Incentives on the Starting Grid in Formula One Racing

Laurence M. McCarthy
Stillman School of Business
Seton Hall University

Kurt W. Rotthoff[1]
Stillman School of Business
Seton Hall University

Spring 2013

Abstract
Starting a race in first place, pole position, is the goal of every race driver. This is even more pronounced in Formula One (F1) racing as the road courses they race are more difficult to pass on, providing an additional advantage to starting on the pole. However, their unique standing starts also create a bottleneck at the first turn, which often leads to contact between cars. Because F1 cars are not designed to make contact, this contact can greatly impact a driver's position on the track. We find that there are certain positions on the starting grid that are more likely to make contact with other drivers than other positions. Specifically the starting position with the highest odds to make contact at the first turn is position 10. This creates the incentive for drivers to avoid this position, which means if they are unable to qualify higher than this position, the incentive exists for drivers to intentionally adjust their behavior to avoid these high-risk (of making contact) positions.

[1] Kurt W. Rotthoff can be contacted at rotthoff@gmail.com or Kurt.Rotthoff@shu.edu; Seton Hall University, JH 674, 400 South Orange Ave, South Orange, NJ 07079. Laurence M McCarthy at Laurence.McCarthy@shu.edu. A special thanks to Andrew J Moreno for his data work as our research assistant. Any mistakes are our own.

I. Introduction

In Formula One (F1) car racing a defined procedure for the assignment of starting positions is used. Mandatory qualifying sessions lead to the fastest car/driver being awarded the starting position one ("pole position") for the race. According to the procedure to reward top qualifying competitors with the best starting positions, it is assumed that performance outcome in formula one car races would benefit (Muehlbauer, 2010).

Although qualifying first or in pole position is the goal for every driver, it is not clear that starting position in the rest of the field follows a linear relationship. Given that F1 races are a standing start, all cars are stopped and the race starts on a light, the cars that qualify further back in the field have a longer distance to the first corner which allows for more acceleration into the turn. This results in drivers further down the grid entering the first corner's breaking zone at a higher speed. This creates a 'Concertina Effect' going into the first corner. It is this effect that leads to contact between cars, and ultimately damage to the cars themselves. F1 cars are not able to be driven effectively, or able to perform as well, with damage; thus this early race contact significantly impacts the entire race for a driver.

Making it through the first corner unscathed is critical to success in a race. As such starting positions can be decisive. We analyze which starting positions are most likely to be involved in first corner accidents. Logically the pole-position holder will be ahead of any wreck, and those starting later in the field will have the chance to avoid the contact. Thus, if there is a certain group of starting positions that is more likely to be involved in a wreck, the optimal strategy of qualifying can be altered. It is possible that the better strategy is to intentionally move back in the starting grid in order to avoid the first turn contact. This information is valuable for both the F1 teams and the Federation Internationale de l' Automobile (FIA, the governing body of F1 racing).

Muehlabauer (2010) examines the relationship between the starting position and finishing position in 70 F1 races. He finds that the correlation between the starting position and finishing position is significant. He finds that there is a correlations coefficient of 0.63; we get similar result for the 2012 season used in our study. This is further confirmed in a working paper by Silva and Silva (2010), where they find that the qualifying performance of a driver was the best predictor of their finishing position in F1 racing. Allender (2008) focused on the National Association of Stock Car Racing (NASCAR), finding that a driver's years of experience plays a significant role in predicting the outcome of NASCAR races. These studies look at how the starting positions of these drivers impact their finishing positions. We take a step back; in order to finish the race you have to make it through the first turn (preferably without damaging your car), especially in F1 where first turn wrecks are common (Benson, 2012).

Optimal decisions in racing have been looked at by Bekker and Lotz (2009) and the tournament structure on between race incentives in NASCAR by von Allmen (2001). We look at optimal qualifying positions by analyzing what positions on the starting grid are most likely to wreck on the first turn. To do this we get data from the 2012 F1 racing season and analyze the probability of wrecking on the starting position. The data suggests that a driver wants to avoid starting at, or near, the tenth position on the grid. Those cars starting tenth, and those cars starting closest to this position, are more likely to make contact in the first turn. The next section describes our data and the methodology. Section three discusses the results and the last section concludes.

II. Data and Methodology

We analyze all 20 Formula One Grand Prix races organized by the Fédération International (FIA) in 2012. We have data on the starting position, finishing position, and contact information, from major wrecks to minor contact, for each driver for each race during this season.[2] The starting and finishing positions were matched to race reports on the first turn accident data. The data were gathered from: formula1.com, ESPNF1.com, bbc.co.uk/sport/0/formula1, grandprix.com, dailymail.co.uk/sport/formulaone, f1fanatic.co.uk, motorsportpress.wordpress.com, telegraph.co.uk, racedepartment.com, paultan.org, edp24.co.uk, and vitalf1.com.

To measure the probability of a given driver, *i*, in a certain starting position making contact in the first turn, we use a probit estimation.

$$Contact_i = \beta_1 (Starting\ Position)_i + \varepsilon \tag{1}$$

The estimation, equation 1, allows us to analyze the odds that a driver makes contact in the first turn based on that drivers starting position. To control for any non-linearities in the data we also estimate equation 2 with a squared term.

$$Contact_i = \beta_1 (Starting\ Position)_i + \beta_2 (Starting\ Position)_i^2 + \varepsilon \tag{2}$$

Equation 2's functional form allows for the relationship between the odds of making contact and starting position to not be linear. We expect this curve-linear specification to be more accurate given that there are two positions on the starting grid that have the ability to avoid contact in the first turn: the pole position and those starting

[2] We attempted to get data for 2010 and 2011, but no accurate data on contacts and wrecks in the first corner were found for those years. Thus the sample is restricted to 2012, where accurate data on first turn contact could be verified.

last, because they have time to react to contact. These estimations are done for the full sample, only races that have some contact, and races that have major contact (more than four cars making contact). These different samples allow us to measure any difference across tracks that could be known for first turn contact.

In addition to the models specified above, we add a robustness check by creating a dummy variable for each starting position. The driver that starts on the pole is position 1, P1, the second qualifier is P2, and so on. This estimation will reveal is there is a relative difference in driver's probability of wrecking on the first turn, relative to the omitted variable: P1. Results for these three specifications are discussed in the next section.

III. Results

The results for each estimation are done on each of the three samples: the full sample, the start of every race throughout the entire 2012 season, and two restricted samples. The first restricted sample looks at the odds of making contact only in races where contact is made in the first corner (which occurs in 14 of the 20 races). The second restricted sample analyzes the impact of races where major contact was made in the first corner. We define major contact as four or more cars making contact, which occurred in five of the 20 races.

Table 1 presents the results from equation 1 on how the starting position changes the odds of making contact in the first turn. For the full sample, column 1, the races where contact was made, column 2, and the sample with major contact, column 3, we find a negative and significant relationship. This shows that the further back a driver is in the grid, i.e. qualifying lower, decreases the odds of making contact in the first turn.

Table 1: Probit estimation on the odds of making contact based off the drivers starting position. Marginal effects reported for the full sample, limited sample to races with contact, and limited sample to races with a major contact.

	Full Contact	Restricted Contact	Restricted Major Contact
Starting Position	-0.004794**	-0.006874**	-0.013988**
	(0.002)	(0.003)	(0.006)
Observations	477	334	118
Pseudo R^2	0.0158	0.0190	0.0414

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

However, as we suggested earlier, we believe that there are two optimal positions: first (on the pole) and going last, because the driver has time to avoid contact. For this reason we analyze a curve-linear relationship between the staring position and odds of making contact in table 2.

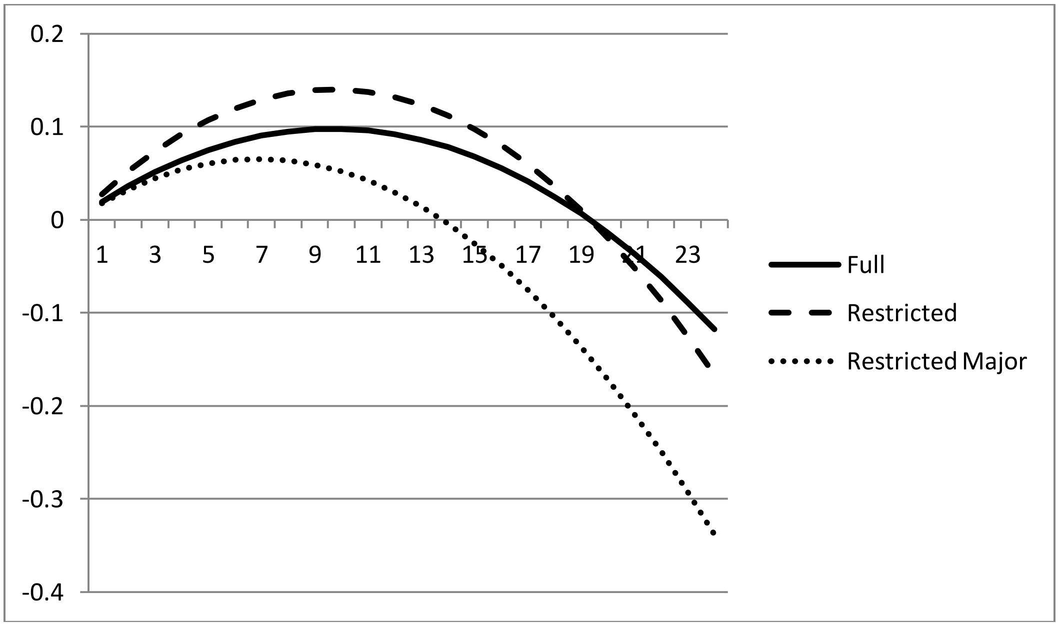
Table 2: Probit estimation on the odds of making contact based off the drivers starting position. Marginal effects reported for the full sample, limited sample to races with contact, and limited sample to races with a major contact.

	Full Contact	Restricted Contact	Restricted Major Contact
Starting Position	0.020212**	0.028901**	0.018988
	(0.009)	(0.012)	(0.024)
Starting Position ^{2}	-0.001046***	-0.001493***	-0.001378
	(0.000)	(0.000)	(0.001)
Observations	477	334	118
Pseudo R^2	0.0433	0.0516	0.0561

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

In table 2 we find that the relationship between starting position and starting position squared is increasing at a decreasing rate. Thus, it is better to start either first or last, but not in the middle. This result holds true for all 3 samples, the full sample and the two restricted samples. However, the peak (the highest odds of making contact on the first turn) changes across the different samples, displayed in figure 1.

Figure 1: The estimated odds of making contact in the first turn, based on the starting position of the driver.



The peaks, which represent the highest odds of making contact in the first turn, exist at starting position 10 for both the full sample and the sample of races where contact is made. The full sample has lower odds of making contact, which is expected since it includes races without any contact. On the other hand, if a track is well known for contact in the first turn, then looking at the sample restricted to contact would reveal more realistic odds of the event occurring. As can be seen, those starting after the 20th position have lower odds of making contact than the pole sitter.

When major contact is made, the peak occurs with P7. The curve also goes negative, again relative to P1, quicker; showing that when major contact is made it impacts the top 13 drivers the most.

Robustness

Table 3: Probit estimation on the odds of making contact based off the drivers starting position. Marginal effects reported for the full sample, limited sample to races with contact, and limited sample to races with a major contact.

VARIABLES	Full contact	Wreck contact	Major Wreck Contact
P2	0.948135***	0.921746***	0.851665***
	(0.012)	(0.017)	(0.039)
P3	0.948135***	0.921746***	
	(0.012)	(0.017)	
P4	0.951087***	0.925535***	0.858156***
	(0.011)	(0.016)	(0.038)
P5	0.949899***	0.923883***	0.851665***
	(0.011)	(0.016)	(0.039)
P6	0.951087***	0.925535***	0.858156***
	(0.011)	(0.016)	(0.038)
P7	0.951087***	0.925535***	0.865446***
	(0.011)	(0.016)	(0.036)
P8	0.944396***	0.918176***	0.843875***
	(0.009)	(0.018)	(0.040)
P9	0.951087***	0.925535***	0.851665***
	(0.011)	(0.016)	(0.039)
P10	0.949899***	0.923883***	0.851665***
	(0.011)	(0.016)	(0.039)
P11	0.951087***	0.925535***	0.843875***
	(0.011)	(0.016)	(0.040)
P12	0.949899***	0.923883***	0.843875***
	(0.011)	(0.016)	(0.040)
P13	0.951087***	0.925535***	0.851665***
	(0.011)	(0.016)	(0.039)
P14	0.952018***	0.926953***	0.851665***
	(0.011)	(0.016)	(0.039)
P16	0.944396***	0.918176***	0.843875***
	(0.014)	(0.018)	(0.040)
P18	0.949899***	0.923883***	0.843875***
	(0.011)	(0.016)	(0.040)
P19	0.948135***	0.921746***	
	(0.012)	(0.017)	
P20	0.944396***	0.918176***	0.843875***
	(0.014)	(0.018)	(0.040)
P22	0.944396***	0.918176***	0.843875***
	(0.014)	(0.015)	(0.036)
P24	0.942759***	0.916124***	0.831175***
	(0.013)	(0.018)	(0.042)
Observations	398	279	89
Pseudo R^2	0.0643	0.0786	0.1352

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

As a robustness check we looked at each starting position's odds of making contact individually. In table 3 we run the probit regression and find that all positions are more likely to make contact than the omitted variable, P1. There are four starting positions that make no contact in our sample, P15, P17, P21, and P23, and thus they do not appear in table 3. In all samples measured, we also find that the odds of making contact fall off for those drivers that qualify lower in the grid.

IV. Conclusion

Although the FIA would like every driver to be trying to qualify in the highest possible position during their qualifying runs, we reveal information that could impact a driver's optimal choice during this qualification process. There is often contact between drivers in the first corner of races and as such drivers have to decide if moving back on the starting grid, to decrease their odds of making contact with other drivers, is the optimal decision. The driver who starts 10th on the grid has the highest odds of making contact with another car. The lowest odds of making contact occur for those drivers starting in the last five positions, or starting on the pole. For courses that are notorious for first turn wreckage, these odds should be taken into account for the driver (after they acknowledge they will not get the pole at that race).

Works cited

Allender, Mary (2008) Predicting The Outcome Of NASCAR Races: The Role Of Driver Experience *Journal of Business and Economics Research* Volume 6, Number 3

Bekker, J. and W. Lotz (2009) Planning Formula One Race Strategies using Discrete-Event Simulation *Journal of the Operational Research Society* 60, 952–961

Benson, A. (2012) Belgian Grand Prix crash: Romain Grosjean banned for one race. Retrieved 3/8/2013from http://www.bbc.co.uk/sport/0/formula1/19458954

Muehbauer, Thomas (2010) Relationship between Starting and Finishing Position in Formula One Car Races *International Journal of Performance Analysis of Sport* 10, 98-102

Silva, Kathleen M. and Francisco J. Silva (2012) A Tale of Two Motorsports: A Graphical-Statistical Analysis of How Practice, Qualifying, and Past Success Relate to Finish Position in NASCAR and Formula One Racing Working paper (access 3/1/13) Retrieved from http://newton.uor.edu/FacultyFolder/Silva/NASCARvF1.pdf

von Allmen, P. (2000) Is the Reward System in NASCAR Efficient? *Journal of Sports Economics* Vol. 2, No. 1, pp. 62-79.