-elasticities of demand from consumers tend be different between Gas#90, Gas#93 and Gas#97. According to a survey by Sohu, Gas#90 is used mainly by commercial vehicles such as taxis and buses. These operators prefer low grade gasoline in order to reduce operating costs, while private car owners tend to pay more attention to the condition of their vehicle and are more likely to use higher grade gasoline. Given private car owners are also more likely to have a preference for quality and non-fuel services, the entry of Shell is more likely to increase competition in the Gas#93 and Gas#97 markets and should therefore have a larger impact on local market prices. In addition, commercial vehicles are often tied into long-term contracts with incumbents and are thus less likely to change brand in response to price competition. As a result, the market price distribution for Gas#90 is less likely to be influenced by the entry of a foreign competitor.

To check the robustness of results, we implement a placebo-control test based on the potential control cities. Within the group of non-Shell cities, a random sample of cities with random entry years were selected as the treatment group (we chose 6 to 10 cities as the treated cities with entry years between 2002 and 2011) and implemented our pairwise DID estimator to investigate the first, second and third period entry effect. The process was simulated 300 times and no significant entry effects were found for any period after the assumed year of entry. This gives us confidence that we are capturing the impact of Shell's entry on prices. In the next section we investigate the potential mechanism through which foreign entry has an impact on gasoline market in China.

6 Pricing mechanism

New foreign gasoline retailers are assumed to exert competitive pressure in localized gasoline markets (Hastings, 2004; Zimmerman, 2012; Lin and Xie, 2015). From the Edgeworth price cycle theory, the average price level tends to decrease due to undercutting initiated by new entrants in an attempt to steal market share. When margins get too low and the market shares have settled down, the average price is thought to rebound. In this section we investigate whether our price dispersion results are driven by the competition from foreign entry conditional on the spatial variation in market maturity. To do this we run regressions using the level of retail prices identifying whether the treated city is located in the western region for each gasoline type using the standard and the extended framework with lags on the entry decision.[17] The results are presented in Table 9.

[Table 9 about here]

Panels A and B in Table 9 present the results investigating the impact of foreign entry on the level of gasoline prices for the whole sample. Panel A shows the static or aggregated impact of foreign entry, while panel B shows the dynamic impact by including the lagged difference-in-difference interaction terms. Furthermore, Panels C and D account for the regional impacts of foreign entry by including triple interaction terms constructed from the difference-in-difference term and the regional dummy ($Treatment \times Post0xWest$ in panel C and $Treatment \times Post0xWest$ and its two lags in panel D). By doing so we avoid undesired influences of changing sample sizes on the estimates when inferring results across panels. Similar to before, the first three columns in each panel display results using the sample for Shell's entry into a city, while the final three columns is a sample that includes the entry of all foreign retailers. A covariate matrix including city-level controls, year fixed effect, and individual fixed effect are included in all specifications. Note that the unit of analysis is an individual city rather

[17] The treated cities located in the western region are Chongqing, Chengdu, Zigong, Xian and Baoji. Beijing, Guangzhou and Shenzhen are regarded as eastern cities and Taiyuan is considered as a central city.

than a comparison across cities which implies that the left hand side variable is simply the price level rather than a measure of price dispersion.

Results in panel A suggest that Shell's entry had no significant impact on retail prices in aggregate. This finding is consistent across the alternative sample of all foreign retailers and matches the results in Table 7. When we take a closer look at the dynamic impacts of foreign entry in panel B, the instantaneous effect of increased competition appears to be negative, although only one of six estimates of $Treatment \times Post0$ is statistically significant at the 95% level or above. A positive effect following the reduction in prices occurs from the first period after Shell's entry ($Treatment \times Post1$ in Columns (2) and (3) of panel B) while it starts from the second period after the entry of any foreign retailer ($Treatment \times Post2$ in Column (5) of panel B).

In panel C a triple interaction term is included to account for the heterogeneous impact of foreign entry on retail prices across regions. An immediate observation is the negative and statistically significant coefficient on the triple interaction term for Gas#93 and Gas#97. This reveals that Shell's presence in a western market increases competition and leads to a reduction in the retail price for two of our three gasoline products. The finding is confirmed by using the alternative sample in Columns (4), (5) and (6). Comparing the results in panel C with those in panel A, it is evident that the impact of the competition is mainly driven by a strong entry effect in western regions in China and is again more significant for highly refined gasoline. Finally, we investigate the dynamic impact of foreign entry on price levels conditional on the spatial variation of market maturity and the results are shown in panel D. It appears again that foreign entry had a greater and more immediate impact on retail prices in western regions. Moreover, prices tend to bounce back in the next two years especially in eastern area as indicated by the positive and statistically significant coefficients on $Treatment \times Post2$ in Columns (1), (5) and (6).

Collectively, the results in Table 9 indicate that foreign entry plays a role in driving local gasoline prices but is not the sole determining factor. The limited impact of foreign entry could be due to the dominant market share of government-owned oil companies as well as periodic pricing adjustment scheme. Nevertheless, results in panels B and D signal that the initial positive effect of entry on price dispersion in Table 8 is likely to be driven by price falls in treated cities and that this effect is particularly strong in western regions and most strongly for highly refined products. Moreover, the price rebound in the periods after entry tend to capture the decrease in price dispersion in Table 8. One explanation for these overall price dynamics comes from the Edgeworth price cycle hypothesis that was discussed in the literature review where an asymmetric pattern of price changes is comprised of numerous small prices decreases that are then interrupted by occasionally large price increases (Maskin and Tirole, 1988; Noel, 2011). As western energy market is less mature, the reallocation of market share is likely to be stronger in the western region (Ma and Oxley, 2012).

7 Conclusions

The recent fall in crude oil prices provides China with an opportunity to implement further energy market reforms. Promoting a more market-orientated pricing regime through a diversified ownership structure is an integral part of such a process. After the opening up of the retail and wholesale market many of the world's largest multinational oil companies sort access to China's energy industry and in many cases they expanded rapidly. In this paper we examine Shell's entry into China and investigate the entry effect on gasoline price convergence across China for three different types of gasoline.

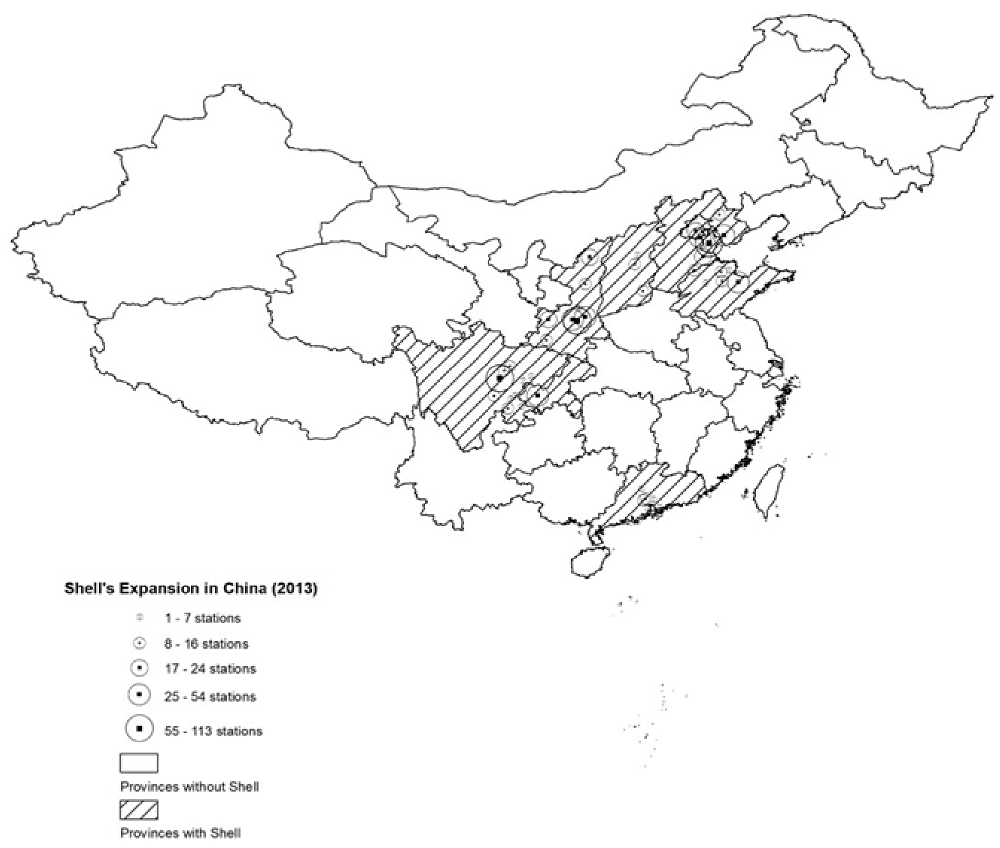
Combining a pairwise approach and the difference-in-difference model we show that the absolute price differential between cities where Shell has a presence and a similar city without any Shell stations increased by between approximately 1.3 and 1.5 percent compared to the price differential between any two non-Shell targeted cities. The increased price gap was found to close again in the following two years. Highly refined gasoline products tend to be the most affected. The price effect is largest for Western China where the competitive effect of a large new entrant is felt most keenly. Then we investigate the pricing mechanism for the retail price of premium products (Gas#93 and Gas#97) we found prices fell in the entry year in western regions before increasing in the following two years. The main results are robust controlling for the entry of other foreign retailers and to a placebo-control test. The difference in the magnitude of the impact across our three gasoline products holds.

The deregulation and market liberalization of China's energy market has been in progress for more than ten years. The outcome appears to be positive for consumers who have experienced lower prices for gasoline. With China's development entering a period of lower growth, energy consumption has slowed dramatically. During the first ten years of 21st century, the annual average growth rate for energy consumption was 9.4%. Between 2011 and 2018 this rate of growth has fallen to 3.2%. As a result, China is likely to have sufficient energy supplies to meet this lower than expected demand. Politically, the current period represents an ideal time to pursue further reform of the energy market. One example of the progress being made is the, so-called, "mixed-ownership" reforms when in 2014 Sinopec sold off around 30% of its retailing business to private investors raising more than 100 billion yuan. Efforts to encourage further reform of the refined oil pricing mechanism could also be considered (Lin and Liu, 2013).

One challenge in the present study is the lack of detailed station network information for China. At present we only know the network distribution of Shell but do not have information of incumbent competitors in a city. If the yearly station counts of Shell as well as other branded retailers (such as CNPC and Sinopec) are available, we could have investigated the relationship between station density (the number of stations in a city), gasoline prices and price dispersion and then explored the exit and entry of domestic retailers. Overall, we consider the price effect of foreign entry to be small economically given the dominant role of state-owned oil companies and the characteristics of the retail gasoline market in China (e.g. the periodic nationwide price adjustment). We hope that our research will provide impetus for future research into the impact of the open market policy on prices in China. In the light of ongoing trade war with the US, the impact of foreign entry on the domestic market may be less than envisaged.

Figures and tables

Figure 1: The location of Shell service stations in China (2013)



Source: Shell website http://www.shell.com.cn/en/aboutshell/our-business-tpkg/china.html

Figure 2: Number of Shell service stations per million people (2013)

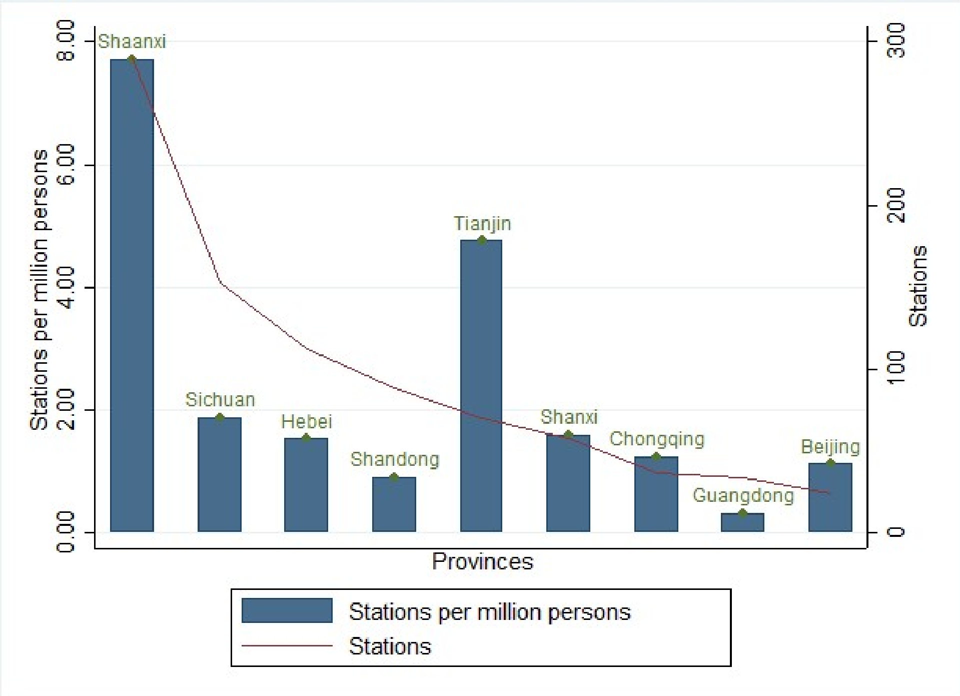


Figure 3: Number of Shell service stations per million vehicles (2013)

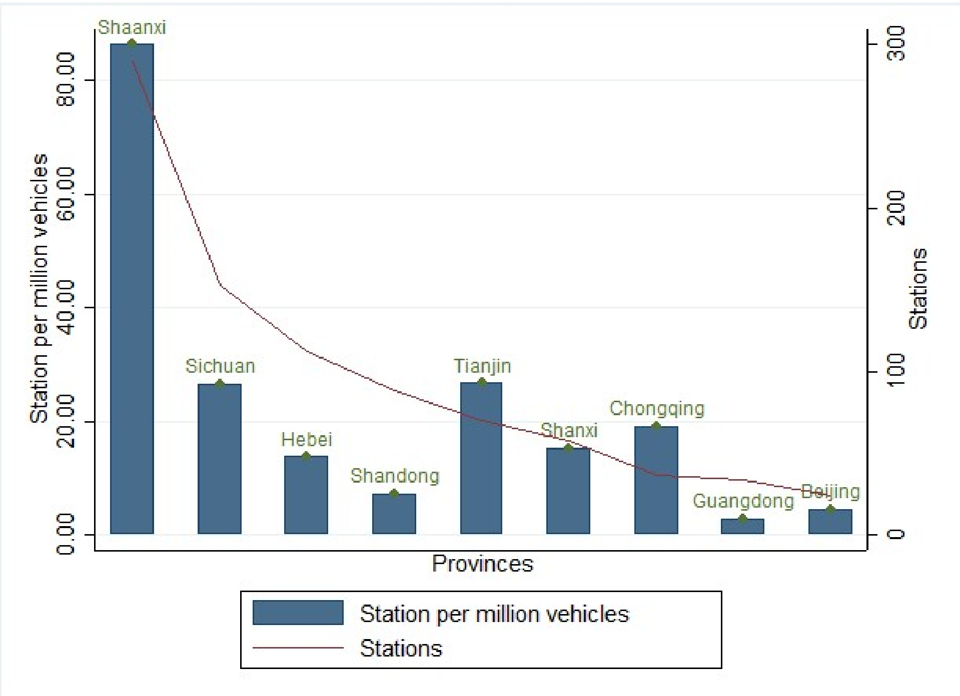
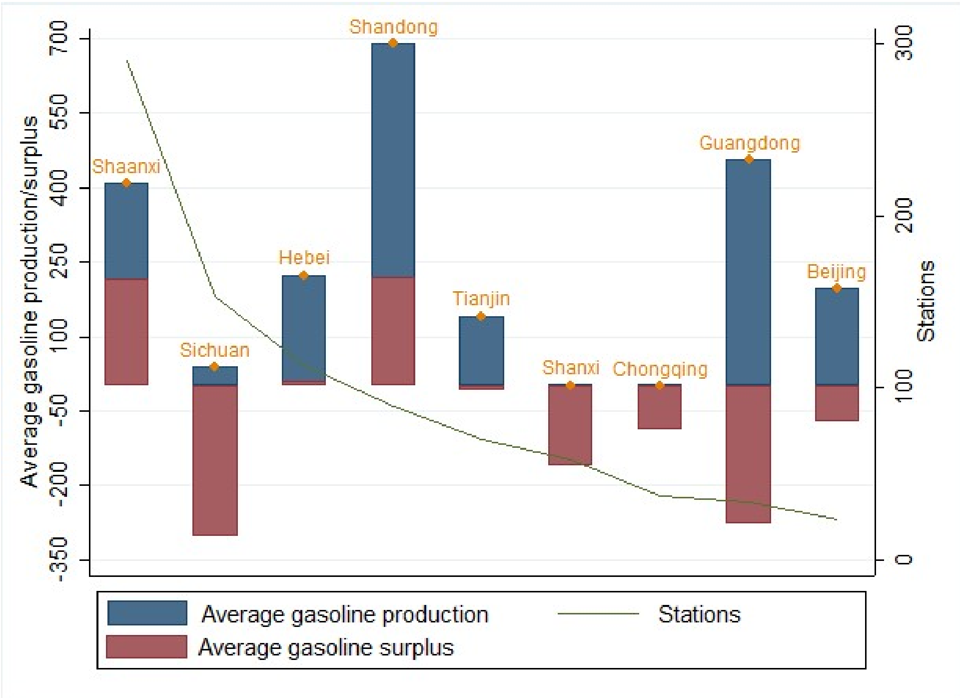


Figure 4: Average gasoline production and surplus at the provincial level (2001-2012)



Source: China Energy Statistical Yearbooks. Unit in 10,000 tons.

Table 1: A summary of Shell service station network across 9 provinces (2013)

Province	Stations	Population	Vehicles	Stations per million persons	Stations per million vehicles	Average annual gasoline production (10,000 tons)	Average annual gasoline consumption (10,000 tons)	Average gasoline surplus/deficit (10,000 tons)
Shaanxi	290	3764.00	336.08	7.70	86.29	426.10	197.89	228.21
Sichuan	153	8107.00	573.03	1.89	26.70	39.59	355.17	-315.58
Hebei	113	7332.61	816.29	1.54	13.84	227.12	219.74	7.38
Shandong	89	9733.39	1199.71	0.91	7.42	722.32	492.47	229.85
Tianjin	70	1472.21	261.58	4.75	26.76	141.88	153.17	-11.28
Shanxi	58	3629.80	378.27	1.60	15.33	2.01	168.87	-166.86
Chongqing	37	2970.00	192.77	1.25	19.19	0.42	91.78	-91.36
Guangdong	34	10644.00	1177.37	0.32	2.89	465.21	767.04	-301.83
Beijing	24	2114.80	517.11	1.13	4.64	200.49	281.18	-80.69

Note: Data sources from China Data Online and China Energy Statistical Yearbook. Units for population and vehicles are 10,000 persons and 10,000 units respectively. Average annual gasoline production, consumption and deficits are calculated for the period 2001 to 2012.

Table 2: Number of Shell service stations in our nine treated cities (2001-2012)

City	Province	Stations	First entry date
Beijing	Beijing	24	2003
Taiyuan	Shanxi	13	2010
Guangzhou	Guangdong	9	2007
Shenzhen	Guangdong	1	2007
Chongqing	Chongqing	37	2007
Chengdu	Sichuan	68	2005
Zigong	Sichuan	4	2005
Xian	Shaanxi	113	2009
Baoji	Shaanxi	20	2010

Note: Data sources from Shell China http://www.shell.com.cn/en/aboutshell/our-business-tpkg/china.html.

Table 3: Treated and control groups (2001-2012)

Group	Treated cities	Control cities	Group	Treated cities	Control cities
1	Beijing	Datong	6	Chengdu	Fuzhou
1	Beijing	Hefei	6	Chengdu	Haikou
1	Beijing	Nanchang	6	Chengdu	Hefei
1	Beijing	Pingliang	6	Chengdu	Nanchang
1	Beijing	Wuhan	6	Chengdu	Wuhan
1	Beijing	Xiamen	6	Chengdu	Zhengzhou
2	Taiyuan	Datong	7	Zigong	Dalian
2	Taiyuan	Jiaxing	7	Zigong	Haikou
2	Taiyuan	Pingliang	7	Zigong	Nanning
2	Taiyuan	Wuhan	7	Zigong	Qingdao
2	Taiyuan	Xingtai	7	Zigong	Quanzhou
2	Taiyuan	Xvzhou	7	Zigong	Shenyang
3	Guangzhou	Nanjing	7	Zigong	Zhengzhou
3	Guangzhou	Xiamen	8	Xian	Nanjing
3	Guangzhou	Xingtai	8	Xian	Wuhan
4	Shenzhen	Jiaxing	8	Xian	Xiamen
4	Shenzhen	Xingtai	8	Xian	Xingtai
4	Shenzhen	Xvzhou	8	Xian	Xvzhou
4	Shenzhen	Zhoukou	9	Baoji	Dalian
5	Chongqing	Nanjing	9	Baoji	Fushun
5	Chongqing	Shanghai	9	Baoji	Guiyang
			9	Baoji	Shijiazhuang
			9	Baoji	Yantai

Note: Only cities which have never been entered by Shell are included as possible control cities. The PSM procedure is based on nearest neighbor matching. A robustness check using radius matching gives similar results which are available upon request.

Table 4: Post-matching tests for balanced assumption

Variable	Matching	Mean Treated	Mean Control	%Bias	%Bias Reduction	T test	P-value	V(T) /V(C)
GDP per capital	Unmatched	10.35	10.15	40.00		2.21	0.03	0.81
	Matched	10.35	10.29	11.60	70.90	0.46	0.65	0.64
Population density	Unmatched	6.46	5.98	90.30		4.72	0.00	0.59
	Matched	6.46	6.42	7.60	91.60	0.35	0.73	0.76
Average wage	Unmatched	10.23	10.05	58.60		3.22	0.00	0.79
	Matched	10.23	10.21	3.80	93.50	0.18	0.86	1.24
Passenger traffic by highway	Unmatched	9.16	9.00	17.30		1.16	0.25	1.91
	Matched	9.16	9.20	-4.40	74.40	-0.20	0.85	2.41
Gas#90	Unmatched	8.66	8.65	3.60		0.18	0.86	0.51
	Matched	8.66	8.68	-8.50	-139.00	-0.45	0.66	1.07

Sample	Pseudo R2	LR chi2	P> $chi2$	MeanBias	MedBias	B	R	%Var
Unmatched	0.21	54.06	0.00	41.90	40.00	132.90	0.64	0.00
Matched	0.01	0.52	0.99	7.20	7.60	16.80	1.40	20.00

Notes: The first table compares the extent of balancing for each variable. The standardized percentage bias (column 5) is the sample mean difference between the treated and non-treated sub-samples as a percentage of the square root of the average of the sample variances in both sub-samples (Rosenbaum and Rubin 1985). A lower bias percentage implies a better match with the covariates. Column 7 shows the t-test for equality of means in both samples. The high p-values in the final column indicate that there is no significant difference between the treatment and the control group. The variance ratio of treated over non-treated is displayed in the last column and it should equal to 1 if there is perfect balance. The second table calculates the overall measures of covariate balance. The likelihood-ratio test of the joint insignificance and the corresponding p-values are present in column 2 and 3. The mean and median bias as summary indicators of the distribution of the absolute bias are shown in column 4 and 5. Robin's B (the absolute standardized difference of the means of the linear index of the propensity score in the treated and (matched) non-treated group) is recommended to be less than 25 and Rubin's R (the ratio of treated to (matched) non-treated variances of the propensity score index) is recommended to be between 0.5 and 2 for the samples to be considered sufficiently balanced (Rubin, 2001).

Table 5: Description of the variable dispersion (2001-2012)

Variable	Obs.	Mean	Std. Dev.	Min	Max
Δlog (Gas#90 retail price)	1,262	0.0402	0.0491	0.0000	0.2716
Δlog (Gas#93 retail price)	1,316	0.0354	0.0373	0.0000	0.2197
Δlog (Gas#97 retail price)	1,302	0.0356	0.0361	0.0000	0.2030
Δlog (GDP per capita)	1,604	0.6893	0.5250	0.0016	2.6764
Δlog (Population density)	1,539	0.4895	0.4290	0.0009	1.8431
Δlog (Average wage)	1,422	0.2021	0.1810	0.0001	1.1389
Δlog (Population)	1,680	0.7086	0.5405	0.0017	2.1276
Δlog (Passenger traffic by highway)	1,434	0.7663	0.5586	0.0006	3.2216

Note: The dispersion rates (in absolute values) for all variables are calculated with the same order which is randomly assigned. All monetary values are deflated in 2005 prices using province level CPI.

Table 6: The impact of distance on price dispersion (OLS regression 2001-2012)

Variables	(1) Gas #90	(2) Gas #93	(3) Gas #97	(4) Gas #90	(5) Gas #93	(6) Gas #97
Trend	0.0027***	-0.0013***	-0.0015***	0.0026***	-0.0013***	-0.0016***
	(0.0004)	(0.0003)	(0.0003)	(0.0004)	(0.0003)	(0.0003)
Log(Distance)	0.0105***	0.0071***	0.0048***	0.0690***	0.0371**	0.0592***
	(0.0015)	(0.0012)	(0.0011)	(0.0191)	(0.0148)	(0.0140)
Log(Distance)^2				-0.0041***	-0.0021**	-0.0038***
				(0.0014)	(0.0011)	(0.0010)
Constant	-5.3664***	2.5557***	3.0918***	-5.3674***	2.5497***	3.0776***
	(0.7703)	(0.6271)	(0.6388)	(0.7606)	(0.6230)	(0.6300)
Observations	1,262	1,316	1,302	1,262	1,316	1,302
R-squared	0.0572	0.0396	0.0321	0.0640	0.0428	0.0430

Note: Robust standard errors in parentheses. ***, ** denote significance at the 1% and 5% levels respectively. All monetary values are deflated in 2005 prices using a province level CPI.

Table 7: The impact of foreign entry on the retail price dispersion – Standard framework (2001 - 2012)

VARIABLES	(1) Shell Dispersion (Gas 90)	(2) Shell Dispersion (Gas 93)	(3) Shell Dispersion (Gas 97)	(4) All retailers Dispersion (Gas 90)	(5) All retailers Dispersion (Gas 93)	(6) All retailers Dispersion (Gas 97)
Treatment×Post0	-0.0076	0.0005	0.0013	-0.0053	-0.0015	0.0027
	(0.0064)	(0.0058)	(0.0062)	(0.0038)	(0.0044)	(0.0043)
ΔLog(GDP per capita)	0.0273**	-0.0149**	-0.0143	0.0003	0.0031	0.0012
	(0.0126)	(0.0073)	(0.0080)	(0.0044)	(0.0037)	(0.0040)
ΔLog(Population density)	0.0027	0.0329	0.0117	0.0099	0.0206***	0.0197**
	(0.0239)	(0.0302)	(0.0288)	(0.0095)	(0.0079)	(0.0077)
ΔLog(Average wage)	-0.0334	-0.0564**	-0.0522**	-0.0085	-0.0116	-0.0114
	(0.0264)	(0.0246)	(0.0230)	(0.0089)	(0.0062)	(0.0073)
ΔLog(Population)	-0.0574**	-0.0039	-0.0042	-0.0185	-0.0236	-0.0331
	(0.0242)	(0.0280)	(0.0246)	(0.0233)	(0.0211)	(0.0211)
ΔLog(Passenger traffic by highway)	-0.0100***	-0.0025	0.0021	-0.0018	-0.0013	-0.0016
	(0.0038)	(0.0063)	(0.0059)	(0.0017)	(0.0015)	(0.0017)
Constant	0.0310	0.0409**	0.0439***	0.0166	0.0088	0.0103
	(0.0179)	(0.0197)	(0.0167)	(0.0103)	(0.0089)	(0.0088)
Year specific effect	Yes	Yes	Yes	Yes	Yes	Yes
Individual specific effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	960	1,010	997	2,928	2,740	2,734
R-squared	0.6636	0.5211	0.5158	0.5862	0.6844	0.6949

Note: The dependent variable is retail gasoline price dispersion. Columns (1)-(3) employ the matched sample considering the entry of Shell in a local market; Columns (4)-(6) employ the matched sample considering the entry of four foreign retailers, including Shell, BP, Total and Exxon. Robust standard errors clustered at individual level are reported in parentheses. ***, ** denote significance at the 1% and 5% levels respectively. Estimates for $Treatment$, $Post_0$ are not reported in the table. All monetary values are deflated in 2005 prices using a province level CPI.

Table 8: The impact of foreign entry on the retail price dispersion – Extended framework (2001 - 2012)

VARIABLES	(1) Shell Dispersion (Gas 90)	(2) Shell Dispersion (Gas 93)	(3) Shell Dispersion (Gas 97)	(4) All retailers Dispersion (Gas 90)	(5) All retailers Dispersion (Gas 93)	(6) All retailers Dispersion (Gas 97)
Treatment×Post0	0.0064	0.0125**	0.0149**	-0.0026	0.0055	0.0115**
	(0.0044)	(0.0055)	(0.0059)	(0.0037)	(0.0046)	(0.0048)
Treatment×Post1	-0.0085	-0.0153**	-0.0179***	-0.0053**	-0.0116***	-0.0132***
	(0.0082)	(0.0060)	(0.0054)	(0.0022)	(0.0029)	(0.0030)
Treatment×Post2	-0.0203***	-0.0029	-0.0029	0.0027	0.0029	0.0017
	(0.0064)	(0.0048)	(0.0047)	(0.0019)	(0.0021)	(0.0019)
ΔLog(GDP per capita)	0.0288**	-0.0127	-0.0126	-0.0000	0.0032	0.0014
	(0.0122)	(0.0075)	(0.0081)	(0.0044)	(0.0037)	(0.0040)
ΔLog(Population density)	0.0077	0.0382	0.0150	0.0100	0.0204***	0.0196**
	(0.0248)	(0.0306)	(0.0291)	(0.0094)	(0.0078)	(0.0076)
ΔLog(Average wage)	-0.0365	-0.0622**	-0.0566**	-0.0086	-0.0117	-0.0116
	(0.0263)	(0.0253)	(0.0237)	(0.0089)	(0.0061)	(0.0073)
ΔLog(Population)	-0.0574**	-0.0035	-0.0040	-0.0171	-0.0195	-0.0281
	(0.0260)	(0.0287)	(0.0255)	(0.0235)	(0.0210)	(0.0212)
ΔLog(Passenger traffic by highway)	-0.0086**	-0.0013	0.0027	-0.0017	-0.0015	-0.0018
	(0.0036)	(0.0064)	(0.0059)	(0.0017)	(0.0015)	(0.0017)
Constant	0.0322	0.0392	0.0429**	0.0167	0.0075	0.0086
	(0.0185)	(0.0201)	(0.0172)	(0.0103)	(0.0089)	(0.0089)
Year specific effect	Yes	Yes	Yes	Yes	Yes	Yes
Individual specific effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	960	1,010	997	2,928	2,740	2,734
R-squared	0.6793	0.5377	0.5267	0.5902	0.6883	0.6996

Note: The dependent variable is retail gasoline price dispersion. Columns (1)-(3) employ the matched sample considering the entry of Shell in a local market; Columns (4)-(6) employ the matched sample considering the entry of four foreign retailers, including Shell, BP, Total and Exxon. Robust standard errors clustered at individual level are reported in parentheses. ***, ** denote significance at the 1% and 5% levels respectively. Estimates for $Treatment$, $Post_0$, $Post_1$ and $Post_2$ are not reported in the table. All monetary values are deflated in 2005 prices using a province level CPI.

Table 9: The impact of foreign entry on the retail price for China (2001 - 2012)

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Shell (Gas 90)	Shell (Gas 93)	Shell (Gas 97)	All retailers (Gas 90)	All retailers (Gas 93)	All retailers (Gas 97)
Panel A: Standard framework						
Treatment×Post0	-0.0189	-0.0019	-0.0026	0.0035	-0.0130	-0.0139
	(0.0124)	(0.0160)	(0.0137)	(0.0094)	(0.0099)	(0.0085)
Observations	444	458	455	1,176	960	959
R-squared	0.9905	0.9893	0.9889	0.9952	0.9919	0.9920
Panel B: Extended framework						
Treatment×Post0	-0.0137	-0.0099	-0.0137	-0.0016	-0.0212	-0.0251**
	(0.0091)	(0.0159)	(0.0138)	(0.0074)	(0.0109)	(0.0095)
Treatment×Post1	-0.0171	0.0154**	0.0180**	0.0064	0.0059	0.0122
	(0.0194)	(0.0058)	(0.0073)	(0.0057)	(0.0082)	(0.0077)
Treatment×Post2	0.0206	-0.0067	-0.0018	0.0005	0.0084**	0.0051
	(0.0169)	(0.0099)	(0.0094)	(0.0055)	(0.0039)	(0.0037)
Observations	444	458	455	1,176	960	959
R-squared	0.9906	0.9893	0.9889	0.9952	0.9920	0.9921
Panel C: Spatial variation – Standard						
Treatment×Post0	-0.0184	0.0270	0.0221	0.0092	0.0012	-0.0013
	(0.0177)	(0.0149)	(0.0143)	(0.0094)	(0.0056)	(0.0061)
Treatment×Post0xWest	-0.0010	-0.0613***	-0.0524***	-0.0266	-0.0360**	-0.0316***
	(0.0229)	(0.0179)	(0.0168)	(0.0198)	(0.0143)	(0.0105)
Observations	444	458	455	1,176	960	959
R-squared	0.9905	0.9897	0.9892	0.9953	0.9921	0.9921
Panel D: Spatial variation – Extended						
Treatment×Post0	-0.0051	0.0161	0.0107	0.0028	-0.0072	-0.0119
	(0.0083)	(0.0135)	(0.0125)	(0.0063)	(0.0088)	(0.0080)
Treatment×Post1	-0.0441	0.0120	0.0115	0.0065	0.0034	0.0094
	(0.0293)	(0.0074)	(0.0080)	(0.0066)	(0.0100)	(0.0092)
Treatment×Post2	0.0795**	0.0272	0.0327	0.0025	0.0120***	0.0074**
	(0.0363)	(0.0290)	(0.0335)	(0.0060)	(0.0033)	(0.0036)
Treatment×Post0xWest	-0.0192	-0.0620***	-0.0580***	-0.0193	-0.0346**	-0.0342**
	(0.0173)	(0.0135)	(0.0127)	(0.0195)	(0.0168)	(0.0134)
Treatment×Post1xWest	0.0535	0.0074	0.0137	-0.0020	0.0101	0.0115
	(0.0329)	(0.0141)	(0.0160)	(0.0118)	(0.0150)	(0.0157)
Treatment×Post2xWest	-0.0865**	-0.0339	-0.0371	-0.0121	-0.0186**	-0.0118
	(0.0376)	(0.0309)	(0.0345)	(0.0085)	(0.0078)	(0.0068)
Observations	444	458	455	1,176	960	959
R-squared	0.9909	0.9898	0.9893	0.9953	0.9923	0.9923

Note: The dependent variable is retail gasoline price. Columns (1)-(3) employ the matched sample considering the entry of Shell in a local market; Columns (4)-(6) employ the matched sample considering the entry of four foreign retailers, including Shell, BP, Total and Exxon. Robust standard error clustered at individual level in parentheses. ***, ** denote significance at the 1% and 5% levels respectively. A covariate matrix including controls, year FE and individual FE are included in all specifications. Controls include city-level income per capital, population density, average wage, population at the end of year and passenger numbers travelling by highway. All monetary values are deflated in 2005 prices using a province level CPI.

References

Barron, J. M., Taylor, B. A., and Umbeck, J. R. (2004). Number of sellers, average prices, and price dispersion. *International Journal of Industrial Organization*, 22(8-9):1041–1066.

Baser, O. (2006). Too much ado about propensity score models? Comparing methods of propensity score matching. *Value in Health*, 9(6):377–385.

Basker, E. (2005). Selling a cheaper mousetrap: Wal-mart's effect on retail prices. *Journal of Urban Economics*, 58(2):203–229.

Bloch, H. and Wills-Johnson, N. (2010). The shape and frequency of Edgeworth price cycles in an Australian retail gasoline market. *SSRN:1558747*.

Borenstein, S. and Rose, N. L. (1994). Competition and price dispersion in the US airline industry. *Journal of Political Economy*, 102(4):653–683.

Byrne, D. P. (2012). Petrol price cycles. *Australian Economic Review*, 45(4):497–506.

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

Byrne, D. P., Leslie, G. W., and Ware, R. (2013). How do consumers respond to gasoline price cycles? *SSRN:1903558*.

Chen, Z. (2003). Dominant retailers and the countervailing-power hypothesis. *RAND Journal of Economics*, pages 612–625.

Dukes, A. J., Gal-Or, E., and Srinivasan, K. (2006). Channel bargaining with retailer asymmetry. *Journal of Marketing Research*, 43(1):84–97.

Eckert, A. (2002). Retail price cycles and response asymmetry. *Canadian Journal of Economics*, 35(1):52–77.

Eckert, A. (2003). Retail price cycles and the presence of small firms. *International Journal of Industrial Organization*, 21(2):151–170.

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

Eckert, A. and West, D. S. (2004). A tale of two cities: Price uniformity and price volatility in gasoline retailing. *Annals of Regional Science*, 38(1):25–46.

Fan, C. S. and Wei, X. (2006). The law of one price: Evidence from the transitional economy of China. *Review of Economics and Statistics*, 88(4):682–697.

Gerardi, K. S. and Shapiro, A. H. (2009). Does competition reduce price dispersion? New evidence from the airline industry. *Journal of Political Economy*, 117(1):1–37.

Gielens, K., Van de Gucht, L. M., Steenkamp, J.-B. E., and Dekimpe, M. G. (2008). Dancing with a giant: The effect of Wal-Mart's entry into the United Kingdom on the performance of European retailers. *Journal of Marketing Research*, 45(5):519–534.

Hang, L. and Tu, M. (2007). The impacts of energy prices on energy intensity: Evidence from China. *Energy Policy*, 35(5):2978–2988.

Hastings, J. S. (2004). Vertical relationships and competition in retail gasoline markets: Empirical evidence from contract changes in Southern California. *American Economic Review*, 94(1):317–328.

Hastings, J. S. and Gilbert, R. J. (2005). Market power, vertical integration and the wholesale price of gasoline. *Journal of Industrial Economics*, 53(4):469–492.

Hausman, J. and Leibtag, E. (2007). Consumer benefits from increased competition in shopping outlets: Measuring the effect of Wal-Mart. *Journal of Applied Econometrics*, 22(7):1157–1177.

Holmes, M. J., Otero, J., and Panagiotidis, T. (2011). Investigating regional house price convergence in the United States: Evidence from a pair-wise approach. *Economic Modelling*, 28(6):2369–2376.

Holmes, M. J., Otero, J., and Panagiotidis, T. (2013). On the dynamics of gasoline market integration in the United States: Evidence from a pair-wise approach. *Energy Economics*, 36:503–510.

Lewis, M. and Noel, M. (2011). The speed of gasoline price response in markets with and without Edgeworth cycles. *Review of Economics and Statistics*, 93(2):672–682.

Lewis, M. S. (2009). Temporary wholesale gasoline price spikes have long-lasting retail effects: The aftermath of Hurricane Rita. *Journal of Law and Economics*, 52(3):581–605.

Lin, B. and Liu, X. (2013). Reform of refined oil product pricing mechanism and energy rebound effect for passenger transportation in china. *Energy Policy*, 57:329–337.

Lin, B. and Xie, X. (2015). Energy conservation potential in china's petroleum refining industry: Evidence and policy implications. *Energy Conversion and Management*, 91:377–386.

Ma, H. and Oxley, L. (2012). The emergence and evolution of regional convergence clusters in China's energy markets. *Energy Economics*, 34(1):82–94.

Ma, H., Oxley, L., and Gibson, J. (2009). Gradual reforms and the emergence of energy market in China: Evidence from tests for convergence of energy prices. *Energy Policy*, 37(11):4834–4850.

Maskin, E. and Tirole, J. (1988). A theory of dynamic oligopoly, II: Price competition, kinked demand curves, and Edgeworth cycles. *Econometrica: Journal of the Econometric Society*, pages 571–599.

Noel, M. (2009). Do retail gasoline prices respond asymmetrically to cost shocks? The influence of Edgeworth Cycles. *RAND Journal of Economics*, 40(3):582–595.

Noel, M. D. (2007a). Edgeworth price cycles, cost-based pricing, and sticky pricing in retail gasoline markets. *Review of Economics and Statistics*, 89(2):324–334.

Noel, M. D. (2007b). Edgeworth price cycles: Evidence from the Toronto retail gasoline market. *Journal of Industrial Economics*, 55(1):69–92.

Noel, M. D. (2011). Edgeworth price cycles. *New Palgrave Dictionary of Economics. Palgrave Macmillan*.

Pesaran, M. H. (2007). A pair-wise approach to testing for output and growth convergence. *Journal of Econometrics*, 138(1):312–355.

Pesaran, M. H., Smith, R. P., Yamagata, T., and Hvozdyk, L. (2009). Pairwise tests of purchasing power parity. *Econometric Reviews*, 28(6):495–521.

Rubin, D. B. (2001). Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, 2(3):169–188.

Tan, X. and Wolak, F. (2009). Does China underprice its oil consumption? *Department of Economics, Stanford University*.

Vita, M. G. (2000). Regulatory restrictions on vertical integration and control: The competitive impact of gasoline divorcement policies. *Journal of Regulatory Economics*, 18(3):217–233.

Wang, Q., Qiu, H.-N., and Kuang, Y. (2009). Market-driven energy pricing necessary to ensure China's power supply. *Energy Policy*, 37(7):2498–2504.

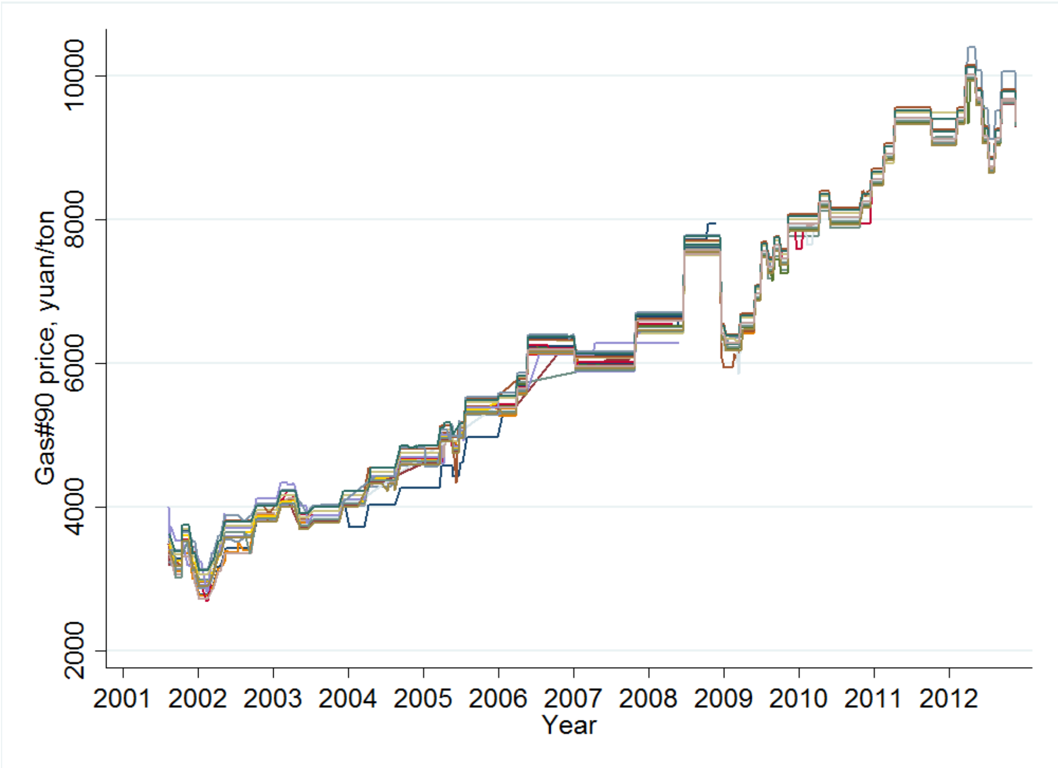
Yazgan, M. E. and Yilmazkuday, H. (2011). Price-level convergence: New evidence from US cities. *Economics Letters*, 110(2):76–78.

Yilmazkuday, D. and Yilmazkuday, H. (2016). Understanding gasoline price dispersion. *Annals of Regional Science*, 57(1):223–252.

Zimmerman, P. R. (2012). The competitive impact of hypermarket retailers on gasoline prices. *Journal of Law and Economics*, 55(1):27–41.

Appendix A

Figure A1: The sport price trend of Gas#90 in China (2011-2013)



Note: Data source CPMC. The spot price shown in this figure is from the anonymous station no.1 in each of our sampled cities. Unit yuan/ton.

Table A1: Gasoline production, consumption and number of Shell service stations at the province level (2001 - 2012)

Province	Gasoline production	Gasoline consumption	Deficit	No. of Stations
Liaoning	936.49	401.80	534.69	0
Shandong	688.31	469.09	219.21	89
Guangdong	454.93	731.21	-276.28	34
Heilongjiang	409.40	333.26	76.14	0
Shaanxi	407.18	190.63	216.55	290
Shanghai	263.64	290.61	-26.97	0
Zhejiang	262.16	409.29	-147.13	0
Gansu	255.09	78.39	176.70	0
Xinjiang	240.73	115.11	125.62	0
Hainan	239.11	38.45	200.66	0
Jiangsu	238.46	496.11	-257.65	0
Hebei	221.88	213.33	8.55	113
Beijing	196.31	267.75	-71.44	24
Hubei	185.70	400.88	-215.18	0
Jilin	165.95	125.32	40.64	0
Henan	159.91	221.12	-61.20	0
Tianjin	140.35	150.03	-9.68	70
Hunan	138.78	215.20	-76.42	0
Fujian	112.95	228.15	-115.20	0
Jiangxi	89.76	93.30	-3.54	0
Anhui	84.03	120.02	-35.99	0
Guangxi	63.80	167.45	-103.65	0
Ningxia	59.91	21.79	38.13	0
Inner Mongolia	39.47	202.50	-163.03	0
Sichuan	37.19	338.90	-301.71	153
Qinghai	29.97	20.01	9.95	0
Shanxi	2.01	162.71	-160.70	58
Chongqing	0.42	89.77	-89.35	37
Yunnan	0.02	159.16	-159.14	0
Tibet	.	.	.	0
Guizhou	.	94.61	.	0

Note: Data source from China Energy Statistical Yearbooks. Data for Tibet is not available. Values in 10,000 tons.

Table A2: Summary for variable definitions and units

Variable	Definition	Unit
Gasoline price	Gasoline price for Gas#90, Gas#93 or Gas#97	Yuan/ton
GDP per capita	Per capital of Gross Domestic Product	Yuan/person
Population density	Population divided by land area	Person per square km
Average wage	Average wage of staff and workers	Yuan
Population	Total population at the end of the year	10,000 persons
Passenger traffic by highway	Passenger traffic by highway	10,000 persons

Note: All monetary values are deflated in 2005 prices using province level CPI.

Table A3: Estimation of propensity scores with logit

Entry	Coef.	Std. Err.	z	P> |z|	95% Conf. Interval	
Log(GDP per capita)	-2.648	0.695	-3.810	0	-4.010	-1.287
Log(Population density)	2.751	0.530	5.190	0	1.713	3.789
Log(Average wage)	7.284	1.603	4.540	0	4.142	10.430
Log(Passenger traffic by highway)	-0.648	0.274	-2.360	0.018	-1.185	-0.110
Log(Gas 90)	-2.386	1.305	-1.830	0.067	-4.944	0.171
Constant	-39.840	9.830	-4.050	0	-59.100	-20.570
N	464					
Prob> χ^2	0.000					
Pseudo-R^2	0.219					

Note: All monetary values are deflated in 2005 prices using province level CPI.

Table A4: Number of foreign retail stations in 20 treated cities (2001-2012)

City	Province	Shell	BP	Total	Exxon	Stations	First entry date
Beijing	Beijing	24	0	21	3	Shell; Total; Exxon	2003(Shell)
Shijiazhuang	Hebei	0	2	5	0	BP; Total	2006(Total)
Xingtai	Hebei	0	0	3	0	Total	2006(Total)
Taiyuan	Shanxi	13	0	0	0	Shell	2010(Shell)
Shenyang	Liaoning	0	0	7	0	Total	2006(Total)
Dalian	Liaoning	0	0	12	0	Total	2007(Total)
Shanghai	Shanghai	0	0	6	0	Total	2007(Total)
Nanjing	Jiangsu	0	0	2	0	Total	2005(Total)
Xvzhou	Jiangsu	0	0	1	0	Total	2005(Total)
Hangzhou	Zhejiang	0	0	3	0	Total	2005(Total)
Ningbo	Zhejiang	0	31	12	0	BP; Total	2005(BP)
Jiaxing	Zhejiang	0	0	6	0	Total	2005(Total)
Juanghzou	Guangdong	9	112	0	0	Shell; BP	2004(BP)
Shenzhen	Guangdong	1	23	0	3	Shell; BP; Total	2004(BP)
Jiangmen	Guangdong	0	35	0	1	BP; Exxon	2004(BP)
Chongqing	Chongqing	37	0	0	0	Shell	2007(Shell)
Chengdu	Sichuan	68	0	0	1	Shell; Exxon	2005(Shell)
Zigong	Sichuan	4	0	0	0	Shell	2005(Shell)
Xian	Shaanxi	113	0	0	0	Shell	2009(Shell)
Baoji	Shaanxi	20	0	0	1	Shell; Exxon	2010(Shell)

Note: Data sources from Shell China and manual searches from the internet. The first entry date was set as the year of the first retailer's entry.