A BIG DATA VIEW OF ON-STREET PARKING

TOM DE RUIJTER

Nijmegen, June 2015

Radboud University

xerox

Tom de Ruijter: *A big data view of on-street parking*
Master thesis Computer Science
Faculty of Science, Radboud University

SUPERVISORS:
Tom Heskes
Onno Zoeter

Nijmegen, June 2015

ABSTRACT

This work is the first study to provide an analysis to the first two and a half years of the LA Express Park project for demand-based parking. The 6300 sensors installed in downtown Los Angeles have generated real-time occupancy data since May 2012, which is later coupled with payment and policy data. In collaboration with Xerox Research I have studied the relationship between rates and behaviour change of parkers, but find no effect of rates on the average occupancy or parking duration, let alone congestion and underuse. In extreme cases such as extensions of operating hours, we do find a significant reduction of average occupancy, suggesting the rate changes are too low in magnitude.

We confirm the suspicion stated in previous research that a large fraction of parkers in LA do not pay for their stay: 45% of the parked time goes unpaid on average. These high values are mostly due to drivers who abuse the handicap placards, allowing them to park free for an unlimited period. Areas suffering from congestion also suffer from above average unpaid use. On average, about half of the parking capacity is not available for paying parkers, mostly due to a minority of nonpaying parkers who arrive early and stay for the rest of the day. This may partially explain why there is no response on rate changes in high abuse areas. Removal of so-called no-parking periods leads to an influx of these long staying abusers.

Regardless of a lacking impact on occupancy, rates have developed well, with no cases where rates have reverted to their original values. I see no reason to change the rate iteration method. More incentives to walk have been created in areas that suffer from parking problems.

Lastly, this work is the first to incorporate the factors nonpayment, relative discount, relative walking distance and parking duration into a predictive model for parking choice. I test this model on historical data to show areas where incentives for behaviour change are in place. More incentives to change have been introduced since the project start, but little to no action is taken by drivers in adapting to these. This may point to either a general dislike in walking, a too simple model or lack of knowledge of cheaper alternatives among the drivers of LA.

To improve demand-based parking in the future, it is vital that satellite navigation software embeds parking software. My colleagues and I emphasise the necessity to change the handicap placard policy; there should be no exceptions for the general public. Also, off-street parking

garages should be included in future demand-based parking studies as most of the parking capacity is concentrated there.

CONTENTS

1

INTRODUCTION

Can we improve the quality of city life with smarter parking solutions? Is it possible to reduce parking congestion and pollution and achieve a better distribution of traffic over a city? If so, what sort of incentives are necessary to achieve a shift in behaviour? With the number of cars still on the rise and more people moving towards cities, city planning and policy making becomes increasingly difficult. Finding a parking spot within the downtown madness is equally difficult. Space and time are scarce and costly resources in city life, but smart and data driven methods for parking may provide a solution.

In 2011 the state of California was given a government grant for improving its parking solutions using computer technology. Road workers drilled 6324 on-street sensors into the pavement of downtown Los Angeles for monitoring the local parking behaviour. These sensors emit real time information on the occupancy of the space directly above it. During 3 years over 3000 block face rates have been adjusted. This process of data collection and rate adjustments was done within the scope of the LA Express Park project with as primary goal to reduce parking congestion. As a result, this work is the first is the first to consider a dataset of on-street parking behaviour at the single user level over a time span of three years. With such a dataset it is possible to verify earlier models and theories on parking economics and the incentives in place for choosing a parking spot.

Many academics argue that the price for parking is not a free market price, but a public price for a public service set to achieve the public goal of effectively managing supply (Vickrey, 1954; Calthrop et al., 2000; Shoup et al., 2011; Pierce and Shoup, 2013a; Zoeter et al., 2014). Parking should be regarded as a valuable and public commodity. Economics tells us if we increase the price of a good with unlimited supply, the quantity of sold goods should decrease. Of course there are some exceptions to this rule – luxury goods such as a Rolls Royce automobile, Chanel No. 5 fragrance or a prized Bordeaux wine – and while parking spaces are not an unlimited good the intuition is this basic rule of economics should hold for selling parking spaces as well, as long as people are willing to walk a bit further to their final destination.

Zoeter et al. (2014) show through simulation in a simple model that pricing can be a way to optimise social welfare, as people with a higher value for a space are willing to pay more. Los Angeles is the second city with demand-based pricing for on-street parking to put this principle into practice, with a parallel project running in San Francisco. For each block face, parking rates have been adjusted frequently depending on local demand. From the outcome of this study, I identify key premises for faster deployment of demand-based methods and faster convergence of rates to market prices in the future.

Specific car manufacturers have already started with incorporating parking technology in their navigation software (Cunningham, 2013). Cars will reserve an available space near your destination according to your preferences and those of the passengers you carry. Less time will be spent cruising for a space, meaning there will be less congestion from parking and less exhaust fumes from this group. Based on a survey from 2005, Shoup et al. (2011) reported that the share of traffic looking for parking space in Los Angeles can be as high as 68%.

SCOPE AND RESEARCH QUESTIONS In this work I look into the efficacy of the rate changes on parking behaviour in the city of Los Angeles from the period May 2012 until March 2015. Have parking customers changed their behaviour as an effect of rate changes by dynamic pricing? What incentives are currently created for customers to change their choice of space? Are these incentives reflected in the behaviour of parkers? If effects are subtle or not measurable, what premises are missing for customers to change spaces and what can other cities learn when deploying smart parking solutions?

RESEARCH CHALLENGES There are numerous challenges that require solving. The project took place as a natural experiment without a control group. Parking habits are most likely not governed by charges alone, meaning patterns vary wildly from street to street. I also expect that periodicity and the weather play a large role. I have to deal with other types of policy changes being applied simultaneously, such as extension of operating hours, extensions of time limits and introduction or removal of no-parking periods. Sensors were not deployed regularly throughout the city and sensors also suffer from failures and noise.

NOVEL CONTRIBUTIONS This is the first work to analyse on-street parking data with high resolution spatial and temporal data on occupancy and payments on a city scale over multiple years. I can identify different groups of customers and by coupling occupancy with payment data it is possible to further discern in paying and nonpaying customers. We know when a customer arrived, how long they stayed for, whether they paid and if so how much. This study is also one of the first to use real-world data to evaluate theoretical models on parking demand and driver rationality. A 'physics of parking' is yet to be formulated and testing primitive theoretical models is the first step to uncovering premises for successful demand-based pricing.

PRACTICAL RELEVANCE To effectively tackle the issues addressed, I use techniques from large scale data processing and economics. I developed new ways to visualise large quantities of spatiotemporal parking data in a clear and concise manner. Other practical relevance of this work lies in clear and concrete recommendations for better deployment of future demand-based pricing systems and a healthy parking economy in general. Results will aid in the development of better pricing iteration schemes. I can accurately identify problem areas and policies requiring change for recommendation to cities. I will show that setting prices based on demand is not trivial for several reasons. For one, demand changes over time and is difficult to measure by humans. Second, computers are much better at coupling information from many sources. Third, automatic pricing methods are more precise in identifying patterns from time series data. These data provide city planners and local governments with insight on patterns they were unable to visualise before, improving their arguments for decision making. The hourly rate equilibrium can be approached faster with less rate changes. This in turn helps the driver. By managing demand with smart software, increasing prices in congested areas and lowering them in underused areas, availability of space becomes less of an issue. Drivers are less likely to cruise in search of a spot. Unnecessary emission of polluting gases also drops when cruising is reduced.

THESIS STRUCTURE I reflect on the analysis methods given new data in section 4. Section 5 goes into further detail and provides statistics on the abuse of handicap placards and unpaid parking as a possible explanation for lacking efficacy. In section 6 I wonder whether the observed lack of efficacy of rate changes is due to a lack of cheaper parking alternatives nearby. These sections are extended versions of earlier work accepted for oral presentation (Clinchant et al., 2015). In section 7 I combine the ideas from sections 4, 5 and 6 into a simple discrete choice model for parking space alternatives and to predict incentives of change. In the last section and remainder of this work I discuss the implications of my findings and provide several recommendations for city planners, policy makers and future deployers of demand-based parking.

2 PREVIOUS WORK

Literature on parking is expanding rapidly over the past decade with contributions from urban planning, economics and computer science. The theory goes back to Vickrey (1954), who won a nobel prize in economics for his work on road pricing and pricing of public services. In this section I have summarised key problems in previously used methodology that require dealing with in the field of demand based pricing for parking. Most importantly, there appears to be no consensus on proper methodology to prove relationships between rate changes and measurements.

Recent trends in publishing revolve around ongoing projects in the cities of San Francisco and Los Angeles. Other ongoing projects include Toulouse (Beardsley, 2010), Moscow (Berishvili and Novosti, 2012), Indianapolis[1], Barcelona (Ross, 2011) and Auckland (Chapman-Smith, 2013). Common problems that also require dealing with are poor data quality, data exploration and visualisation, finding a proper metric for mapping effect of policy changes and dealing with background trends and variance in occupancy.

2.1 SFPARK

With the financial support of the United States Department of Transportation (USDOT), San Francisco and after it Los Angeles have replaced coin meters with smart and electronic parking meters and installed street sensors with the ultimate goal to improve the parking situation and reducing the amount of cruising for parking. The SFpark project is managed by the San Francisco Municipal Transportation Agency (SFMTA), who supply processed occupancy averages to interested parties. In total the project area includes 398 on-street blocks divided over 12 neighbourhoods, of which 4 were selected as a fixed-rate control group.

Choosing a representative spatially distributed control group is nearly impossible. Keeping the rate at one block fixed while increasing the rates in the surrounding area will likely cause an inflow of parkers from the treatment group towards the control block. Similarly comparing entirely different areas with each other is difficult since drivers

1 http://parkindy.net

in each area display different behaviour. A remaining option would have been to randomly apply rate changes and compare these with the non-random rate changes. The SFMTA does not provide a report on how the control group was selected.

TARGET OCCUPANCY The city of San Francisco set the monthly average occupancy target between 60% to 80%. This is a heuristic that is commonly used by urban planners and city officials. Millard-Ball et al. (2013a) were the first to look into the city target heuristic with real-world data, specifically from SFpark. Millard-Ball et al. show that in San Francisco, below 95% hourly average occupancy little to no cruising should occur according to their metric. Above 95% the distance cruised increases dramatically, which can be explained from the perspective of a critical mass for parking occupancy being reached; the system does not contain enough free spaces to maintain car flow and breaks down. The 70-90% target in Los Angeles would then be a good though conservative target to reduce cruising.

RATE ITERATION Rate iteration in San Francisco is done through reactive adjustments based on the average occupancy per month. If the monthly average occupancy lies between 60% and 80% no action is taken. Otherwise rates are adjusted up or down by a small amount ($0.25 or $0.50). Since the pricing iteration scheme is reactive, there is a risk of rate oscillation due to background variance. The SFMTA reports that in 9% of cases, meter prices were adjusted both upward and downward at least once. My colleagues and I do not know of any deeper analysis on the development of the rates and their convergence.

Three independent studies were conducted on the efficacy of the initial period of SFpark. While the methods used differ, authors agree there is little to no observed effect measured in average occupancy and a longer measurement time span is needed. For completeness the studies and their methodologies are reviewed in the remainder of this section.

2.1.1 *Price elasticity*

The first study on the efficacy of SFpark was done by Pierce and Shoup (2013a), who have calculated price elasticity of parking demand after rate changes. Price elasticity is defined as the relative change in demand divided by the relative change in price, assuming an unlimited supply. Pierce and Shoup collaborated with the SFMTA on the deployment of the SFpark project. Both parties wondered whether rate changes influences occupancy as wanted, i.e. increasing rates meant a decrease of occupancy. Elasticities vary greatly between area, time of day, initial price and price change, demonstrating the variety of natural patterns. The found price elasticities by Pierce and Shoup were often positive also indicating the rates in effect were not the most important drive of the variation in occupancy. To clarify, a negative price elasticity indicates a decrease in goods after an increase in price. Of particular interest to us are possible causes for this phenomenon. These may indicate missing premises or conditions for successful function of demand based pricing.

Millard-Ball et al. (2013b) have provided a methodological comment on this first analysis. They show by replication of the elasticity results that the claimed effects by Pierce and Shoup (2013a) are mostly due to the reactive rate adjustment and the fact this procedure depends on average occupancy; a phenomenon called regression towards the mean in statistics.

Occupancy may be temporarily high due to natural variance, but the reactive rate iteration responds by increasing the price at that block. During the next month rate changes are most likely closer to the true mean occupancy, suggesting an effect by the rate change. A strong influence of variance in occupancy unrelated to rates would explain positive elasticities.

Of course there may be some cases that are truly influenced by a change of rates as prices are not independent of occupancy. Pierce and Shoup (2013b) respond with a specific example that suggests a shift of parking traffic from one street to a neighbouring street due to occupancy. It supports the hypothesis that average occupancy is insufficient to clearly measure effects of rate changes.

A comment not made by either group of authors is that in the calculation of price elasticity lies the assumption a product cannot sell out. If the product, in this case parking spaces, were to sell out, blocks would seem inelastic while in reality demand is sky high. In reality every congested block suffers from frequent sell-outs.

2.1.2 *Metrics for cruising and space availability*

Average occupancy has been and is the most popular metric for cities to quantise parking. It summarises parking on a group of spaces to a single number that explains the average parking situation during a given period of time. The point with on-street occupancy is that it is so dynamic and may change rapidly for different moments of day. Average occupancy is only a proxy of other quantities cities want to reduce: cruising and full blocks. Its fitness for describing these aspects of parking has not been widely researched.

In their SFpark efficacy study, Millard-Ball et al. (2013a) were the first to provide a thorough analysis of the effects of using average occupancy. They developed two metrics to estimate cruising and the probability of finding a full block based on queueing theory. An overlooked aspect of averaging in parking is the impact of the duration of averaging. The longer the averaging period, the more opportunities come by of a block becoming full, meaning a higher probability of finding a full block and cruising. This is illustrated in figure 1, showing the number of blocks cruised and the probability of finding a full block for averaging over an hour and over a month.

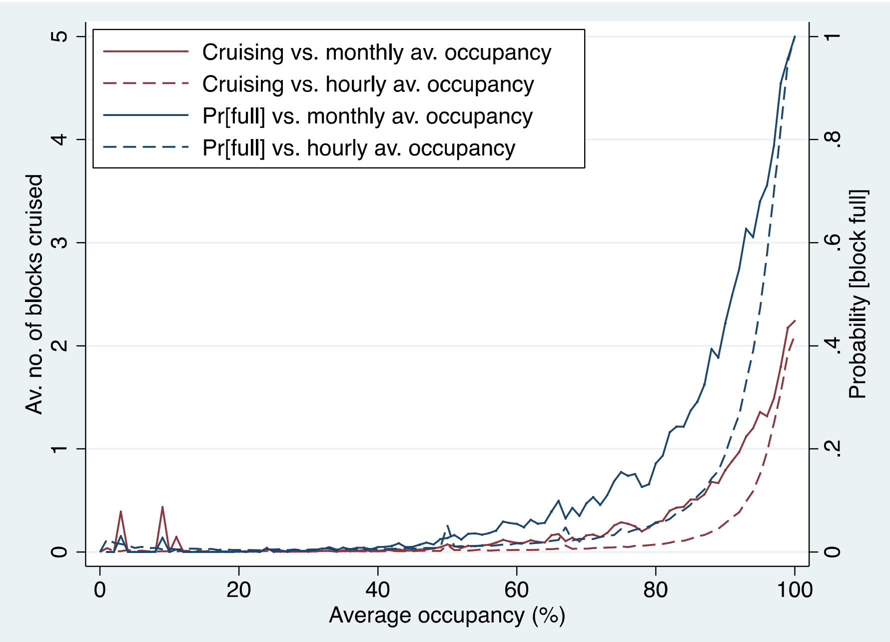


Figure 1: A comparison of cruising and the probability of finding a full block using different averaging durations. Values are estimated on the first year of San Francisco data. Image taken from Millard-Ball et al. (2013a).

An aspect used in the development of these metrics is the convex relationship between the probability of finding a full block and the length of the time span used for averaging. Imagine we measure the occupancy on some block face. Measuring 85% occupancy over one-hour will mean there are relatively few possibilities of the block becoming full. If we were to measure 85% over the course of a month, it seems unlikely the block would have been exactly 85% occupied for the entire month. Instead there were probably moments when the block was completely full as well.

The solid lines in figure 1 shows the probability of a block being full over the course of a month, while the dashed lines show the same probability but averaged per hour. The dashed line lies below the solid line, exemplifying the convex relation between the period of averaging and the probability of a block being full; the longer the period of averaging, the higher the probability the block was full.

Lastly, averaging over specific times of day will give different results as more people will be trying to park at high demand periods, such as morning or evening rush hours or even lunch time.

Even with these metrics Millard-Ball et al. (2013a), Millard-Ball et al. observe little to no change in average occupancy, cruising and parking availability compared to the SFpark control group. No surveys were done to compare the metric predictions with actual cruising.

2.1.3 *Regression discontinuity analysis*

In their comment, Millard-Ball et al. (2013b) performed a regression discontinuity analysis, while also including the second year of occupancy data from SFpark. While this method does not suffer from regression towards the mean, sadly no concrete relationship between average occupancy and rate changes is apparent. Regression discontinuity analysis assumes block faces just below and above an occupancy threshold are similar in all ways, except pricing. The findings by Millard-Ball et al. (2013b) also suggest rates are currently not the most important decision factor for parking.

2.1.4 *Linear Model study*

A third group of researchers, Chatman and Manville (2014) have independently surveyed 40 randomly selected blocks in the SFpark project area during three measurement moments in the first year. This allowed them to compute more direct metrics to measure driver satisfaction, features that were not used in earlier work. They collected statistics on employment, weather transitions, double parking incidents per hour, the number of occupants per vehicle, the share of minutes parked illegally unpaid and the share of minutes parked with handicap placards. The model specification reported by Chatman and Manville (2014) is:

$$\begin{aligned}\Delta y_{jk} = \alpha + \beta_1 P_j + \beta_2 \Delta P_{jk} + \beta_3 C + \beta_4 \,\mathrm{Fin} + \beta_5 \,\mathrm{Mis} + \beta_6 \,\mathrm{SOMA} \\ + \beta_7 \,\mathrm{Emp} + \beta_8 \Delta S_{jk} + \epsilon i,\end{aligned} \tag{1}$$

where P_j is the beginning meter price; ΔP_{jk} is the change in price from round j to round k; C indicates whether the block was in the control group; Fin, Mis and SOMA compensate for fixed environmental effects for the SFpark neighbourhoods Financial District, Mission Street and South of Market, with Civic Center as reference category; Emp is the total employment in the proximate block from the US Census's survey of Longitudinal Employer-Household Dynamics; lastly, ΔS_{jk} is the change in the share of hours with sunny and fair weather conditions as measured by the surveyors.

The variables mentioned but not included in this formula were included in other regression models which were not reported explicitly. Later, Chatman and Manville (2014) discerned price change in both increase and decrease of rates, as well as logarithmic variations of their model.

While the authors find a slight correlation between average occupancy and rate change, they find no relationship between parking duration, vehicle turnover and carpooling. They offer three possible explanations for their results. First, the city heuristic of optimal monthly occupancy between 60-80% does not mean enough spaces are available during peak hours. This contests the findings by Millard-Ball et al. (2013a) discussed in section 2.1.2, who concluded that a range of 60-80% was sufficient to stimulate space availability.

Second, for political reasons SFpark did not let parking prices float and restricted how fast and high prices could rise. As a consequence this could interfere with the observations and can possibly explain the presence and lack of the found correlations.

Third, the city updated the coin meters with electronic parking meters and also reduced or removed time limits in many city areas. This may have diluted the effect of rising prices.

This study regards blocks as independent and assumes the control group is indeed independent in every way of the treatment group. While this is partially inherent to linear regression models, Chatman and Manville (2014) could have included variables on the relative walking distances between blocks and relative discount between blocks. The authors did not explicitly compensate for the statistical effect of regression towards the mean in their reported correlations. While some regression coefficients were reported to be significant, an analysis on the model fit to the data would have been insightful.

2.2 LA EXPRESS PARK

LA Express Park is a parallel initiative to improve on-street parking in downtown cities in California. Most of the work on San Francisco is also applicable in Los Angeles. The project includes over 6000 spaces on more than 800 block faces. LA Express Park has been ongoing since early 2012 and has been under control of Xerox and the Los Angeles Department of Transportation (LADOT). Its goal is very similar to that of SFpark: reducing cruising and improving driver satisfaction.

The rate change iteration as described by Zoeter et al. (2014) is shown in algorithm 1.

Algorithm 1 The LA Express rate change iteration

1: **for** each block b **do**
- Compute the *congestion index* $I_c^{(b)}$ as the fraction of operating hours in the review period that b is congested (occupancy $> 90\%$).
- Similarly compute the *underuse index* $I_u^{(b)}$ (occupancy $< 70\%$).
- Define the *congestion-underuse balance* as
$$B_{cu}^{(b)} = I_c^{(b)} - I_u^{(b)}.$$

2: **for** each block b with $B_{cu}^{(b)} > 1/3$ (congestion dominant problem) **do** Increase the rate by one step in the ladder.

3: **for** each block b with $B_{cu}^{(b)} < -1/3$ (congestion dominant problem) **do** Decrease the rate by one step in the ladder.

By updating rates based on the congestion-underuse balance with thresholds $-1/3$ and $1/3$, the algorithm has some robustness against rate oscillations. For a detailed explanation I refer to the method section of Zoeter et al. (2014).

The analysis by Zoeter et al. (2014) also revealed that a simple partition into three parts for weekdays is optimal for over 90% of all metered spaces: open to 11AM, 11AM to 4PM, and 4PM to close. Time-of-day pricing was gradually introduced in different areas of the city, starting in August 2012.

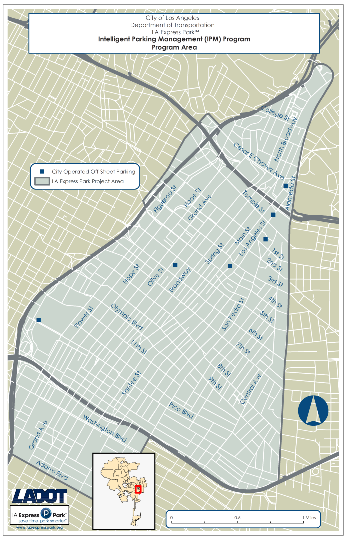


Figure 2: The project area of LA Express Park in downtown Los Angeles. Image courtesy by the LADOT.

LADOT uses the heuristic of having monthly average occupancy between 70% and 90%. These boundaries are also in line with the findings by Millard-Ball et al. (2013a) to reduce cruising. Anything above this boundary is considered to be congestion and below to be underuse. If congestion is the foremost problem, rates are increased a step on the price ladder, with a maximum of $\$6/h$. Otherwise if underuse is the foremost problem rates are reduced a step, with a minimum of $\$0.50/h$. Ghent et al. (2014) clearly describe the LA Express project goals and premises.

Since the LA Express data is still private at time of writing, only few studies have been performed, all by researchers from Xerox. I highlight important findings from an observational study.

OBSERVATIONAL STUDY From observing parkers, Glasnapp et al. (2014) have come to a number of interesting conclusions. Most importantly they confirmed the abuse of handicap parking as an important pitfall, reporting numbers as high as 75% of parkers not having to pay for their stay. "We observed many people who parked with handicap placards easily walking presumably to their place of employment nearby." This is due to lean laws in California regarding parking for the disabled. The necessity of a demand-based distribution of parking *without* such exceptions was also emphasised by Vickrey (1954).

Glasnapp et al. (2014) found that on streets with parking restrictions there was a lower amount of unpaid parking. I look into this in section 5, to see whether introduction or removal of restrictions is paired with change in unpaid use.

Second, the study confirmed many parkers are unaware of rates in the surrounding area, another key problem potentially blocking the success of the project.

Third, Glasnapp et al. (2014) found that parkers find proximity the most important value, over cost and time. Moreover, they found that short-stay parkers would be willing to walk about 3 blocks on average for cheaper parking. I use these findings for the modelling assumptions in section 7.

2.3 REGRESSION TOWARDS THE MEAN

I stress that price elasticity and other relative measures are very sensitive to small changes that could be driven by natural variance and sensor noise. Of special importance is the statistical law of Regression Towards the Mean (RTM). If not careful, an experimenter may wrongly attribute effects to a false or non-existent cause.

For example, imagine a large class room of students about to take a multiple choice test. No student bothered to study, meaning the outcome of the test is nearly fully determined by chance. Approximately half of the students passed. The teacher is a merciful person and decides to give a re-test to the 10 worst performing students the day after. Again, no one bothered to study. All 10 students perform significantly better than their first attempt, no thanks to their skill. The teacher is happy and concludes his retesting had a positive effect on the outcome. This is RTM in practice.

Simply put, any signal depending on a specific probability distribution in some way displays RTM. The occurrence of an extreme value with small probability mass is likely to be followed by a value closer to the distribution mean. The effect is especially applicable when deciding whether or not to apply some treatment depending on the measured value. In the case of the above example, the teacher decided to give a retest to the 10 worst scorers. Of course the improved scoring had nothing to do with the test, but is simply an effect of the laws of probability.

In experimental design one can compensate for the effect by having a specific control group also undergoing treatment and by randomly assigning subjects to control or treatment group. The effect of RTM within measurements can then be quantified and compensated for by using the correlation between the dependent and the independent variable. RTM and ways of prevention and compensating in experimental design are clearly described by Barnett et al. (2005).

The effect is very much present in the SFpark and LA Express studies. Here, treatment is the adjustment of rates. The adjustment of rates depends on prior occupancy and some threshold. The effect of treatment is also measured in occupancy. A block displaying high occupancy receives an increase in rates, which may be followed by lower occupancy. The actual influence of the rate change is mixed with the influence from RTM. Millard-Ball et al. (2013b) were the first to notice this problem in the context of demand-based pricing. Sadly, LA Express does not have a control group nor random application of rate changes. This makes it difficult to split small influences from rate changes from the effect of RTM. Of course, if there is no measured effect there is also no need to compensate.

2.4 conclusions

To summarise the observations from SFpark and LA Express, there is little to no observed effect on project goals with average occupancy. The different academic groups agree that this may be due to average occupancy only being a proxy for driver satisfaction and the amount of cruising. Changes might be more visible in statistics of unpaid use, turnover or arrival-duration. Further investigation may reveal a change in the composition of parking populations, although findings byChatman and Manville (2014) suggest the opposite.

Another cause may lie in the strong seasonality and high variance in on-street occupancy. Data from a longer measurement period is necessary to clarify the matter. The data for this work was obtained directly from the sensors, meaning the full preprocessing procedure was done by us, rather than the SFMTA or the LADOT. This gives us additional expressive power in the analysis of this dataset; we can monitor parking on the level of individual drivers. The processing is described in the next chapter.

All groups stress that unpaid parking and in particular abuse of handicap parking policies is a problem for parking in California. With the exception of Chatman and Manville (2014), previous work did not account for unpaid use. This work is the first to give special attention to this issue, in section 5.

Pierce and Shoup (2013a) and Millard-Ball et al. (2013a) mention that some congested blocks could be so inelastic that more drastic price changes are required before observing short-term change. This is supported by the study of Zoeter et al. (2014) and Ghent et al. (2014) where an extension of evening operating hours caused people to consider alternatives.

Furthermore, Glasnapp et al. (2014) confirmed the awareness of drivers of surrounding rates in Los Angeles to be alarmingly low, even though there exists the willingness to walk to a cheaper alternative. This may also have been a problem in San Francisco. To look into this issue, I build a choice model – see section 7 – that regards block faces in respect to each other, rather than assuming they are independent as is done in earlier work. I model the walking distances between spaces to take into account parking duration and relative discount.

3 DATA PREPROCESSING

Data aggregation and price adjustments were done with methods developed by Xerox in assignment for the LADOT. This includes data from sensors, meter payments and space policies. Upon inquiry representatives from Xerox shared snapshots of the LA Express data solely for academic purposes. All further processing was done with Matlab R2014b.

3.1 SENSOR DATA

Each time a sensor detects a change in its state, it emits a record <current location, current time, new state, sequence number>. Sensor states are either *occupied*, *vacant* or *unknown*, where unknown indicates detectable sensor failure. These occupancy messages represent parking events, which can be extracted using simple but admissible heuristics. Admissible in this sense means the occupancy variable is never overestimated and the sensor is marked as offline in cases of ambiguity.

In total there are transmissions for over 6300 parking spaces. Database records dated before May 2012 and after March 2015 are removed. Records without location identifier are considered beyond recovery and are also removed. It may happen a sensor transmits the same message more than once; only a single copy is kept. Apart from the location identifiers, I lack data on the physical sensors such as placement, replacement or repairs, making the mapping between spaces and sensors unclear. To my knowledge, sensors have not been moved between spaces. There is one space with more than one sensor active simultaneously at a specific point in time, for which I only keep transmissions of the earliest active sensors during the collision period. I found five cases where a sensor was replaced by one with a different identifier.

Although sensor time is provided in UTC, the parking operating hours are defined in local time. All transmission times are converted to Los Angeles local time. This introduces some ambiguity when the clock is set back an hour on the first Sunday of November between 2:00AM and 3:00AM for daylight saving time. It may happen that a person arrives at 2:30AM and leaves at 2:15AM in this encoding. For the analysis in the rest of this work I have removed parking events starting and ending in this period, unless stated otherwise. I removed all parking events on national holidays. This leaves over 100 million sensor transmissions between 26 May 2012 and 3 March 2015.

While occupancy tells us how many cars were physically detected at a time and location, capacity tells us how many sensors were available for reliable measurement. The order of application of the heuristics is important and changes the semantics of the total filtering procedure.

SEQUENCE NUMBER FILTERING Each sensor keeps a sequence number for the number of messages it has sent. A gap in sequence number indicates an unrecoverable loss of information and the occupancy is set to unknown during the gap period.

UNKNOWN STATE FILTERING When receiving one or more transactions with unknown state, the period of time from the first unknown message until the first vacant or occupied message is marked as unavailable for that location.

TIME FILTERING Some occupied or vacant events last too long or short to be realistic. It is likely these indicate sensor failure respectively sensor noise. Sensors whose occupied states last longer than two days and whose vacant states last longer than two weeks are marked as offline during this period of time. Sensors with occupied states lasting shorter than 7 seconds are unlikely to encode a parked car and are marked as vacant during this time. Sensors with vacant states lasting shorter than 2 seconds are unlikely to encode a leaving a car and are marked as occupied during this time.

The result of all these filtering steps is an occupancy function specific for each space that says: is the space occupied at time t? This occupancy function can also be converted to a dataset of parking events, where each event is marked by the arrival and departure time of a car. The time of arrival arrival is the time of occupation and the duration is the amount of time until the space is vacated.

Obviously smarter filtering steps can be done. For example, Zoeter et al. have demonstrated sensor noise patterns are specific to groups of sensors and it is possible to fit noise models to better predict street occupancy (Zoeter et al., 2012). In this work I continue with simple heuristics as this has proven sufficient.

3.2 METER AND PAYMENT DATA

The collected dataset on payments is a set of records $<$space identifier, paid amount, time of payment, time paid for$>$. There is a separate table specifying the relationship between spaces and meters. Payments for more than 24 hours of parking and for spaces further than 1 km of the meter were removed. In total there we included 20M payments between 26 May 2012 and 3 March 2015 with a total value of $30M.

There are two types of meters: multi space meters and single space meters. Multi space meters are typically placed on the corners of the street and allow payment for spaces in a very wide radius. Single space meters are placed directly at each space and only allow payments for that space. In rare cases single space meters have recorded payments for neighbouring spaces. After verification, these neighbour payments happen when the space's meter was taken offline for repairs.

When inspecting the distribution of payments throughout the project period, I found days lacking payments entirely. While some of these dates corresponded with national holidays, I did verify payments with data obtained directly from the single space meter vendor for the six month period of June 2012 to January 2013. I found found approximately 5% of payments to be missing throughout the day compared to the vendor standard, without a clear spatial or temporal pattern. I did no such comparison for data from the multi space vendor.

3.3 POLICY INFORMATION

Parking policies encode when people should pay, how much they should pay, how long they can stay for and whether parking is allowed at all. I call these concepts respectively operating hours, hourly rates, time limits and no-park periods. Any car parked in a no-park period will be towed, as the purpose of these periods is to use the parking lane for traffic or cleaning the street. By the end of 2012 time-of-day pricing was introduced to most spaces in downtown Los Angeles. This means different times of day may have different hourly rates to provide incentives to park outside of peak hours.

The city of Los Angeles works with a discrete rate ladder, where each space has an hourly rate from the list $[\$0.50, \$1, \$1.50, \$2, \$3, \ldots, \$7, \$8]$. A rate change may take at most one step up or down the ladder, never more expensive than $\$8/h$ or cheaper than $\$0.50/h$. Generally, all spaces on a block face have the same policy. A small amount of spaces had hand-set policies throughout the project diverting from these rules. These spaces were removed from the dataset.

3.4 SPACE DATA

Associated with each space are its geographical coordinates, the block face it lies on, sensor and payment transactions and policy information. Spaces without geographical location, sensor transactions, payments or policy informations were marked incomplete for further analysis. In total, data of 5429 spaces was used for processing.

3.5 DEALING WITH MISSING OBSERVATIONS

Like any physical equipment, parking sensors fail and suffer from noise. In fact, the amount of available sensors deteriorates rather quickly with time. See figure 3.

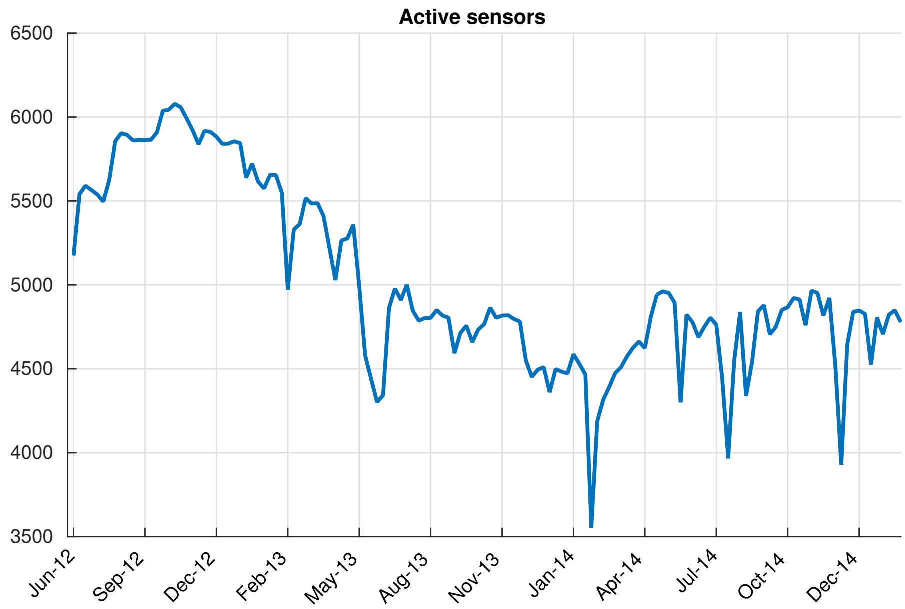


Figure 3: The number of sensors that is available on a weekly resolution. A sensor is included in the count if it was deemed online after the filtering steps. The downward spikes are large system failures, not sensor failures.

An important aspect for estimating street-level occupancy is how to deal with missing sensors. Imagine we have the measurements taken on a single street over three weeks as seen in table 1. We are interested in what fraction of the street is occupied. Given the data from the first two weeks it is likely the street remains completely sold out in the third week as well. A simple approximation is assuming sensors fail at random. We scale the observed occupancy by a quantity depending

Week	Occupancy	Sensors active (capacity)	Occupancy fraction
1	20	20	1
2	20	20	1
3	18	18	1

Table 1: Example occupancy and capacity

on the number of observed sensors, a procedure called backfilling in statistics. The estimated occupancy of a set of locations is then

$$\hat{Z} = Z \, \frac{C_{max}}{C}, \tag{2}$$

where Z and C are short notation for the occupancy and capacity functions for a set of locations This approach yields occupancy fraction, the fraction of observed spaces to be found occupied:

$$OF = \frac{Z}{C} = \frac{\hat{Z}}{C_{max}}. \tag{3}$$

When a sensor fails, backfilling corrects the failure depending on the quantities $\frac{Z}{C} \approx \frac{Z-1}{C-1}$. The accuracy of this approximation is proportional to the number of remaining active sensors. The average block face has 12 sensors.

Zoeter et al. (2012) found that sensor failure is unique per location, time and sensor. By fitting sensor noise models a more accurate prediction of a street's state is available. One could also use a continuous interpolation of this quantity that works well even for few sensors. Given the limited time and simplicity of this approach I use proportional backfilling.

4

ANALYSIS: EFFECTS ON BEHAVIOUR

One may wonder whether LA Express has come closer towards fulfilling its goals: reducing the share of cruisers in traffic. In the natural experimental setting of LA Express I have no pre-project data, baseline or control group to compare with, making this a difficult setting for analysis. I only have partial observations and it is not possible to measure cruising directly. From Millard-Ball et al. (2013a) it is known that a reduction in average occupancy should also be followed by a reduction in the number of blocks cruised on average if no externalities apply. Related to average occupancy is the parking duration. If parkers stay longer, average occupancy goes up.

I first need to discover city wide trends throughout the project. I then try to find an answer to the question whether rate changes have had an effect on average occupancy and parking duration in downtown Los Angeles. Lastly, I wonder whether the change in monetary incentive is high enough to motivate people to search for alternatives. Since the highest rate change is at most $1 up or down, I look at extensions of operating hours and their influence on average occupancy.

4.1 CONGESTION AND UNDERUSE

The LA Express pricing iteration scheme as described by Zoeter et al. (2014) uses the fraction of time a block face spent in a specific state. The indexes describing these states are called the Congestion Index, the Underuse index and Just-right index. A block is considered to be congested when it is over 90% occupied and underused when a block is below 70% occupied. Lastly, a block is considered to be just-right when occupancy is between 70% and 90%. By averaging which state a block is in over the course of a week, we can calculate the indexes. For example, a block having a congestion index of 0.5 would mean that during half of the week this block is congested.

Looking at the influence of rates on parking, I expect the biggest change, if any during the midday period from 11AM to 4PM; it is the segment that has the highest number of parkers on average. I have calculated the parking indexes for each block during this period over a time span of nearly 3 years. I aggregated the indexes per block weighted by the number of spaces per block and plotted the result in figure 4. All found trends are highly significant. I observe an upward

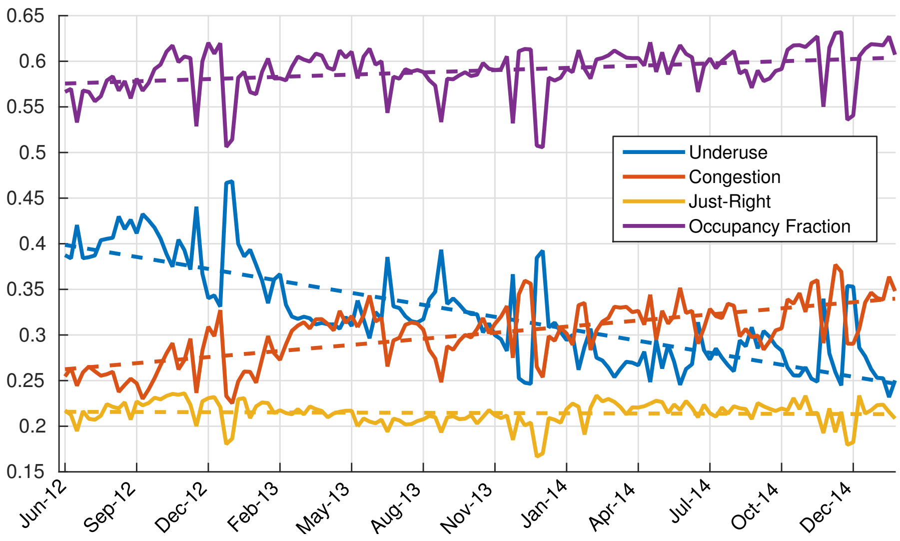


Figure 4: The parking indexes and occupancy in Los Angeles averaged per block per week during weekday operating hours from 11AM to 4PM. The dotted lines are linear fits. A congestion index of 0.5 means on average a block spends half of the week congested.

trend in overall occupancy, indicating more parkers have come to the city. There is an upward trend in congestion and a downward trend in underuse. No trend is found in the amount of time spent just-right.

The seasonality component of the data is clearly displayed. The December holiday season has a distinct spiking pattern, visible three times in the graph. The start and end of the school summer break also has a distinct repeating pattern.

From this figure it is unclear whether the decreased underuse and increased congestion are due to rate changes, but given the increased average occupancy fraction it seems unlikely. LA Express Park lacks a control group, meaning we cannot rule out nor confirm that trends in congestion and underuse have been altered.

An important disclaimer is that, with the exception of occupancy fraction, the trends of the indexes depend on how spaces and blocks are selected and aggregated. Computing just-right per block face is poorly defined for small block faces; what is 70-90% for 3 spaces? Ghent et al. (2014) have done the same analysis, but aggregate on the 20 nearest spaces rather than geographical block faces. While agreeing on an increase in congestion and a reduction in underuse, they report a rising trend in just-right blocks. This may be since their metric is fairer to small block faces.

Not all measured underuse is necessarily bad. Some block faces in the project area have low demand throughout the day since they lie in the peripheral or more industrial parts of downtown. These blocks also have no congestion nearby, meaning they are also not feasible alternatives to walk to. Even if parking would be free on these blocks the block would still be underused.

Also of importance is sensor availability and active sensor noise, illustrated in figure 3. I have seen local trends inflect upward or downward depending on which sensors are excluded.

4.2 influence on occupancy

From the data I know that the variance in occupancy is large, making it difficult to establish any influence from rate changes. It also increases the odds of drawing conclusions based on statistical fallacies. From work by Pierce and Shoup (2013a) and Millard-Ball et al. (2013b) I know that average occupancy is useful for summarising parking behaviour, but poor at showing subtle changes. I did not explicitly compensate for Regression Towards the Mean (RTM) due to a lack of resources, but, skipping ahead, this is also not necessary as the measured effect of rates on occupancy is marginal.

I considered the 13 rate changes that took place between June 2012 and March 2015. I compare the 6 weeks of activity before and after each rate change, the same time window as the studies on SFpark. Any sensor inactive for more than 5% during these 12 weeks was removed from further analysis. The same goes for block faces with less than 3 active sensors after filtering.

I first looked at the influence of rates on occupancy fraction per block as other authors have done before us. Of course, rates only have influence on the paying fraction of parkers. I observed little to no difference in the response after rate changes when excluding unpaid parking, which is why I only included paid use in figure 5. It shows the average price change versus the difference in average paid occupancy before and after each change. The yellow line illustrates the correlation between occupancy change and rate change. I find a Pearson correlation value of $\rho = -0.19$ with a P-value of 6×10^{-8}. This means occupancy decreases as price increases but only very slight: an increase of \$1 corresponds to a decrease in average decrease in occupancy of 3 percent point with large variance. Part of the observed effect is due to RTM since the correlation between rate change and occupancy change is not perfect; chance plays a limited role in the updating of rates. I have not included any figures on separate city districts and rate changes as they show similar results.

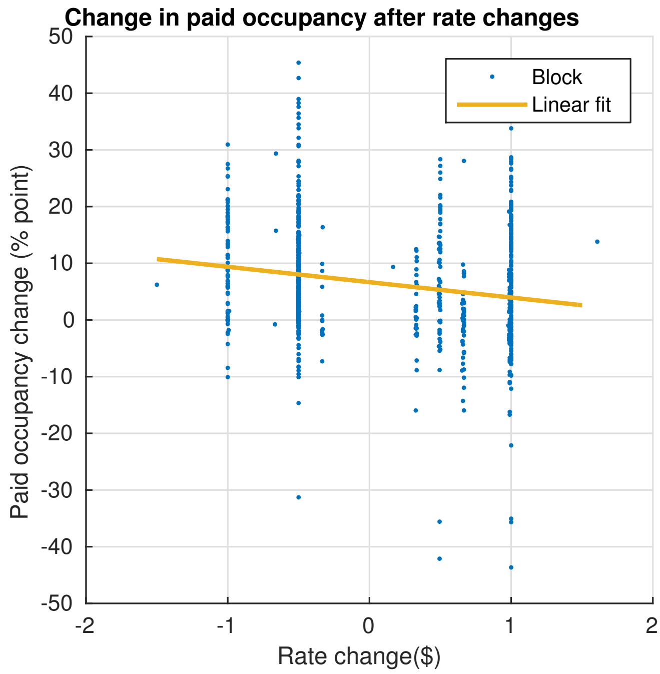


Figure 5: The change in rate versus the change in paid occupancy fraction during midday operating hours 11AM to 4PM not including weekdays for each rate adjustment from June 2012 to March 2015. A dot is a single block face in one change. The yellow line is a regularised least squares linear fit to these points to illustrate correlation.

4.3 PARKING DURATION

The effect in figure 5 – price goes up, occupancy goes down – may partially be due to regression towards the mean, even though I observed no cases where rates have reverted to their previous value. Statistics on the level of the single parker, such as duration of stay and space turnover, suffer less from regression towards the mean. The effect is not completely negligible since occupancy goes up when parking durations increase.

The cost of parking scales linearly with the duration of stay. If prices go up, the intuition is that people park shorter, to save money or because they decide to park elsewhere. This does not hold for unpaid use of parking facilities, which is why I filter out parking events for which no payments were made. I have plotted the average paid parking duration before and after the same rate changes as before in figure 6.

I find a Pearson correlation value of $\rho = -0.19$, with a P-value of 9×10^{-8}. On average, for each dollar prices are increased, people park a minute shorter. While the direction of change is small, it is of the same magnitude as that observed in figure 5.

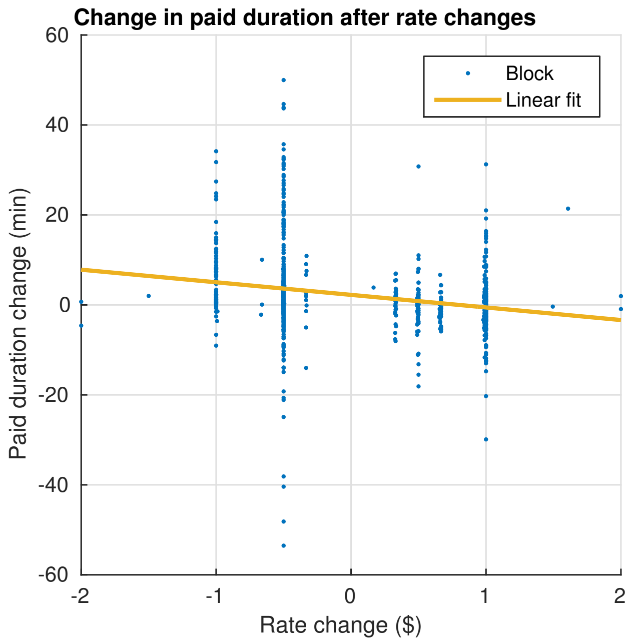


Figure 6: Change in mean paid duration versus rate change during midday operating hours, from 11AM to 4PM and not including weekdays for all rate changes from June 2012 to February 2015. The yellow line is a linear fit illustrating the correlation.

4.4 Operating hour extension

Both Millard-Ball et al. (2013a) as Pierce and Shoup (2013a) have suggested the rate changes to be of too small magnitude to show short term influence. Figures 5 and 6 also support this. Further evidence for this claim may be found in the extension of evening operating hours from 6PM to 8PM on many block faces. During these periods the cost of parking went from $0 to $4-8, while the largest midday rate change was an increase of (only) $1. Extension of evening operating hours was done in May and June 2012. I only have partial observations of the extension since data capture began halfway May 2012.

Ghent et al. (2014) did report on the full period of operating hour extension. They note a reduction in occupancy during these periods, which they attribute to the increased cost for drivers. To support their argument they provide a concrete example of a single block where occupancy has drastically decreased after the extension of operating hours. I did a similar analysis and found no reduction averaged over all blocks whose operating hours were extended.

I did however see a decrease in blocks that had high prior occupancy between 6PM and 8PM. This result comes from my analysis towards blocks with high evening occupancy. The blocks taken into consideration all had a consistent average occupancy of at least 70% with little variance between 6PM and 8PM. Due to data limitations I only consider two weeks of occupancy data prior to the time limit extension. Of the 152 blocks where operating hours were extended, 85 of these had high prior occupancy. They are highlighted in blue in figure 7. Highlighted in red are all other blocks, a count of 63, that also had high prior occupancy but whose operating hours were not extended.

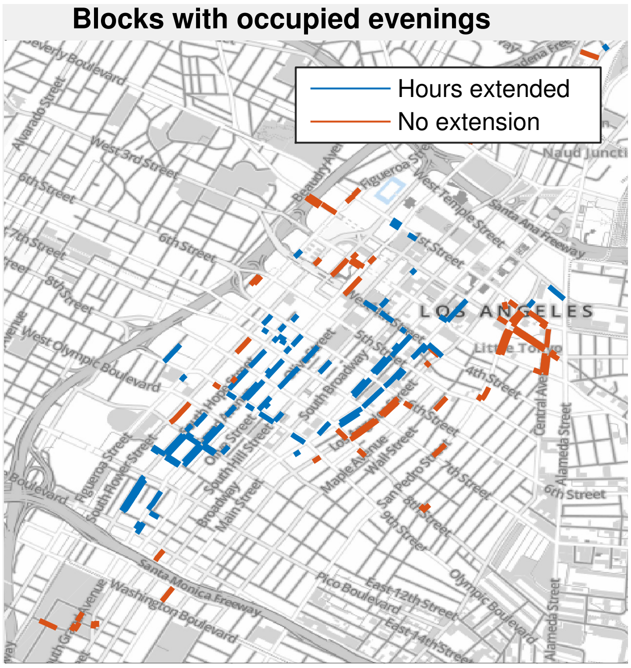


Figure 7: Map showing blocks with evening occupancy between 6PM and 8PM of at least 70% two weeks before the extension of operating hours in June 2012. Shown in blue are the blocks whose hours were extended to 8PM, located mostly toward the center. Shown in red are all other blocks with evening occupancy at least 70%.

Figure 8 shows progression of the daily mean during the 6PM-8PM period on weekdays two weeks before and two weeks after the evening extension. The bars are the evening mean occupancy for each two week period. A clear comparison is difficult since evening rates and no-parking periods were also changed inconsistently within the groups. Outflow of occupancy onto other block faces is not directly visible, making it difficult to pinpoint why occupancy decreases. I see a downward trend in both groups. The group whose operating hours were extended show a decline of 7 percent point on average, while the non-extended group only shows a decline of 3 percent point on average, both values significant with $P < 0.05$. The decline in both groups is due to a city-wide trend.

Further investigation is necessary before attributing the decline of occupancy solely to the extension of operating hours. Also no rate was present and products that go for free require special treatment in

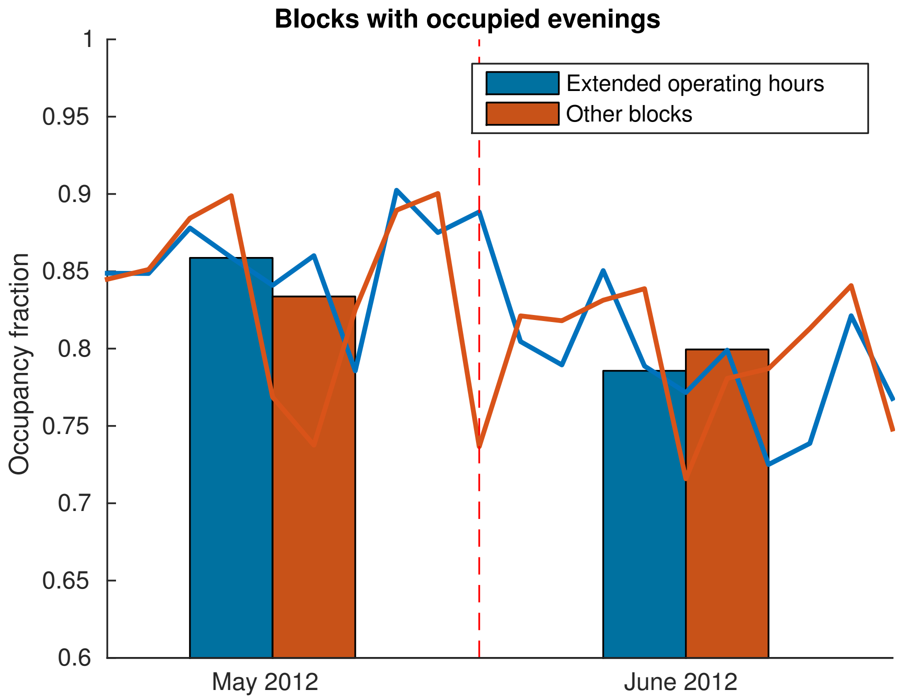


Figure 8: The extension of operating hours in June 2012. In total, 152 blocks had their evening operating hours extended from 6PM to 8PM. Depicted in blue, 85 of these had 70% average occupancy in the period prior to the change. Depicted in red are 63 other blocks whose operating hours were not extended, but also had evenings with a mean occupancy of at least 70%.

economics. This analysis does support the hypothesis that extension of operating hours has a significant effect on the parking population.

4.5 CONCLUSIONS

Overall, my results align with those from the San Francisco studies and the earlier LA Express studies. Measuring driver satisfaction and cruising in terms of average occupancy has not revealed any conclusive results and it is possible externalities drive the observed changes. On average, occupancy decreases after rate changes, in line with the findings by Pierce and Shoup (2013a). I also found that on average parking durations decrease when increasing rates. However, as these effects are so marginal and the role of Regression Towards the Mean as mentioned by Millard-Ball et al. (2013b) has not been quantified no real conclusion can be made.

Further, the price changes appear to be too small to have a drastic short-term effect. The analysis towards the extension of operating hours supports this claim as I observed a significant reduction in occupancy after extension towards operating hours, compared to blocks that did not undergo this change.

I wonder what else might cause the lacking response towards rate changes. Glasnapp et al. (2014) mentioned driver awareness as a possible leading cause. Another possible cause is that the amount of spaces available for paid parking is too low due to unpaid use. I look into this matter in the next section.

5 UNPAID PARKING

Earlier academic work and news reports have suggested many people abuse the current system for handicap parking in California (Shoup et al., 2011; Shoup, 2011; Manville and Williams, 2012; Millard-Ball et al., 2013a; Pierce and Shoup, 2013a; Chatman and Manville, 2014; Glasnapp et al., 2014; Ghent et al., 2014; ABClocal, 2012; Goldstein, 2013). While being a cause of missed revenue, these so-called handicap placard abusers would also lead to a reduced parking capacity for those who are willing to pay. I wonder to what extent this is the case. I can imagine that the response on rate adjustments is lower in areas with high unpaid use due to lower parking capacity. Glasnapp et al. (2014) and Zoeter et al. (2014) have suggested the use of no-parking periods to demotivate placard abusers from parking their cars. I wonder whether changing no-parking periods has an effect on the amount of unpaid parked time.

Problems of interest are how to measure placard usage and how to discern it from regular unpaid use. We need to know how unpaid usage is spread over the city and how this relates to areas already suffering from parking congestion. Lastly, I look at how daily parking capacity is divided over different usage groups to see whether there is indeed reduced capacity for paid parking.

In California it is possible to obtain a special parking permit that enables its owner to park limitless and free-of-charge on any space, with the exception of parking during no-parking periods (California Department of Motor Vehicles, 2005). Most importantly, pricing mechanisms do not effect placard users. While originally intended for the disabled, my colleagues and I and the previously cited academics have strong suspicion their use goes far beyond just helping the less-abled. Glasnapp et al. (2014), Zoeter et al. (2014) and Pierce and Shoup (2013c) reported numbers as high as 75-90% of parked cars on specific block faces in Los Angeles having placards. Regardless of abuse, Vickrey (1954) already stated that all exceptions to parking rules are impeding the success of demand-based pricing.

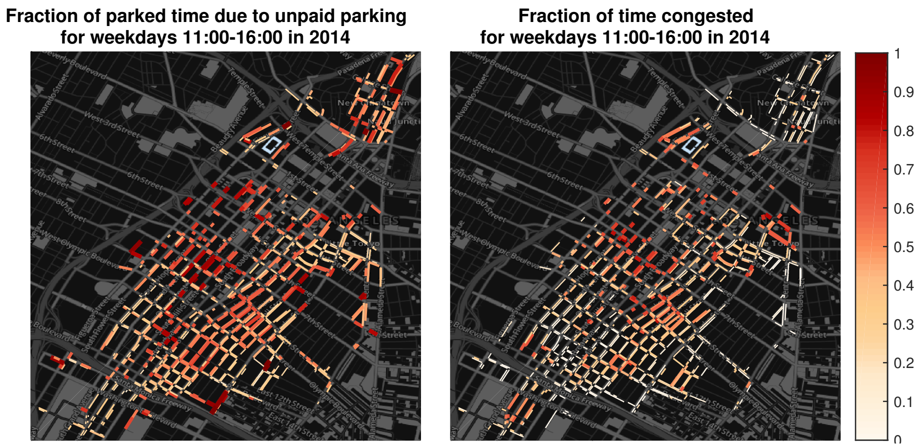
5.1 MEASURING PLACARD USE

We have no means to directly detect placard users from the data, but it can be estimated. Paid use only makes sense when considered for cars that were parked during some part of operating hours. A person in possession of a placard has no need to pay. When counting unpaid use, we count precisely those events where no payment was received at all.

Since my focus lies with placard abusers, I do not consider events where a payment was done, but the amount was insufficient to cover the stay. *Piggybackers*, parkers who did not have to pay since there was time left on the meter, are also excluded. It is impossible to further discern between violators and placard users with full certainty.

5.2 UNPAID USE AND CONGESTION

In figure 9a I show the fraction of parked but unpaid time averaged per block during the midday weekday operating hours in 2014. I consider the first five minutes of stay to be a slack period during which the parker has time to pay. On the right, figure 9b shows the fraction of time a block face was in a congested state, i.e. more than 90% occupied during midday weekday operating hours in 2014.



(a) Unpaid parked time

(b) Midday congestion

Figure 9: Maps of downtown Los Angeles showing the percentage of unpaid parked time next to the fraction of time streets are congested, having an average occupancy of 90%. The financial district, roughly the middle of the map, suffers from highest congestion and unpaid use.

The fraction of unpaid parked time is never lower than 25% while going as high as Base percentages are almost never lower than 25%, while going as high as 90%, confirming results by Glasnapp et al. (2014), Zoeter et al. (2014) and Pierce and Shoup (2013c). On average, block faces have about 45% of the parked time go unpaid.

In particular the business and financial district suffer from very high congestion, spending more than 70% of operating hours congested. When comparing with the fraction of unpaid parked time on the left, I notice that it is exactly these areas that also suffer from high unpaid use.

When looking at the effect of rates on unpaid use through time I see no effect. Little has changed since the start of the project in 2012, shown in figure 10.

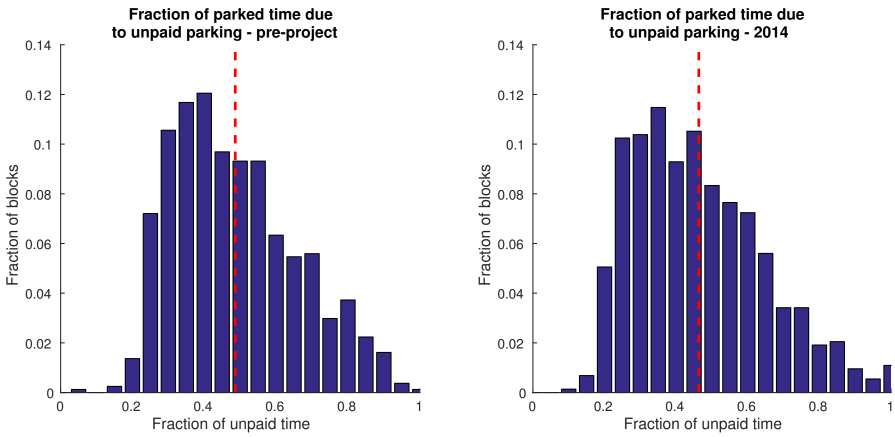


Figure 10: The distribution of fraction of unpaid parked time per block between May and June 2012 and January to December 2014. The red dotted line indicates the mean fraction of unpaid time, in operating hours of course. On a whole-city scale the project has had little influence on the overal distribution of unpaid parking.

In most streets with high unpaid use I found it possible to identify cars with placards on Google street view. While it is not an accurate method of measurement – not all rear-view mirrors are visible – I did find it anecdotal. See for instance figure 11 taken on North Fremont Avenue.

5.3 UNPAID USE AND CAPACITY

I found high percentages of unpaid use, going as high as 90% on specific block faces. How this reflects the leftover capacity and parking opportunities depends on the duration of stay and the fraction of cars not paying for their visit. In figure 12 I show for different neighbourhoods of the Los Angeles downtown district how occupancy is distributed among different groups of daytime parkers. The data is averaged for weekdays during 2013 and 2014.

The severity of unpaid use throughout the day is present in all neighbourhoods. Especially early on the day around 8AM and later on the day at around 5PM parking is dominated by non-paying parkers. I es-

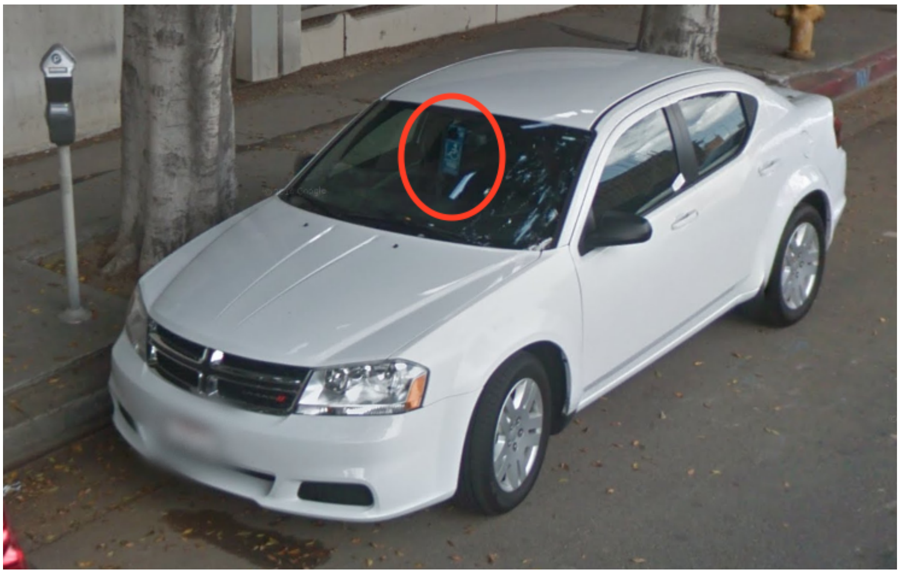


Figure 11: An example of a car with a handicap placard parked on North Fremont Avenue. Taken from Google street view, May 2014.

tablished earlier that many downtown blocks are in a congested state during peak hours. If during these moments, the majority of these cars is unpaid for, I imagine this contributes to additional cruising. On many other moments of day more than half of the parked cars are not paid for, in particular in the civic center, Chinatown and the central business district.

While their appears to be a fraction of spaces vacant most of the time, this is actually a side effect from averaging over so many blocks. If I was just to include blocks suffering from congestion it would provide a biased view.

5.4 THE INFLUENCE OF NO-PARKING PERIODS

Glasnapp et al. (2014) and Zoeter et al. (2014) have suggested the use of no-parking periods to demotivate placard abusers from parking their cars. I suspect these periods specifically influence the group of non-paying parkers that arrive early and stay late, due to the interruption in their schedule. I reviewed 28 cases where no-park periods were removed in the mornings (7-9AM) and late afternoons (4-6PM). I observed only one case of introduction of a no-parking period, in the late afternoon. Sadly, these periods do not fully intersect with common office working hours making it difficult to test this hypothesis.

In all cases of removal of a no-park period there was an increase in parker turnover and revenue from different groups of parkers. In 12 cases of removal of a no-park period I saw a significant increase in

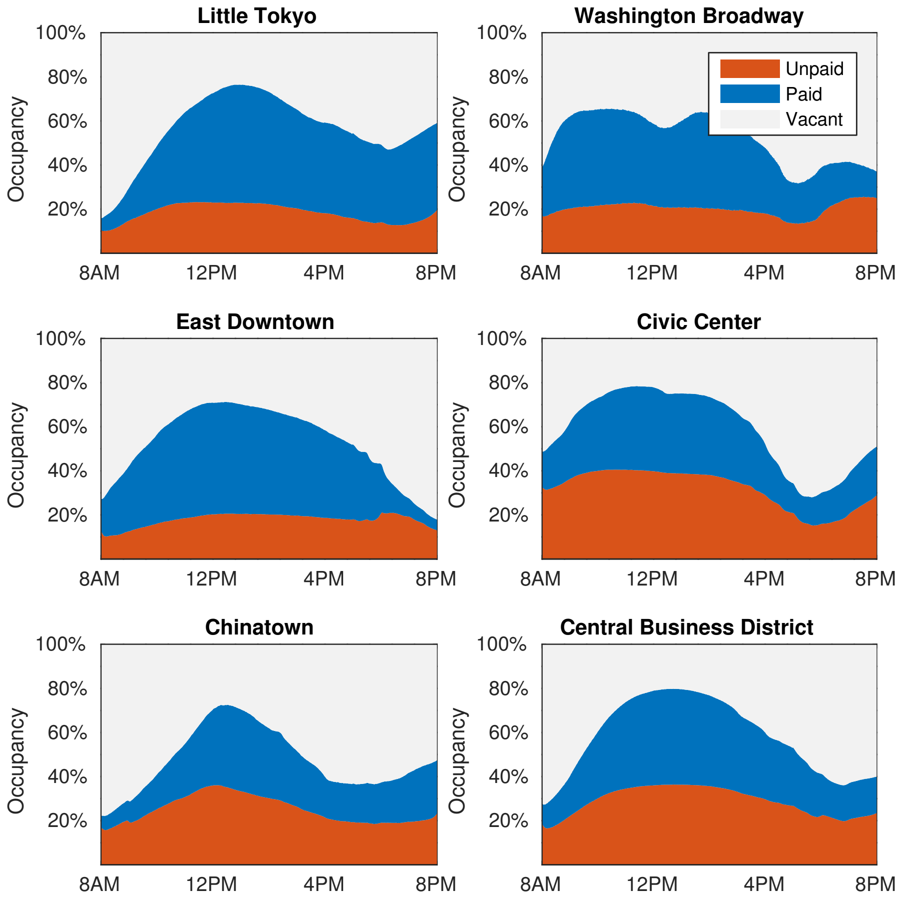


Figure 12: The distribution of parking capacity in downtown Los Angeles averaged over January 1st 2013 to December 31st 2014. In particular the civic center, Chinatown and the central business district suffer from reduced capacity due to unpaid use.

the amount of nonpaying parkers staying longer than 6 hours. Figure 13 shows in red the increase in percent point in occupancy on these 28 blocks due to unpaid use. The last blue bar is the mean change in unpaid occupancy for more than 1800 other changes throughout the project.

Compared to other types of rate changes, removing no-parking periods leads to an increase of nonpaying parkers that is disproportionate to the additional turnover. The largest of these changes are due to an influx of long-staying nonpaying parkers, on for example 700 South Olive street depicted in figure 14. Ghent et al. (2014) also confirmed this, while not directly using payment data.

This example shows an influx of non-paying parkers around 8AM who stay for the rest of the day. While this is just one block, most blocks in the analysis of this thesis have a similar group of parkers that do not pay and stay for the entire workday. This is a group of parkers returning frequently throughout the entire city. To us this suggests

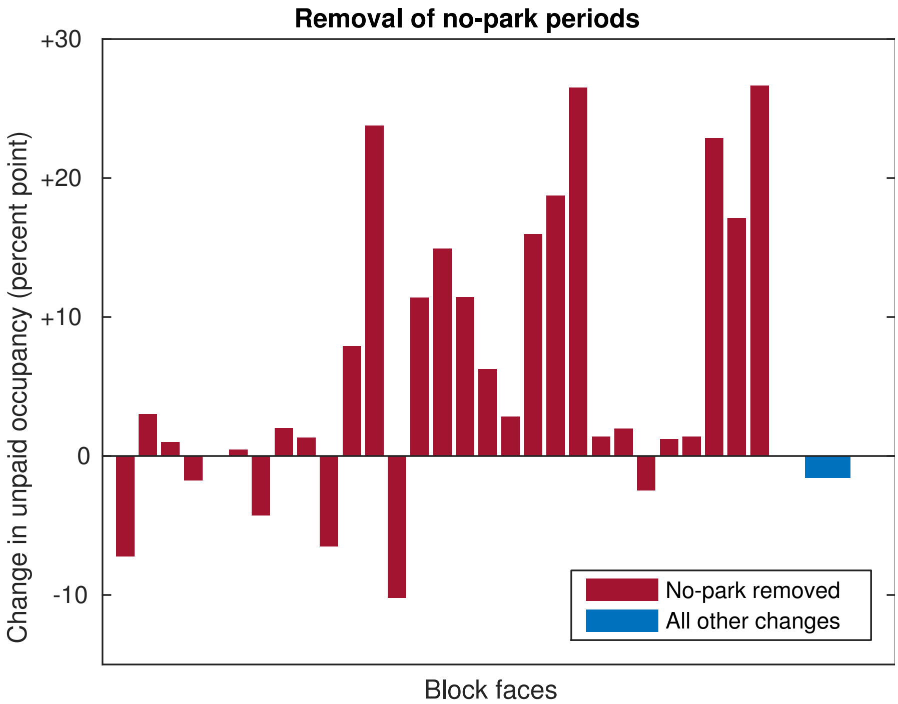
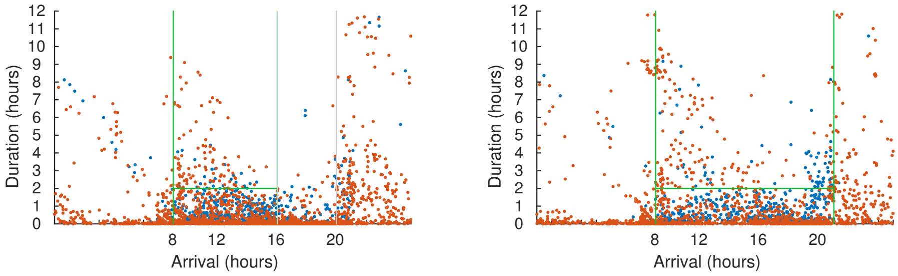


Figure 13: The change in occupancy fraction due to unpaid parking activity for two types of rate changes during the entire project. In red: removal of no-park periods from 7AM to 9AM and 4PM to 6-8PM on 28 block faces. In blue: the average of over 1800 other policy changes.



(a) Before removal (b) After removal

Figure 14: Arrival-duration plots before and after the removal of no-parking from 4-8PM at 700 South Olive street. The green bars indicate the start and end of operating hours. After removal there is an influx of non-paying parkers arriving around 8AM, who stay for the rest of the day.

this group consists of commuters who use a handicap placard to cover their trip to work.

Since there were no cases of removal or introduction of no-parking periods in the middle of the work day, I cannot draw conclusions on whether or not this is beneficial for tackling placard abuse. Best would be to get rid of the law allowing placard abuse in its current form. All other procedures to tackle unpaid use would be proxies of a true solution. Modification of no-park periods should be regarded case-to-case as they are quite intrusive procedures that largely change the type of parker that can be served and the total block turnover.

5.5 EFFECTIVE SUBSIDISATION AND MISSED REVENUE

As a further emphasis on the problem regarding handicap placard abuse I calculated the effective subsidisation the city spends on unpaid use; how much revenue is missed due to unpaid use.

LADOT (2013) reported meter revenue figures of $55M for 37,560 metered on-street and off-street spaces from July 1st 2013 until July 1st 2014. Upon inquiry, LADOT reported $13.1M of this amount is from the 6,150 metered on-street spaces from the LA Express Park project. I calculated the cost of stay for each individual paid parking event, not taking into account the first 5 minutes after arrival for spaces following the filtering described in section 3. Doing so I obtained the figure of $9.4M in earned revenue.

I did the same cost calculation for all events that went unpaid, obtaining the figure of $9.5M in missed revenue, or effective subsidisation towards unpaid parking. It suggests one could double revenue by getting rid of placard abuse. This sum, about 50% of the total parking cost, also corresponds with the earlier calculated average of 46% of parked time to go unpaid. A possible explanation why the missed revenue is higher than the actual revenue may be found in the fact there is more unpaid time in areas with higher hourly rates. I have illustrated this in figure 15, where areas around the high rise part of town have been found to have the highest missed revenue.

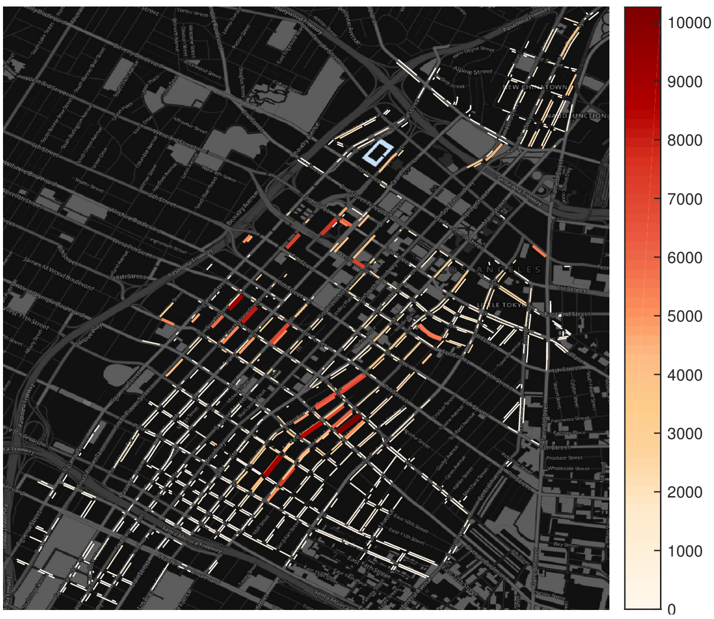


Figure 15: Missed monthly revenue per block from July 1st 2013 to June 30th 2014 during midday operating hours, averaged per block. Midday is from 11AM to 4PM. Most revenue is lost in the fashion district, the area in the lower middle of the map.

Figure 16 shows how effective subsidisation and revenue fluctuate throughout the project. Large downward spikes match up with system failures for multiple blocks.

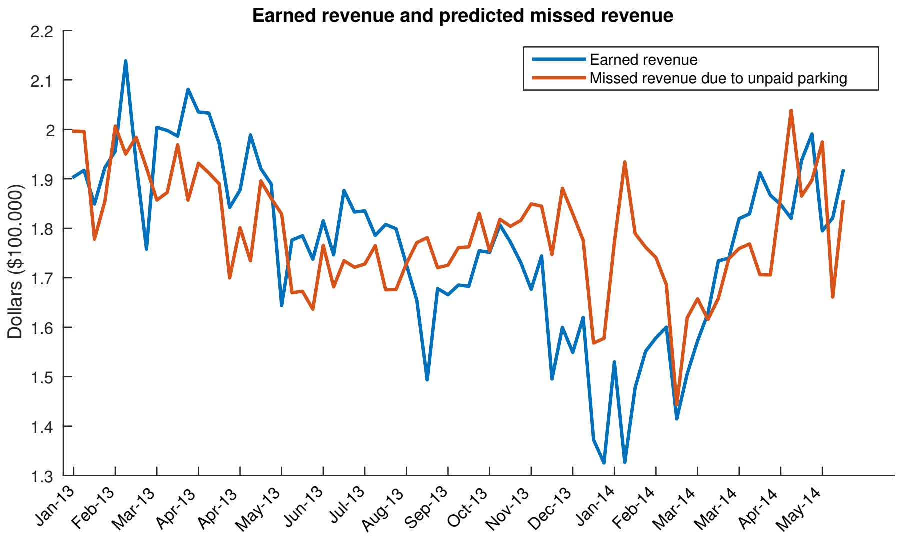


Figure 16: Earned and missed monthly revenues from on-street meters in downtown Los Angeles during midday operating hours. Midday is 11AM to 4PM.

The differences in the calculated revenue and the reported revenue can be explained from a number of factors. I calculated revenue based on duration of stay and at least one payment done during stay, opposing actually summing the paid amounts. I choose this approach as this allows me to calculate the cost of unpaid events. Further, I considered only the available meters, instead of the full 6,150. Third, I obtained payment data indirectly, while the LADOT has direct access to payments from the meter vendor. In section 3.2 I reported 5% of the payments to be missing compared to data directly from the meter vendor. I emphasise that the amounts reported are merely indications of the severity of the situation and should not be followed by the letter.

5.6 CONCLUSION

I found unusually high values of unpaid parked time in downtown Los Angeles, which I think is due to the abuse of the special handicap placard system. The average unpaid parked time of both placard users and violators in Los Angeles is about 45%, with numbers going as high as 90% in congested areas. Most unpaid use is concentrated in high-demand areas as many congested areas also have high unpaid usage. The majority of unpaid parked time is due to long-stay unpaid usage, most likely from commuters using a handicap placard. Changing rates has had little influence on unpaid usage throughout the project. Most revenue is lost in areas with high turnover and costs, such as the fashion district and revenue can be nearly doubled if the city was to tackle unpaid use.

Regarding no-parking periods, removal mostly leads to an influx of long-stay unpaid usage. I observed no cases of introduction of no-parking periods. Future research is necessary to draw conclusions about their influence on unpaid use.

It may be the case though that previous rate changes may have been wrongly applied as these did not take placard use and unpaid parking into account, specifically on block faces in the business and financial districts. Rate changes had no visible influence on the amount of unpaid use over the course of the project.

I recommend the city of Los Angeles and all other cities with similar laws on parking rules to adjust their legislation to no longer exempt specific groups of people. An alternative mechanism would be to pay out a fraction of the current effective subsidisation in parking vouchers so to return the market mechanism. Another option is to adopt the European system of handicap parking and reserve a fraction of on-street spots specifically for handicap parking.

This dataset on unpaid use can also be used for the detection and prediction of unpaid use on a minute resolution. Such predictions and detections serve many purposes, such as to direct meter maids towards areas of high expected unpaid use, or to modify flawed parking rules.

6 DISCOUNT BY WALKING

In section 4 I looked into the efficacy of rate changes and noticed a low response. I sought a possible explanation for the lacking response in the high levels of unpaid parking and handicap placard abuse that may reduce available capacity for paying users in section 5. A different reason may be that rates are increased by the same amount in the entire area and no cheaper alternative is created nearby. In this short section I look into the development of cheaper alternatives as function of the walking distance.

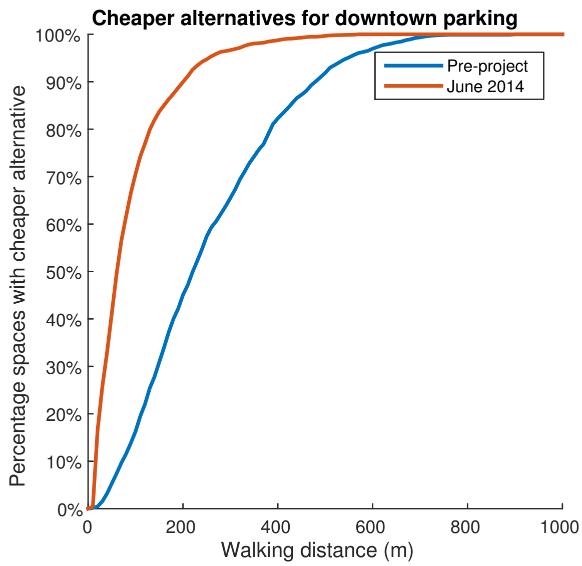


Figure 17: Fraction of downtown on-street parking spaces with a cheaper alternative in a specific walking distance.

I calculated the walking distances between spaces using the Open-StreetMap routing by Luxen and Vetter (2011), rather than as the crow flies. As a measure of heterogeneity I computed the fraction of spaces downtown that have a cheaper space nearby as function of the walking distance. I compare between the rates from early 2012 and those as of June 2014. Figure 17 shows these fractions. The blue line describes the city as it was before the project early 2012 when rates were still zone based. The red line corresponds with the city parking rates as of June 2014 after 10 rate update iterations. Both lines hit 100% since I excluded the cheapest spaces ($\$0.5/h$) from having a cheaper alternative, while they still may serve as an alternative to other spaces. All spaces on the previously shown maps as well as spaces outside of the project border area are taken into account. I obtained rate data for these neighbouring areas directly from representatives of the LADOT.

This graph shows that rates are much more heterogeneous in 2014. Already most spaces have a cheaper alternative nearby when walking just 100 meters. From Glasnapp et al. (2014) I learned that most people would be willing to walk $2 \sim 3$ blocks, approximately 250 meters, to park at a cheaper alternative. To get a better idea, I show the spaces with a cheaper alternative nearby for 250 meters walking distance in figure 18. The left side shows the rates as they were in January 2012 and June 2014. The right side shows which spaces have a cheaper alternative nearby.

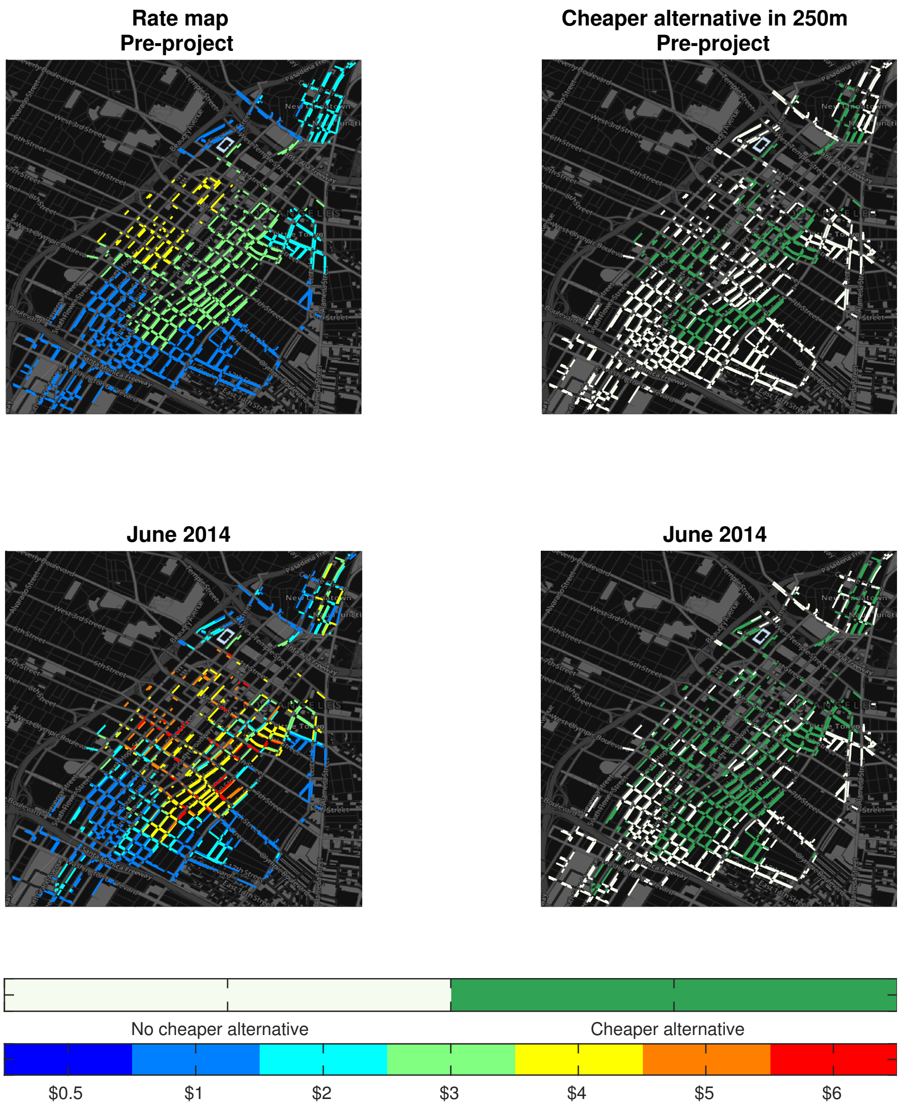


Figure 18: Left: a comparison of rates in 2012 and 2014. Right: spaces with a cheaper alternative within 250m walking distance.

The figure supports my intuition that only traffic near zone boundaries is incentivised to change spaces and reduce parking congestion.

Having demand based rates on the other hand will provide incentives to reduce congestion at nearly all locations. The plot and maps graphically emphasise that demand based rates are a way to maximise the area where parkers are incentivised to change their behaviour. The blocks where no cheaper alternatives are available, depicted as white in figure 18 are those that are already on the lowest step of the rate ladder, $\$0.5/h$. Figure 9b in section 5 shows these blocks do not suffer from major congestion problems, reducing the need for incentives of change.

Like Ghent et al. (2014) points out, the most expensive areas in 2012 are still the most expensive in 2014. Also those areas that were the cheapest in 2012 are still the cheapest in 2014. Rates have reached up to $\$6/h$ in the central business district, while this was $\$4/h$ previously. Similarly, underused areas south of town have dropped in price, from $\$1/h$ to $\$0.50/h$. From the LADOT I learned that rates in the central business district are kept artificially low; the rate change algorithm has suggested rates to go up to $\$8/h$. However, due to the high number of nonpaying parkers the city is hesitant to increase the rates further in that area.

In practice, the incentive of change is also determined by the size of a possible discount which depends on the duration of one's stay; walking cost is incurred once while the hourly rates scales with duration. Other factors, such as the probability of finding a space and the parking duration should be taken into account before quantising the incentive to change parking behaviour.

7 RATIONALITY MODEL

I wonder how the city would look like if we were to re-arrange cars according to their best parking alternative. McFadden et al. (1978) developed discrete choice theory, motivated by problems from transportation. For this work McFadden was awarded the nobel prize in economics. Using this and utility theory I describe a simple discrete choice model with the primary goal to show where incentives for cheaper parking are currently in place and how people would have parked if they had better understanding of their environment. Discrete choice models describe how people choose between a limited number of options given their interests and their environment.

This discrete choice model describes how people pick a parking space given a limited number of options and what they could gain by switching to a cheaper alternative. The previous two sections already hint at this subject: areas with high unpaid use are less flexible and areas with heterogeneous rates provide a wider range of parking options. To provide a full view we also need to consider the local driver behaviour and the probability of finding an empty space. Parkers staying for longer periods get a higher total discount from a per hour rate reduction than those staying shorter. These factors have not yet been accounted for in previous parking research.

Utility represents the satisfaction experienced by the driver when successfully parking his car. As it cannot be measured directly, utility is taken to be correlative to demand and observed parking choices; people's willingness to pay different amounts for different spaces. See for instance Marshall et al. (1920).

7.1 A MODEL FOR PARKING VALUE

To describe the value of parking, I first introduce some notation. In utility theory, each parker has a specific value v_0 how much he wants to be at his final destination l_0. This determines how far the parker is willing to walk, what approach route to take and how much he would pay for his stay. A single parking event is a combination of location l, arrival time a, duration of stay d and whether or not the stay was paid for p, where $p \in \{0, 1\}$.

My first assumption is that if a parker decides to stay at location l, their respective value is proportional to the incurred monetary cost and time spent driving and walking. I assume these costs scale linearly. The parkers' personal values of parking and time are encoded by $\theta = \langle v_0, \alpha \rangle$ as linear tradeoff coefficients. The parker's travel is described by $x = \langle l_0, a, d, p \rangle$. Together, this amounts to the following equation for the value of a parking event:

$$v_{net} = V(\theta, x, l) = v_0 - 2\alpha \mathrm{T}(l, l_0) - p\mathrm{R}(l, a, d). \tag{4}$$

Here V is a function of parker and space properties. Any distance traveled to a final destination has to be traversed back as well, hence the factor 2.

The time spent T is a distance function of the distance between the parking location and the final destination: $\mathrm{T}(l, l_0)$. The monetary cost is a function of the location parked, arrival time and duration of stay: the rate function $\mathrm{R}(l, a, d)$. Other types of policies can be encoded as well. Parking during a no-parking period results in a towed car and a fine. Staying longer than the allowed time limits likely results in a fine. Parking outside of operating hours means parking is free.

7.2 CHOOSING A SPACE

In a practical setting we have many drivers and many parking alternatives. I assume drivers want to maximise their parking utility. Further, arrival times and durations are assumed to be constant regardless of where one would park. I encode the choice of a user n for alternative i as the choice variable Y_{ni}. While equation 4 is deterministic, describing the choice process also includes driver considerations that remain unobserved. These latent properties are encoded in the random error variable ϵ. The maximum utility choice is then also a random variable described as

$$\begin{aligned} P_{ni} &= P\left(Y_{ni} = 1\right) \\ &= P\left(v_{ni} - \epsilon_i > v_{nj} - \epsilon_j, \forall j \neq i\right) \\ &= P\left(v_{ni} - v_{nj} > \epsilon_i - \epsilon_j, \forall j \neq i\right) \\ &= P\left(V(\theta, x_n, i) - V(\theta, x_n, j) > \epsilon_i - \epsilon_j, \forall j \neq i\right). \end{aligned}$$

Put in words, if the relative difference in costs between space i and all others is large enough, parker n chooses space i with high probability.

Written in full, the above becomes

$$\begin{aligned} P_{ni} &= P\left(V(\theta, x_n, i) - V(\theta, x_n, j) > \epsilon_i - \epsilon_j, \forall j \neq i\right) \\ &= P\left(2\alpha \left[\mathrm{T}\left(j, l_0\right) - \mathrm{T}\left(i, l_0\right)\right] + p\left[\mathrm{R}\left(j, a, d\right) - \mathrm{R}\left(i, a, d\right)\right] \right. \\ &\qquad \left. > \epsilon_i - \epsilon_j, \forall j \neq i\right). \end{aligned}$$

Utilities are interpreted as ordinal quantities. We do not have to know the actual utilities v_0, just the utility in respect to other spaces. In fact, after the subtraction v_0 has disappeared from the above equation.

I adopt the soft-max multinomial choice model as described by McFadden et al. (1978) and Ben-Akiva and Bierlaire (1999). The model is also known as multinomial logistic regression, a generalisation of logistic regression to problems with more than two outcomes. See for instance Bishop et al. (2006). The error term ϵ is assumed to be i.i.d. type 1 extreme value distributed. By integrating over the error term ϵ, we obtain the following expression for the choice probability:

$$P\left(Y_{ni} = 1 \middle| x_i, \theta\right) = \frac{\exp\left(\gamma V(\theta, x_n, i)\right)}{\sum_{j=1}^{J} \exp\left(\gamma V(\theta, x_n, j)\right)}, \tag{5}$$

where γ is a smoothing parameter describing how well people are informed. A low value of γ would spread the probability mass over multiple spaces, resulting in multiple viable alternatives. In reality there is also never one ideal space.

The previous equations become useful as soon as we have distributions describing how people park and what their odds are of finding an available space. In a Bayesian approach, we want to find posterior probability distributions describing the optimal drive choice model parameters $\theta = \langle \alpha, \gamma \rangle$ given a dataset of observations $\mathcal{D}$. The posterior is described by Bayes Rule for probabilities:

$$P(\theta|\mathcal{D}) = \frac{P(\mathcal{D}|\theta)P(\theta)}{\int P(\mathcal{D}|\theta')P(\theta')d\theta'} \propto P(\mathcal{D}|\theta)P(\theta). \tag{6}$$

Here $P(\mathcal{D}|\theta)$ is the parameter likelihood which describes how probable the data could have been generated from the model given a set of parameters. The term $P(\theta)$ is a prior describing the *a priori* belief of the parameters' values.

In a full Bayesian approach we would have probability distributions over the best parking spaces, which provides the modeller with a sense of certainty over the model predictions. This requires calculation of the integral in equation (6), which is often computationally heavy or even intractable. Instead people often choose parameters that maximise the proportional quantity on the right hand side. While this approach, appropriately named Maximum A Posteriori (MAP), does not allow the computation of a full predictive distribution we are still able to incorporate prior beliefs over parameters into the model fit.

The likelihood term $P(\mathcal{D}|\theta)$ depends on the composition of the previous observations. In the Los Angeles data we observed parking events without knowledge of trip information such as the different driver approach routes and final destinations. The latter is a problem since I explicitly included the walking distance from a space to the final destination in the value model. Assuming we have a discrete and finite number of final destinations L, we may decompose the likelihood term to try and sum out this latent value:

$$\begin{aligned} P(\mathcal{D}|\theta) &= \prod_{n=1}^{N} \sum_{l_0 \in L}^{L} P(l_0, \{l, a, d, p\}_n \,|\theta) \\ &= \prod_{n=1}^{N} \sum_{l_0 \in L}^{L} P(l_n|l_0, a_n, d_n, p_n, \theta) P(l_0, a_n, d_n, p_n), \end{aligned}$$

where we've decomposed the likelihood into the model distribution from equation (5) and the trip demand process.

In total there are L^N possible assignments of final destinations to parked cars. Since final destinations are dependent on trips and the choice of parameters, we need to sum over all of these to obtain an exact estimate for θ. This is not easy to do. Even if we limit the number of final destinations to a walking radius around the parked location, run-time complexity would still be of order $O(L'^N)$, where L' is the number of parking spaces considered. Given the size of my dataset, exact inference is intractable. Approximations with Expectation Propagation or MCMC methods are possible, but are left for future work. Instead I simplify the model through modelling assumptions.

The first assumption is that the driver's final destination is negligibly close to their observed parking space. If we disregard driving towards the space, we can calculate distances using the observed space. For some observed event $\langle i, a, d, p \rangle$, where i is the parked location, we say $\mathrm{T}(i, l_0) = 0$ and $\mathrm{T}(j, l_0) = \mathrm{T}(j, i), \forall j \neq i$.

While it is certainly a doubtful assumption, we no longer have to sum over final destinations. This allows us to operationalise the model while taking into account the essential factors of parking duration, non-payments, relative discount, distance and availability in arguably the simplest possible way.

Second is that since we do not know the driver's true arrival time at his final destination, we cannot modify the parking arrival time and duration based on different alternatives and the walking distances between them. I assume the observed arrival time and duration to be independent on the choice of space alternative.

A consequence of these assumptions is that it is no longer possible to find a meaningful value for α, the value of time, based on the observations. Since I assume people could park in front of their final destination, the best model choice to match with past observations is the one matching exactly, thus requiring no walking at all. The best value for α and γ would be infinitely large, meaning people *really* hate walking. Instead I could try a sensitivity analysis for realistic values α and γ to determine the impact of variance on the model outcome.

Even with these assumptions, the model is arguably the simplest model to account for important factors – relative walking distance, relative discount, non-payment – that previous work did not include. The model is realistic enough to highlight unexpected causes for mismatches between prediction and observation.

7.3 COUNTERFACTUAL REASONING

Instead of fitting distributions of demand, there is the simpler way of sampling from past parking events. We can replay past events from the dataset of observations $D = \{\langle l, a, d, p \rangle\}_N$ to see what people would have done if they had full information about their surroundings as a form of counterfactual reasoning. This is similar to a situation where all drivers have onboard navigation directing them to an available parking space.

To realistically model the redistribution of past events, we need to account for capacity and availability of streets and spaces. This requires a simulation model that keeps track of the city's parking state. We could distribute each event to the space with the most probability mass given the current occupancy state, but then we have to deal surplus cars that don't fit in the new distribution. I have no data on traffic flow so we cannot directly sample new events to replace the cars elsewhere in the city. As available time for this work was limited I instead relaxed the constraint of capacity in the model; it is possible for more cars to park than capacity allows. This simplification is also motivated by the final use of the model: I will only use the model to identify locations with a particularly large increase or decrease of occupancy under the simplified rationality model. For this usage an occupancy over 100% is an indication of high demand or the presence of incentives that motivate parking at this block.

Ideally, whether a space is a viable alternative follows from the detailed simulation, but this is harder to implement and more computationally expensive to run. Its predictive performance would also be very susceptible to the model fit. Looking ahead (at for instance figure 20) we see that the simplified model does not fit perfectly. To still include a notion of space availability I weigh the earlier notion of utility with the probability a space is likely to be available; it may be that a cheaper street is nearby, but there's no point in moving if it is likely to be full.

A space l is said to be available if $Z(L) = 0$. The new notion of utility would then be:

$$v_{net} = V(\theta, x, l) = P(Z(l) = 0|a)\left[v_0 - 2\alpha \mathrm{T}(l, l_0) - p\mathrm{R}(l, a, d)\right].$$

To clarify, on a street level, availability is the opposite of occupancy, defining what percentage of spaces remain vacant. We estimate the probability of a space being available as the mean availability on weekdays at the time of arrival in the 6 weeks before the event. As a form of smoothing I take the mean occupancy per quarter of day. For example, if a car arrives on 1:04 PM June 4th 2015, the mean availability is determined by the average availability between 1:00PM and 1:15PM from April 23rd to June 3rd, excluding weekends.

7.4 INCENTIVE PREDICTION

The observed mismatch between historical data and the model redistribution is a measure of how irrational drivers behave according to the model. We can shift to predicting incentives by replaying past events under different conditions, for instance a change in rates or removal of a no-parking period. I am interested in how well rate changes explain large changes in behaviour by replaying past events under the new conditions after a rate change.

I considered the 13 rate changes that took place between June 2012 and March 2015. To map the local incentives, I look at the 6 weeks of activity within operating hours before and after each rate change, including weekends. I first redistribute each event under the model with the same conditions as when the event originally took place before the rate change. Subsequently, I replay this same event under the conditions which would be present if the rate change was applied. If done for all parking events in the 6 week period, the two redistributions have different occupancy patterns which I call respectively Z_{model}^{before} and Z_{model}^{after}, abbreviated Z_m^{t1} and Z_m^{t2}. The model does not use any parking events from after the application of the rate change. I may compare these quantities to the actual observed occupancy distribution before and after each rate change, $Z_{observed}^{before}$ and $Z_{observed}^{after}$, abbreviated Z_o^{t1} and Z_o^{t2}. These four quantities are in fact functions of time and location determining how many cars were parked at a specific time and space.

I wonder how well the model before-after distribution correlates with the observed before-after distribution as this indicates the predictive value of the model and how well drivers listen to the incentives created for them. To compare block faces, I look at average occupancy fraction per block over the full 6 week period. The function μ_b indicates the averaging of occupancy for a specific block face b over the comparison period of 6 weeks during operating hours. I then compute the following quantities for each block face:

$$\text{Predicted occupancy change} \stackrel{def}{=} \mu_b(Z_m^{t2}) - \mu_b(Z_m^{t1}) \tag{7}$$

and

$$\text{Observed occupancy change} \stackrel{def}{=} \mu_b(Z_o^{t2}) - \mu_b(Z_o^{t1}). \tag{8}$$

since the only change between Z_m^{t1} and Z_m^{t2} is a change in rates, matching signs of predicted occupancy change and observed occupancy change support that the observed change may be driven by the applied rate change.

PARAMETER SELECTION My model assumption that observed drivers park negligibly close to their final destination makes it is impossible to fit a value for α based on the data. Instead I choose values that seem reasonable for α and γ based on current beliefs. Glasnapp et al. (2014) found that on average people were willing to walk 2 to 3 blocks or 250 meters, corresponding with a value of $\alpha = 0.005$. I choose $\gamma = 1$ to spread probabilities over the neighbouring streets as well.

If given more time, I would have analysed how different values for α and γ influence the final model results compared to observed parking events; would the conclusions be qualitatively different if the values for α and γ change slightly? An option would have been to apply one-versus-all cross-validation for each rate change for a range of reasonable choices for α and γ comparing on the model's predictive value.

7.5 RESULTS

I have calculated the model occupancy change from equation (7) and the observed occupancy change from equation (8). I have plotted these two quantities for each block and rate change in figure 19, where I compare the signs of both. I say the model has correctly predicted a shift in parkers when the observed trend under the model matches the observed trend, i.e. when the signs of the quantities match.

I see the model has little to no predictive value of how rate changes influence parking behaviour. In total there are 1233 block face adjustments over the 13 rate changes. 338 of these saw a decrease that was predicted correctly (true negative), 334 saw an increase that was predicted correctly (true positive), 223 blocks saw a decrease while the model predicted differently (false positive) and 338 blocks saw an increase while the model predicted a decrease (false negative).

This means that for $\alpha = 0.005$ and $\gamma = 1$ the model has an accuracy of 0.540: just over half of the proposed changes were predicted correctly, a bit better than random.

While it was never the goal for the model to accurately predict future trends, I seek an explanation for the lacking prediction value in drivers

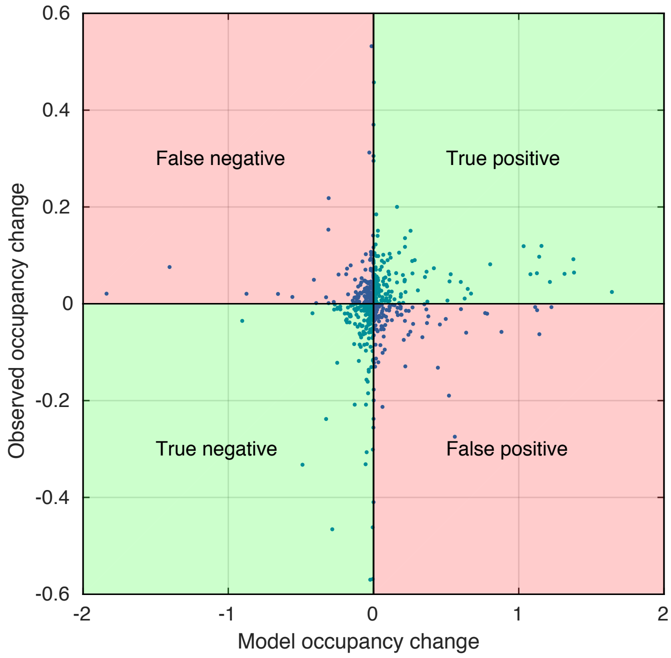
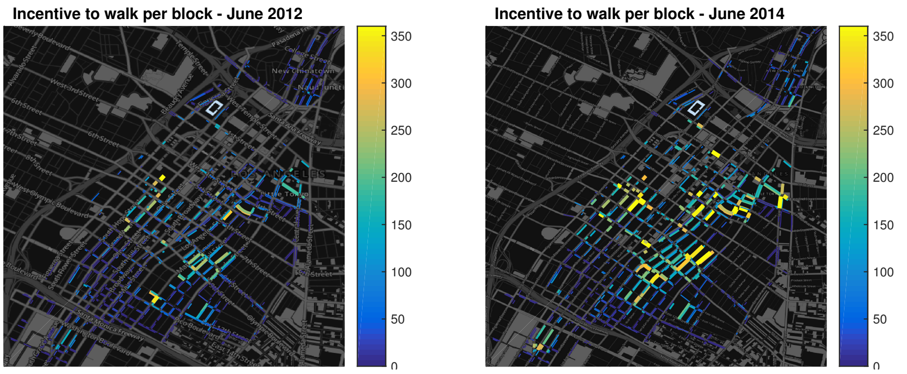


Figure 19: Quadrant analysis of the predictive value of the choice model given 6 weeks of observations before each of the 13 rate changes between June 2012 and March 2015. The values on both axes are in percent point.

not following incentives. I have replayed events from the 6 weeks after the rate changes in June 2012 and June 2014 to get an idea of how and how many cars are redistributed. I show the difference between the observed distribution of cars and the model's distribution of cars for these two periods in figures 20 and 21.

Figure 20 shows the outflow of cars from blocks when comparing the model to the observed situation. This shows the scale at which the model finds drivers could have found better parking opportunities nearby. This outflow or *incentive to walk* is quantised by the amount of dollars that could have been saved per space during the 6 weeks, averaged per block by choosing an alternative space according to the model.

By model definition, only paying drivers are incentivised to consider other spaces; as drivers are assumed to park next to their final destination, non-paying drivers gain nothing by walking. Vice versa, areas with high unpaid use are unattractive as an alternative as most spaces are likely to be already occupied. While in 2012 only drivers on specific blocks of the central business district had incentives to park elsewhere, in 2014 this has expended towards entire areas. Over the two years, the model distributes more drivers to a different spot than they originally parked. This is due to parkers not following the incentives created for them.



(a) Outflow in 2012 (b) Outflow in 2014

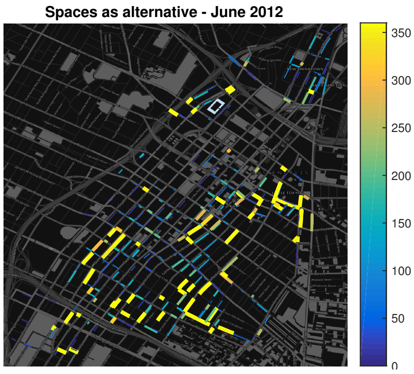
Figure 20: The incentive to walk is expressed in the number of dollars that could have been saved by moving away from a specific space under the model prediction for $\alpha = 0.005, \gamma = 1$. It shows how attractive it is to move out from a specific block face.The amounts displayed in the maps above are the total amount over the course of 6 weeks per space. Much more incentives to move spaces have been created in 2014 than in 2012.

One reason for this increased difference between the model and the observed is that there are more incentives that drivers may choose not to follow; rates have increased and have become more heterogeneous.

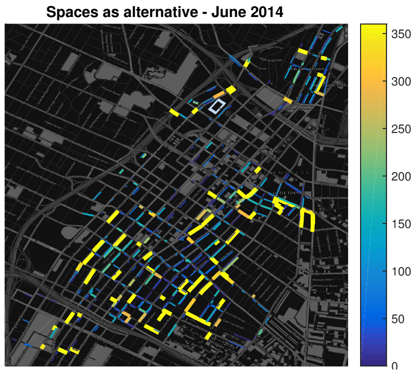
Figure 21 shows the number of times a space has been selected as an alternative compared to the observed during these 6 weeks, averaged per block. The color axis has been truncated at 360 visits as a few number of spaces were selected an order of magnitude more often. While in 2012 the alternative spaces lie near the rate zone borders, the spaces selected by the model for redistribution are located more heterogeneously throughout the city, away from the most congested parts and away from parts with high unpaid use. In total, more spaces were selected as alternative for the rates and events as they were in June 2014 than in June 2012.

7.6 CONCLUSIONS

Los Angeles' drivers do not park rational according to the model, while I expected the model would show some predictive performance. When compared to a straightforward independent-per-block analysis, as all previous work has done, I included relationships that exist between blocks such as walking distance and the relative discount, as well as parking durations and nonpayment. While my assumptions may make the model a little rough on the edges, I do believe the propositions head in the right direction. The gap between the calculated incentives and observed behaviour point to three possible conclusions.



(a) Inflow in 2012



(b) Inflow in 2014

Figure 21: Spaces as alternative is the number of times a space has been selected as a viable alternative under the model prediction for $\alpha = 0.005, \gamma = 1$. Since I do not perform a detailed simulation to exactly include availability, some parking spaces function as sinks which are preferred over the other spaces. For this reason the count has been truncated at 360, after inspecting the distribution of values.These spaces all lie on the edges of what used to be the rate zones in 2012. More spaces have been marked as alternative in 2014 over 2012.

The most obvious is the model is too simple and a follow-up study is necessary, even though I included factors not yet taken into account in previous work.

Second, it could be that parkers are truly not willing to walk. The willingness to walk is governed by α. I can explain the current data by using a value for α that says that people are not willing to walk. By assuming drivers park within negligible distance to their final destination, I already include an error in the prediction if an event would be distributed differently. If it is indeed the case drivers are hesitant to walk, a further increase in rates could motivate these drivers to walk. Rates have not yet reached the cap of $\$8/h$ meaning better results could come at some point in the future.

The third and to us the most likely possible conclusion, parkers are unaware of the parking rates in their close surroundings. The proposed model reasons based on the fact that rates are an important factor in deciding where to park. If this is not the case, the model mismatch with reality will be great, no matter the drivers' behaviour. This idea is supported by Pierce and Shoup (2013a), Zoeter et al. (2014) and the survey by Glasnapp et al. (2014) which showed most drivers are unaware of the hourly rates. If drivers are unaware of the rates, this would also explain the poor response on rate changes in general in earlier work. We can easily confirm this by performing additional surveys. The areas with high calculated incentive could then be used as seed locations for possible awareness campaigns with signs or stickers.

I believe it is too soon to propose complicated models for predicting the influence of rate changes due to the lack of additional trip data. If price does not play a significant part in the driver's decision proces, additional data is required. By integrating parking software in onboard satellite navigation (SatNav) I remove the need for the driver to memorise rates. The incorporation of SatNav software in parking systems would solve the problem of having a latent destination with a two-fold improvement. First, we gain the additional data required to improve the model, such as the driver departure point, approach route, time spent cruising and the driver's final destination. While these variables can already be inferred through complex modelling procedures, integrating parking within SatNav would make these procedures obsolete. In a matter of years most car manufacturers and navigation software producers will start doing so, while some already have (Cunningham, 2013). Secondly the scheduling and distribution of cars over the city and remainder of vacant spaces can be orchestrated by a central parking system once cars communicate with SatNav servers.

8 CONCLUSION & DISCUSSION

Demand-based parking is a promising solution to improve environments in cities. In this way, many cities of ongoing projects hope to reduce the share of traffic looking for parking.

PREVIOUS WORK The field of parking theory has expanded rapidly in the last few years. There are close ties in literature to road pricing, which already has an abundant body of literature. The main theories of demand-based parking as deployed today were written down many decades ago by Vickrey (1954), but technology was lagging behind to realise his ideas. Most existing work is from researchers working in the field of urban planning, economy and computer science.

Analysis of parking data depends heavily on filtering steps. Current methodology seems to be too simple to uncover the possible effect of rate changes on parking behaviour. Combining theory from economics, statistics and computer science is necessary to develop proper methodology for parking time series analysis.

Others have tried to use price elasticity to measure the response of parkers on rate changes, but did not compensate for biases, side effects and statistical fallacies such as regression towards the mean. While useful in other parts of economics, conclusions drawn from elasticities are heavily biased by variance when a good is limited and may sell out, such as parking spaces.

(Generalised) Linear Models are another simple and viable approach but require extensive feature engineering, as well as compensation for regression towards the mean if one were to claim an effect.

Continuing on these studies, this work is the first to include non-payment, parking durations, relative discount and walking distance into an analysis of parking behaviour. Durations are important to include next to average occupancy, since average occupancy is a derived notion of an arrival-duration process. I provide a start to a general framework of data preprocessing and research methodology of real-time parking data. Preprocessing is particularly important as I have seen different processing steps lead to different qualitative results. Research difficulties include sensor failures, the fact that control groups are non-existent or heavily biased, the leading roles of periodicity and weather conditions in measured results and different types of policy changes being applied simultaneously.

ANALYSIS Within my analysis I find that more drivers have come to park on-street in Los Angeles. While underuse has decreased, parking congestion has gone up. I find no change in the amount of block faces that are just-right. This is partially due to the city's 70-90% metric being inappropriate for small blocks or blocks with few sensors.

In general I find no impact of rate changes on duration or occupancy. In specific extreme cases, such as extension of operating hours, I do see a significant impact of rate changes on occupancy. This suggests rate changes of amounts larger than $2 could help.

I discerned between paid parking and unpaid parking. The rate adjustment algorithm was defined without this group in mind. Reflecting on past rate changes, I see no faulty rate adjustments or rates reverting to their previous value. Although the method could be leaner towards paying parkers in areas with high unpaid usage I see no reason to update the rate adjustment algorithm.

UNPAID PARKING I have confirmed suspicions of earlier work that placard usage is a significant problem for the functioning of demand-based parking in downtown Los Angeles. Unpaid parking may take up as much as 90% of the total time parked on certain high-demand block faces and on average about 45% of the total parked time in operating hours is not paid for.

Half of the on-street parking capacity is lost for paid parking due to unpaid usage. Areas suffering from congestion also suffer from a high share of unpaid parking.

As expected, I find no effect of rate changes on the distribution of unpaid parking on spatial distribution during the project.

I found that removal of no-parking periods attracts long-stay non-paying parkers. Future work will have to determine whether introduction of midday no-parking periods deters non-paying parkers.

I found estimates that the downtown parking revenue can be doubled if the city succeeds in reducing unpaid use to marginal amounts.

DISCOUNT BY WALKING Rates have developed well through the city, in spite of a lacking response in occupancy. New incentives to walk have been introduced and these are placed conveniently based on current parking problems.

RATIONALITY MODEL I describe a simple choice model for choosing a parking space in downtown Los Angeles. This model is arguably the simplest to include non-payment, parking durations, relative rates and relative walking distances, all features that were not included in previous work. I find a large gap in my model predicted incentive and the observed parking behaviour. The model specifically identifies areas where many more incentives were created and observations show drivers do not pursue these incentives. We identify three possible causes for this gap.

First, people really dislike walking. A possible answer is to wait for further rate adjustments. No area has yet reached the rate roof of $\$8/h$. In many areas the rate can be increased further, which may push drivers to search for cheaper alternatives in spite of their dislike of walking.

Second, the model is not complex enough and a more detailed or complex model is required. While the model is simple, the assumptions may go too far in describing the observed data. To overcome the computational complexity, one could try approximative methods such as Expectation Maximisation or MCMC methods to estimate the posterior distributions.

Third, and in my opinion the most likely, is that drivers are not aware of cheaper alternatives. This may be confirmed by additional surveys, adding to the work by Glasnapp et al. (2014). The model predictions show the areas where the most incentives are ignored, making these viable seed locations for an awareness campaign with signs or stickers.

Due to time limits I was unable to include a thorough analysis on the model's parameter sensitivity.

FUTURE WORK I have identified possible causes for the lack of response towards rate changes which require looking into. Surveys would provide a definitive answer to whether drivers are aware of rates. As for choice models, additional driver trip data is required to improve the quality of model predictions. Parties such as TomTom and Google are already in possession of trip data necessary to greatly improve demand-based parking. Access to this information is vital for the successful development and deployment of demand-based parking systems.

I stress the need of a method to establish a causal influence of rates on driver behaviour. Without a clear effect of rates on the behaviour of drivers, cities will not see a reduction of parking congestion.

Calthrop et al. (2000) suggest a combination of imperfect methods from road pricing and imperfect methods from demand-based parking to improve the overall driver experience. There have not yet been studies that look into parking congestion in cities that introduced road pricing or a congestion charge.

Off-street parking garages have not yet been included in studies relating to demand-based parking. I am in possession of off-street rate data, but lacked the time to include an analysis in this work. It would be interesting to see how pricing of garages was influenced by the LA Express project. Have garage operators changed the rates of their facilities based on the on-street rate changes? Since off-street is the main alternative for on-street parking and the off-street capacity is much larger, I wonder whether off-street rates are an approximation of equilibrium market rates. How does off-street capacity relate to on-street congestion?

RECOMMENDATIONS I conclude this work with a number of recommendations for city and parking officials. For the success of demand-based parking it is vital that satellite navigation is included in parking systems as soon as possible.

Further, the handicap placards are a harmful exception to the parking economy and it needs to be changed. Alternatives, such as a budget for parking or a limited number of handicap spaces may provide an intermediate solution.

Lastly, off-street parking garages should be included in further demand-based pricing programmes. Most of cities' downtown parking capacity is within parking garages. I expect parts of the on-street outflow to choose for often cheaper off-street alternatives instead. By including these in a project it will be easier to establish a project's efficacy.

ACKNOWLEDGEMENTS

The process of writing a master thesis is a proof of competence; that the student was able to bring all things learned into practice. This proof of competence was not an easy one and I would like to thank the people who made it possible.

Foremost I would like to thank my colleagues and collaborators at Xerox Research Center Europe. Onno Zoeter and Stephane Clinchant for their vision, help and guidance. Christopher Dance and Tomi Silander for valuable discussion and helpful remarks. Roberto Rodriguez for the indispensable libraries and listening to my frustrations about Matlab. Luis Ulloa for helping with OpenStreetMap. Franck Tison for being a database genius. And of course all my friends who made my time in France an unforgettable one, Dominik, Audi, Diana, Milan, Alexandr, Kunal, Chunyang and Pajolma. Thank you.

I owe a lot to Tom Heskes, who helped me numerous times when I was clueless, whether it be about a subject, specific section or going abroad in the first place.

Important to me are my Dutch friends, Caroline, Wouter, Judith, Martijn. Many thanks for all the coffee, tea and mental support.

My dear Susanne, you were there when I needed it most. Thank you.

Lastly I would like to thank my parents, Paul and Saskia. None of this would have been possible without their unconditional support and trust. Thank you for believing in me from the start.

BIBLIOGRAPHY

ABClocal (2012). Fraudulent use of handicapped placards on the rise in LA.

Barnett, A. G., van der Pols, J. C., and Dobson, A. J. (2005). Regression to the mean: what it is and how to deal with it. *International Journal of Epidemiology*, 34(1):215–220.

Beardsley, E. (2010). French city implements new public parking spot sensors.

Ben-Akiva, M. and Bierlaire, M. (1999). Discrete choice methods and their applications to short term travel decisions. In *Handbook of transportation science*, pages 5–33. Springer.

Berishvili, N. and Novosti, M. (2012). November parking revolution.

Bishop, C. M. et al. (2006). *Pattern recognition and machine learning*, volume 4. springer New York.

California Department of Motor Vehicles (2005). California vehicle code V C Section 22511.5 Disabled parking Authorized Parking Zones.

Calthrop, E., Proost, S., and Van Dender, K. (2000). Parking policies and road pricing. *Urban studies*, 37(1):63–76.

Chapman-Smith, B. (2013). Parking sensor trial launched on busy akl road.

Chatman, D. G. and Manville, M. (2014). Theory versus implementation in congestion-priced parking: An evaluation of SFpark, 2011–2012. *Research in Transportation Economics*, 44:52–60.

Clinchant, S., Dance, C., De Ruijter, T., Ghent, P., and Zoeter, O. (2015). Using analytics to understand on-street parking; the impact of special permit use and the benefit of demand based rates over zones. In *Proceedings of 22nd ITS World Congress*.

Cunningham, W. (2013). Audi tells you where to park it.

Ghent, P., Mitchell, D., and Sedadi, A. (2014). LA Express Park™-Curbing Downtown Congestion through Intelligent Parking Management. In *19th ITS World Congress*.

Glasnapp, J., Du, H., Dance, C., Clinchant, S., Pudlin, A., Mitchell, D., and Zoeter, O. (2014). Understanding dynamic pricing for parking in Los Angeles: Survey and ethnographic results. In *HCI in Business*, pages 316–327. Springer.

Goldstein, D. (2013). CBS2 Investigation Looks Into Use Of Disabled Parking Placards.

LADOT (2013). Anual report of the fiscal year 2013-2014.

Luxen, D. and Vetter, C. (2011). Real-time routing with openstreetmap data. In *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, GIS '11, pages 513–516, New York, NY, USA. ACM.

Manville, M. and Williams, J. A. (2012). The price doesn't matter if you don't have to pay legal exemptions and market-priced parking. *Journal of Planning Education and Research*, 32(3):289–304.

Marshall, A., Marshall, A., Marshall, A., and Marshall, A. (1920). Principles of economics: an introductory volume.

McFadden, D. et al. (1978). *Modelling the choice of residential location*. Institute of Transportation Studies, University of California.

Millard-Ball, A., Weinberger, R., and Hampshire, J. (2013a). Is the curb 80% full or 20% empty? Assessing the efficacy of San Francisco's parking experiment. In *Annual Meeting of the Transportation Research Board*.

Millard-Ball, A., Weinberger, R., and Hampshire, R. (2013b). Comment on pierce and shoup: Evaluating the impacts of performance-based parking. *Journal of the American Planning Association*, 79(4):330–336.

Pierce, G. and Shoup, D. (2013a). Getting the prices right: an evaluation of pricing parking by demand in San Francisco. *Journal of the American Planning Association*, 79(1):67–81.

Pierce, G. and Shoup, D. (2013b). Response to Millard-Ball et al.: Parking Prices and Parking Occupancy in San Francisco. *Journal of the American Planning Association*, 79(4):336–339.

Pierce, G. and Shoup, D. (2013c). SFpark: Pricing Parking by Demand. *ACCESS Magazine*.

Ross, V. (2011). Smart parking systems steer drivers to open spaces.

Shoup, D. (2011). Ending the abuse of disabled parking placards. *Access*, (39).

Shoup, D. C., Association, A. P., et al. (2011). *The high cost of free parking: updated edition*, volume 206. Planners Press Chicago.

Vickrey, W. (1954). The economizing of curb parking space. *Traffic Engineering*, 29(1):62–67.

Zoeter, O., Dance, C., Clinchant, S., and Andreoli, J.-M. (2014). New algorithms for parking demand management and a city-scale deployment. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1819–1828. ACM.

Zoeter, O., Dance, C. R., Grbovic, M., Guo, S., and Bouchard, G. (2012). A general noise resolution model for parking occupancy sensors. In *19th ITS World Congress*.