STOCKHOLM SCHOOL OF ECONOMICS
Department of Economics
5350 Master's thesis in economics
Academic year 2016-2017

Labour Supply Decisions of Stockholm Taxicab Drivers

Joel Wållgren
(22218)

ABSTRACT: In this paper, panel data on hours worked and wage rates of taxicab drivers in the city of Stockholm are used to analyse whether reference-dependent preferences play a significant role governing labour supply decisions. The results suggest that only a small fraction of wage variation (about 1/5) is unanticipated implying that reference-dependence (which is relevant only in response to unanticipated variation in the wage rate) plays a limited role determining labour supply decisions. The result is confirmed by applying the discrete-choice stopping model (Farber, 2005) and the findings imply that the probability of ending a shift is positively related to accumulated hours (conditional on accumulated income) and seemingly unrelated to accumulated income (conditional on accumulated hours). This is inconsistent with the model of preferences dependent on a reference income level. The paper finds no evidence of a dual target suggested by Crawford and Meng (2011). In addition, heterogeneous behaviour across drivers' labour supply decisions is found. One possible factor explaining the differences is that driving a taxicab may be a "learning by doing"-process (LBD). Heterogeneous behaviour across drivers' labour supply decisions also highlights the importance of data quality and it could be argued that some of the spurious results found by Camerer et al. (1997) and Chou (2002) can be explained by a non-random sample selection. The findings of this paper are generally consistent with Farber's (2015) results. Conclusively, extending the time horizon, some evidence for a longer planning horizon is found and precise wage elasticity estimates are positive and ranging from 0.20-0.35.

Keywords: Labour supply, Reference-dependence theory, Neoclassical theory, Taxicab drivers

JEL Classification: D02, D03, J22

Supervisor: Erik Lindqvist
Date submitted: 12th of December
Date examined: 20th of December
Discussant: Gustav Fredriksson
Examiner: Maria Perrotta Berlin

Acknowledgements

I would like to thank my supervisors Michael Lechner, at the University of St. Gallen, and Erik Lindqvist, at the Stockholm School of Economics, for their guidance during the writing process. I would also like to thank Juanna Schrøter Joensen, at the University of Chicago, for working with me for all these years and for her encouragement of my eccentric fascination for taxicab drivers. Finally, I would like to thank all those who have borne with me during this project. Thank you.

Table of Contents

Table of Tables

Table of Figures

1. Introduction

The most important factor of production in conventional macro models is labour. The standard estimate is that two thirds of all production revenue is dedicated to labour compensation, which implies that labour costs represent the largest share of total production cost. Acknowledging the importance of the labour share of total production cost, scholars have attempted to understand the dynamics of the compensation process and how income affects the labour supply of agents. This paper presents two different models of labour supply. The neoclassical labour supply model implies that hours worked should be positively related to transitory fluctuations in wage. The model of reference-dependence predicts the opposite relationship with transitory wage fluctuations. The question answered in this paper is which one of these models is the dominant explanatory model of individual labour supply.

Assuming that the income effect is negligible, the prediction of life-cycle models of labour supply is that there is a positive relationship between hours supplied and transitory wage fluctuations. The hypothesis is that workers intertemporally substitute consumption and leisure, supplying a higher quantity of hours when the alternative cost of leisure raises due to a transitory wage increase (Lucas and Rapping, 1969). The prediction is proved to be hard to verify empirically. Utilising aggregate (Mankiw, Rotemberg, and Summars, 1985), cohort (Browning, Deaton, and Irish, 1985) or panel (Altonji, 1986) data, scholars find the estimates of wage elasticity to be low and insignificant. The estimates of the wage elasticity also diverge in the micro and macro dimension respectively (see Chetty at al., 2011 for a meta-study of wage elasticity estimates).

The labour supply decisions of taxicab drivers have received much attention in recent years. Taxicab drivers are examined because of two features: they are exposed to temporary wage fluctuations and they freely decide their own working hours. This enables a comparison between the two labour supply models in order to decide which one contributes the strongest explanatory value of the labour supply decision (Barberis, 2013). In a seminal paper Camerer, Babcock, Loewenstein and Thaler (1997) estimate negative wage elasticity in their data sets that include the hours worked and income earned of taxicab drivers in New York City (NYC), and are supported by findings in Singapore (Chou 2002). The findings contrast the projection of the standard neoclassical model of labour supply, and suggests that drivers end their shifts earlier on days when they receive high earnings, whereas they continue their shifts longer during days when earning opportunities are scarce. The literature suggests that the negative elasticity is consistent with a labour supply model where agents are reference-dependent and seeking to reach an income target (originally articulated by Kahneman and Tversky (1979; 1991)). However, the econometric methodology employed by Camerer et al. (1997) may induce downward bias in the wage elasticity estimates (Stafford 2013; Farber 2005; Oettinger 1999). Thus, negative wage elasticity estimates may be a consequence of econometric issues, rather than a reflection of the drivers' true behaviours. Nevertheless, the debate continues and Agarwal et al. (2015) find estimates of labour supply models for taxicab drivers with negative slopes. Using different approaches, Crawford and Meng (2011) and Doran (2014) also find support for reference-dependent preferences when analysing the labour supply of NYC taxicab drivers.

In light of these developments, Farber (2015) gathers an extensive data set of NYC taxicab drivers and applies the model of expectations-based reference points presented by Köszegi and Rabin (2006; 2007; 2009). Farber (2015) explores the question of taxicab drivers labour supply decision from three different angles. First, he distinguishes between anticipated and unanticipated daily wage variation since reference dependence and neoclassical theory predict the same line of behaviour in regards to anticipated wage variation. Second, he uses a discrete-choice stopping model, i.e. the driver makes the decision whether to keep working or ending the shift after each completed fare, to analyse if accumulated income or hours worked during the shift are correlated to the discrete intra-day labour supply decision. Third, he investigates the heterogeneity across drivers in their driving behaviours and the potential of a learning process, with more experienced drivers having more positive wage elasticities. Farber (2015) finds that unanticipated daily wage variance is only a small fraction of the overall daily wage variance and that workers labour supply is positively correlated with both anticipated and unanticipated wage variations. Moreover, the probability of ending a shift is positively correlated to the number of hours worked (conditional on accumulated income) and is seemingly unrelated to accumulated income (conditional on accumulated hours worked). He also finds that taxicab drivers differ significantly in their driving behaviours, providing evidence that more experienced drivers have more positive wage elasticities. Farber (2015) concludes that reference-dependence plays a limited role in determining labour supply.

Farber (2015) presents two weaknesses in his paper, which simply can be summarized as the institutional settings of the NYC market. First, the taxicab drivers in NYC work in shifts and are not completely free to set their own working hours. Second, tipping accounts for a substantial part of the driver's wage, which is not included in his data (see Haggag et al. (2014) and their finding that a substantial part of NYC taxicab drivers' income consists of tips). These shortcomings raise the question whether Farber's (2015) result that reference-dependence only plays a limited role in the labour supply decision of taxicab drivers reflects the true behaviour of the drivers, or if the results are driven by the institutional setting. Furthermore, Farber (2015) does not have reliable data on tenure, i.e. how long the driver has worked as an authorized taxicab driver and instead the time each individual driver can be observed within the sample period is used as a proxy for experience, leaving the question of experience-based heterogeneity across drivers' behaviours open for further investigation.

The aim of this paper is to replicate Farber (2015) by applying his methodology in a setting free from the institutional constraints associated with the NYC taxicab market and with actual data on experience to assess the possibility of a learning process and consequently drivers' behaviours heterogeneity. This paper is also a development of my previous work (Jonasson and Wållgren (2013)) in which we found that the driving behaviours of the taxicab drivers in Stockholm are inconsistent with a one-day income target theory of labour supply (analysing the wage elasticity). Our findings are in line with Farber (2015), but are reached in a much less sophisticated way. The extension is two folded. First, the Farber (2015) paper had not been published when I wrote my previous paper and thus the empirical implementation of unanticipated wage variation was not reviewed. Second, my new assembled data set includes trip-by-trip information and tenure. This enables analysing the discrete-choice stopping model (requiring specific information of each fare of each driver) and a potential learning behaviour (requiring specific information of experience of individual drivers). Therefore I am able to replicate Farber's (2015) methodology and test the external validity of his paper and to some extent the external validity of the cumulative evidence presented on the driving behaviour of NYC taxicab drivers. The research question is formulated as follows:

Does reference-dependence play a significant role in determining the labour supply decision of Stockholm city taxicab drivers?

I handle the weaknesses of Farber's (2015) paper by using newly gathered data with three main advantages. First, unlike the NYC taxicab drivers who are constrained in their labour supply decisions because of shifts lengths, the drivers in Stockholm can freely choose when to work. In my data set all drivers have their own car available 24 hours a day. Second, the data set has minimal measurement error as it derives directly from electronically gathered data that disaggregate the tipping, assuring that all income is taken into account when conducting the analysis. Third, I observe each driver's unique identity over a period of three months with distinct data on tenure and how long the driver has worked as an authorized taxicab driver, which allows this paper to investigate a potential learning process.

In addition to the problems mentioned above, there is also an issue of the time aspect of the labour supply decisions of taxicab drivers. Even though Farber's (2015) latest contribution incorporates a temporal aspect of the labour supply decision, one could argue that it is a very narrow timeframe (one day). Farber's (2015) anticipated versus unanticipated wage variation allows drivers in the model to make daily decisions at the intensive margin but never to look ahead. This view may seem myopic since drivers can arguably have longer time horizons. One way to approach this issue is to apply the discrete-choice stopping model with a longer time frame. To get a more precise estimate of wage elasticity I also compute the wage elasticity for different time frames.

The results provide general support for the neoclassical model of labour supply. The unanticipated wage variation only accounts for 1/5 of the total wage variation implying that reference-dependence plays a limited role determining the labour supply decisions of Stockholm taxicab drivers. The results are in line with the replication of the discrete-choice stopping model (Farber, 2015). The probability of ending a shift after a given fare is positively related to hours worked during the shift (conditional on accumulated income) and seems to be unrelated to the income level (conditional on accumulated hours). Furthermore, I find no evidence of a dual target as suggested by Crawford and Meng (2011). However, I find heterogeneous behaviours across drivers' labour supply decisions and an indication of a potential learning curve. But first and foremost, heterogeneous behaviour among taxicab drivers emphasizes the importance of having an extensive data set in order to make plausible evaluations of the behaviours of taxicab drivers. Lastly, when extending the time horizon, some evidence for a longer planning horizon is found which emphasises the need of looking at the labour supply decisions of taxicab drivers in a more dynamic way than in previous research, including Farber (2015). Precise wage elasticity estimates are positive and range from 0.20-0.35.

The paper is organized as follows: Section 2 covers the previous research on labour supply decisions. Section 3 describes the Stockholm taxicab data collected for the study, while section 4 contains a review of the conceptual framework of Farber's (2015) paper and the empirical implementation. Thereafter, section 5 presents the empirical results, which are then discussed in section 6 where findings are put into context. Lastly, section 7 presents concluding remarks.

2. Previous Research

This section defines the concept of wage elasticity and I present previous research relating to the taxicab market, a theoretical evaluation of the reference-dependent theory, and evidence from various markets. Lastly, LBD behaviour is defined and I articulate the relevance of the research question.

Wage Elasticity

The measurement of interest is wage elasticity, which is conventionally separated into two distinct terms. First, the Frisch (marginal utility constant) elasticity reflects intertemporal substitution response to transitory wage fluctuation. Second, the Hicksian (wealth constant) elasticity reflects steady-state responses and the welfare cost of taxation (Chetty et al., 2011). Conventionally, when examining taxicab drivers' labour supply decisions, the Frisch wage elasticity is used.

Derived from theory the relation of the Frisch wage elasticity and wage fluctuation is the following: if wage fluctuations are purely transitory, it is reasonable to assume that the income effect is negligible. Prediction of the neoclassical model of labour supply prescribes the existence of a positive relationship between hours supplied and transitory wage fluctuation. The relationship between working hours supplied and wage fluctuations has been hard to verify empirically (Camerer et al., 1997, and see introduction).

Potential problems with the early studies of labour supply are that wage changes are seldom purely transitory causing serial correlation in wage rate. Hence the neoclassical theory is jointly tested together with agents' expectations regarding the wage shock persistence and future wage (Camerer et al., 1997). In addition, since wages are often positively serially correlated and changes are often non-transitory the assumption of a negligible income effect is not reasonable. This affects the prediction of the labour supply response of the neoclassical model and consequently the result will be more difficult to interpret. Indeed, if the wage increase is permanent, empirical observations suggest unresponsiveness of the labour supply. A possible explanation is that the income and substitution effects cancel each other out (Kimball and Shapiro, 2008). Furthermore the neoclassical model assumes that workers are able to react to transitory wage fluctuations, although this is rarely the case since the vast majority of workers have fixed working hours (Farber, 2005; 2008; 2015). The presented caveats may bias the wage elasticity toward zero (Keane, 2011).

Starting with Camerer et al. (1997), several empirical studies are trying to mitigate some of the inherent problems with earlier research by analysing occupations where the workers are free to choose their own working hours and in addition are exposed to transitory wage fluctuations. Taxicab drivers are recognized as such a group (Camerer et al., 1997; Barberis, 2013). Camerer et al. (1997) regress daily log hours worked on the daily log mean wage rate. They incorporate adverse weather condition control for demand fluctuation and use driver fixed effect to control for time-invariant heterogeneity across drivers. The authors' result indicates statistically significant negative estimates of wage elasticity (IV estimates around -0.5 in one of their samples)[1]. Their findings are inconsistent with the neoclassical theory of labour supply since drivers respond with negative temporal substitution to temporary wage shocks. Individuals seem to supply less hours as the wage temporarily increases. To address the concern of measurement errors in reported income, the authors suggest using an instrumental variable (IV) approach and propose that a valid instrument is other

[1] Three data set used are 1) TRIP – 70 shifts for 13 drivers from 20 days in 1994, 2) TLC1 - 1044 shifts for 484 drivers from six days 1990 and 3) TLC2 – 712 shifts for 712 drivers from two days in 1988).

drivers' recorded wage (driving the same day). The measurement error may induce downward bias since the explanatory variable, the hourly wage, is derived from the ratio between daily income to hours worked (Camerer et al., 1997). Their estimates of wage elasticity are statistically significant with a value around -0.5. The authors conclude that that drivers set loosely defined income targets one day at a time, and when the target is reached drivers stop working.

The authors argue that the result is in line with prospect theory, a model of risk preferences first articulated by Kahneman and Tversky (1979, 1991) that combine a reference point and the notion of loss aversion. Reference-dependence is the basic idea that agents draw utility from gains and losses and where the reference point is some target level. Around this point losses are weighted more heavily than gains. Thus, this implies a "kinked" preference curve around the target. This is in contrast to the neoclassical theory, where utility is derived from the absolute level of wealth (see Chou (2002) and Jonasson and Wållgren (2013) for mathematical intuition).

In order to make reference-dependence a plausible explanation of their finding of negative wage elasticity, Camerer et al. (1997) need to prove that the wages of taxicab drivers are not serially correlated across days. The reason is that negative estimates of wage elasticity can be obtained even if the drivers follow the neoclassical model but with a longer timeframe. For example, a driver may stop working either because he has reached an income target or because he anticipates higher wage tomorrow. Camerer et al. (1997) argues that a short time horizon is consistent with research examining the concept of narrow bracketing (Read et al., 1999).

Chou (2002) conducts a replication exercise of Camerer et al. (1997) in Singapore obtaining drivers' self-reported daily income and hours worked (following around 20 drivers for 5 consecutive days). Consistent with Camerer et al. (1997), the results identify significant negative wage elasticity estimates in various specifications, (OLS: -0.40, IV: -0.56, and FE: -0.51).

A practical impediment of Camerer et al., (1997) and Chou's (2002) findings is that their limited data may influence their assessment of a one-day earnings target. The data used does not consist of balanced panels and the time period used does not allow examining the temporal aspect of the wage fluctuation. Another aspect is that the proposed instrument must be technically difficult to obtain, since in their samples there are very few drivers driving on the same day (especially the data set in Camerer et al. (1997) called TRIP).

Finding a Target

A behavioural reason why drivers may choose a daily income target is to mitigate the self-control problem. Since drivers are free to choose their working hours and quit whenever they want, an income target (reference level) would constrain the possibility to simply give up and end the shift too early. Evidence supporting various aspects of prospect theory is found in different fields within decision-making (see Camerer (2000) for survey).

Barberis (2013) acknowledges that not articulating the reference level is one of the shortcomings of the work of Kahneman and Tversky (1979; 1991). Köszegi and Rabin's (2006; 2007; 2009) work suggests that drivers' reference income levels depend on drivers' expectations. Their model allows drivers to draw utility from absolute levels of income and hours worked, but also to take into account differences between actual and expected daily income as well as the difference between actual and expected daily number of hours worked. The findings of Köszegi and Rabin (2006; 2007; 2009) question the validity of the methodology and findings of Camerer et al. (1997) and Chou (2002) in multiple dimensions.

First, the implication of defining a reference point based on expectations has theoretical and empirical consequences, limiting the potential scope and changing the empirical prediction of the reference-dependent theory. The model can only explain negative wage elasticity estimates to be consistent with reference-dependence theory if an agent is exposed to unanticipated variation in the wage rate. When agents react to anticipated wage variation, i.e. when wages are expected to be high, individuals will adjust their reference points upward implying higher labour supply and hence a positive wage elasticity. Thus, reference-dependent preferences prescribe the same behavioural response as the neoclassical model and agents will supply more hours given an anticipated transitory wage increase. Thus, reference-dependence is only a relevant explanation to negative wage elasticity estimates when agents react to unanticipated wage variations (Köszegi and Rabin, 2006).

Second, reference-dependence is a local phenomenon, so to test the validity of the prediction of a wage elasticity of -1, the actual wage must be relatively close to the expected wage. Thus, the theory of expectation-based reference-dependence has verifiable predictions: on days when the wage rate substantially varies from the expected rate, labour supply is likely to vary with wage (Farber, 2015). Consequently, if the wage rate is close to the expected value, drivers are likely to react inversely to the unanticipated wage changes. Experimental trial of the theory of variation in effort levels supports the notion of a reference point based on expectations (Abeler et al., 2011).

The model articulates a tension when testing the validity of the reference-dependence theory empirically: the labour supply variation is explained by the theory only if a large fraction of the total wage variation is unanticipated. However, when examining the reaction to unanticipated wage fluctuation the actual wage needs to be close to the expected wage and only in this region of wages one should expect negative estimates of wage elasticity if the reference-dependence theory is valid (Farber, 2015). See mathematical derivation in *Appendix 2*.

The work of Köszegi and Rabin (2006; 2007; 2009) gives theoretical grounding to an expectation-based reference point, has verifiable empirical predications and questions the empirical implementation of Camerer et al. (1997). However, the findings of Camerer et al. (1997) received great attention and the basic set-up of their study is replicated in various markets.

Experience From Other Field

As previously mentioned, the findings of Camerer et al., (1997) and Chou (2002) inspire numerous studies to investigate labour supply response to transitory wage fluctuations in other markets where workers completely can decide their own working hours.

Oettinger (1999) finds significant positive estimates of wage elasticity when conducting a field study analysing the extensive margin of baseball match vendors. Fehr and Goette (2007) conduct a field experiment varying the wage rate paid to bicycle messengers. The wage rate change is through the variation of the piece rate given to the workers, allowing the authors to analyse the effort decision. The findings suggest that the labour decision is in line with the neoclassical model reflecting that messengers work more during months with high piece rates. The evidence presented in the paper is mixed in one aspect, messengers work fewer hours per day during the days with the relatively high piece rates. This can be seen as evidence supporting that reference-dependent preferences still influence the labour supply decision. Farber (2005; 2015) argues that this is not necessarily inconsistent with the neoclassical model of labour supply. Reducing hours worked can be optimal if taking into account the effort level of more intense work. The evidence from bicycle messengers is not coherent. For example, a study in Kenya finds strong evidence of reference-dependence (Dupas and Robertson, 2016).

Stafford (2013) argues that the estimates of negative wage elasticity in the original study by Camerer et al. (1997) could be suffering from bias. One of the determining factors of obtaining precise estimates is the access to a complete panel data. The lack of such did not allow Camerer et al. (1997) to control for self-selection in participation. If daily wage fluctuations affect both hours worked and the participation probability in the same direction, it may induce negative correlation between the wage and the error term in the hours equation and negatively bias the wage elasticity estimates (Stafford 2013). Conducting an analysis of daily labour supply variation of lobster fishermen in Florida, Stafford (2013) estimates a small but significant (Frisch) wage elasticity (0.07) and in addition a substantially higher (participation) elasticity at the extensive margin (covering estimates of 1.29 to 1.42 which are all statistically significant). Thus the results are coherent with the standard neoclassical model. However, the evidence from the fishery industry is mixed. One study conducted in Hawaii finds heterogeneous behaviours across captains but overall concludes that reference-dependence significantly influences the labour supply decision (Nguyen and Pingsun, 2009). Another study analysing the pear-packing industry, finds evidence of the reference-dependent theory using (quasi)-exogenously varying piece rates when analysing workers labour supply decisions (Chang, 2014).

An overall assessment of the evidence in various markets is that there is no distinct model of labour supply that solely governs the labour supply decision of workers.

Battle of Models

Essentially, the focus of the literature is the daily working hours decisions of taxicab drivers, where various specifications are used while the foundation of the methodology is not changed substantially from the original specification articulated by Camerer et al. (1997): a regression of log daily hours worked on the log of daily average hourly earnings to distinguish which model of labour supply governs the behavioural responses of taxicab drivers. However, Farber (2005) argues that it is not possible to circumvent the standard problems with wage elasticity when examining taxicab drivers. First, taxicab driver's wages are serially correlated across days making the argument of transitory wage increase not plausible. Second, the wage is not constant or positively correlated during a shift causing the average wage not to be a valid measurement of the alternative cost of leisure.

Farber (2005) gathers a data set of NYC taxicab drivers (593 trip sheets for 22 drivers over several months) and develops a model capturing factors that may influence the decision to end a shift. He suggests a dynamic model where taxicab drivers make the decision of whether to end a shift or not after each fare. (The model in reduced form will be presented in Section 3). His paper suggests that the decision to end a shift is predominantly correlated to accumulated hours worked during the shift, conditional on accumulated income. His findings are in line with the neoclassical model. Farber additionally examines the original data set from Camerer et al. (1997) and finds that applying his method, reference-dependence is not necessarily the only inference that could be drawn from their data set.
Additionally, Farber (2005) finds several data handling errors by the original authors.

Farber (2008) re-examines his evaluation of the reference-dependence theory. Even if a reference income target is important, the reference level is likely to vary across drivers and across different days of the week. Given the theoretical constraint of not knowing how to model the reference point, Farber concludes that reference-dependence may be a crucial factor influencing taxicab drivers' labour supply decisions (note that Farber's (2008) paper did not adopt any ideas enunciated by Köszegi and Rabin (2006; 2007; 2009)).

Crawford and Meng (2011) empirically apply the concept, developed by Köszegi and Rabin (2006), of an expected-based reference point and they construct proxy targets for hours supplied and income earned for every shift worked by all individual drivers (using Farber's (2005) data). The proxy targets are derived from an individual driver's expectation, using the hours worked and income earned the previous week on a given weekday. For example, the third observed Wednesday, the hours target and income target for a distinct driver are the means of the two earlier Wednesdays he drove (and consequently the fourth observed Wednesday the hours target and income target us the mean of the three earlier Wednesday etc.). Crawford and Meng (2011) do not find that neither an earnings target nor an hours target exclusively influence the labour supply decisions of NYC taxicab drivers. Instead the probability for a driver to end his or her shift increases rapidly after targets are reached, independent of the order in which the targets have been reached. However, Crawford and Meng (2011) analyse the consequence of intra-day wage fluctuations and find that if the fares are well below the expected value early during the shift (i.e. substantially below the intra-day proxy target), the income target turns out to be the formative factor. In contrast, if early earning opportunities are above the expected value, the hours target becomes decisive.

Conclusively, there are two modelling approaches where the research is addressing the shortcomings of the original paper of Camerer et al. (1997): the discrete-choice stopping model and the formulation of proxy targets. Recent research is re-examining the validity of the claims of Farber (2005; 2008) and Crawford and Meng (2011) with better data.

Searching for Better Data

A limitation with previous research on workers with flexible hours is that it has relied on self-reported data (such as Camerer et al., 1997; Farber 2005; Farber 2008; Crawford and Meng, 2011; Oettinger, 1999 etc.). As a consequence of the self-reporting data collecting procedure, spurious observations with drivers reporting the same hours worked and / or the same income for multiple days are commonly found. Often the accumulated sum of income does not match the reported income and no breaks are included in the data (Jonasson and Wållgren, 2013).

Due to the limitations of previous work data, Jonasson and Wållgren (2013) gather a new data set of 47 cabdrivers in Stockholm containing trip-by-trip data, participation, and earnings for a 3-month period derived from an electronic log system. Analysing the summary statistics the results suggest that the wage changes are correlated across days, and that the wage elasticity estimates are positive in numerous specifications.

Even though not able to complete Farber's (2005) and Crawford and Meng's (2011) work, some evidence of a cumulative hourly target is found.

The literature has diverged during recent years. Agarwal et al. (2015) find negative wage elasticity estimates for taxicab drivers. Utilizing a different approach (exploiting a permanent fare change), Doran (2014) also finds support for reference-dependent preferences in his analysis of labour supply decisions of NYC taxicab drivers. In light of this, Farber (2015) gathers a new data set, covering all trips taken in all taxicabs of NYC for five years from 2009-2013, now using an electronic log system.

Farber (2015) distinguishes between anticipated and unanticipated daily wage variation and presents evidence that only a small fraction of wage variation (about 1/8) is unanticipated, concluding that reference-dependence is not a dominating factor influencing the labour supply decisions of NYC taxicab drivers. This is consistent with Köszegi and Rabin's (2006; 2007; 2009) argument that reference-dependence is only relevant to unanticipated variation. Moreover, Farber (2015) does not find evidence of the prediction of the reference-dependent theory. The days when the average hourly wage is close to the expected average hourly wage, negative wage elasticity estimates are not found. The conclusion is that the scope of the theory is limited since a large fraction of the wage variation is anticipated, and even on days when the prediction can be tested accurately, no evidence is found of its existence. The empirical contribution of Farber (2015) is the acknowledgement to Köszegi and Rabin's (2006; 2007; 2009) contribution of the data implementation of expectation-based reference point and the improved validity of his own 2005 discrete-choice stopping model. Farber (2015) revisits his earlier findings with better data and finds that the probability of ending a shift is strongly positively correlated with accumulated hours worked and weakly related to accumulated income earned.

Furthermore, Farber's (2015) analysis shows that the research based on NYC taxicab drivers has an inherent problem: the drivers work in shifts. Typically there are two shifts per day, of which the conventional duration is around 12 hours long and also requires the driver to return the vehicle within a certain timeframe. Consequently, day shift drivers will be constrained at the end of their shift when the taxicab will be turned over to the next taxicab driver, and drivers will select hours worked before information on unanticipated daily earnings opportunities can be observed. In contrast, the night shift drivers can experience the evolution of unanticipated earning opportunities, since the lack of discretion of labour supply is relevant for their starting time. If labour supply is affected by unanticipated transitory variation, only night time drivers will have the opportunity to adjust their labour supply responses compared to day shift drivers (Farber 2015).

In summary, while wage elasticity estimates in Jonasson and Wållgren (2013) and the sophisticated analysis by Farber (2015) do not rely on self-reported data and subsequently avoids earlier mentioned problems of previous research, new information regarding the institutional setting of the NYC taxicab market arguably question the validity of Farber's (2015) findings. This and the fact that the results have not been replicated in other markets, raise further questions about the external validity of Farber's (2015) work.

Learning-by-Doing

New research addresses the possibility that taxicab drivers are not a homogenous group and analyses a potential learning behaviour. There is an extensive amount of literature that supports learning effects in many settings (see Thompson, 2010 and 2012, for surveys). Essential limitations in many of the studies on LBD are the identification of who is improving, what they improve, and how strongly individual improvement is encouraged. Haggag et al. (2016) argue that taxicab drivers are one of the few occupations where the above-identified problems are not binding. Their paper finds a strong learning behaviour, in the sense that more experienced drivers are better at finding new costumers. Both Camerer et al. (1997) and Farber (2015) discuss the existence of a learning curve but employ another definition: positive labour supply elasticity that grows with experience level. While Camerer et al. (1997) do not support the claim of a learning curve with reliable quantitative estimations, Farber (2015) uses time in his data as a proxy for experience and presents evidence that newer drivers have smaller estimates of wage elasticity and that it becomes more positive as time passes. French and Stafford (2015) find strong LBD behaviour revisiting the Florida lobster fishermen.

To conclude the literature review, Farber (2015) mitigates many of the limitations of the current research on workers with flexible hours. However, given the institutional setting in NYC, where drivers are not completely free to set their own hours and tipping may be important, the empirical validity of Farber's (2015) findings and the theoretical assessment of Köszegi and Rabin (2006; 2007; 2009) need re-examination. New evidence of reference-dependent models has in recent years been found for taxicab drivers in Singapore (Agarwal et al., 2015) and this raises the question weather or not Farber's (2015) findings are contingent on taxicab drivers to be limited in their labour supply response. This paper seeks to answer the question if the institutional setting in NYC limits the generality of the Farber's (2015) findings. In addition, my paper contributes to the existing literature on taxicab drivers using real data on experience. The relatively new literature exploring the LBD behaviour in various fields will benefit from additional empirical grounding from new markets. The differences in behaviour across drivers might alleviate the conflict of imposing a single model for all agents and allow for several existing models to have explanatory value predicting the driving behaviours of taxicab drivers. A learning curve is one potential explanation for heterogeneous labour supply behaviour across drivers (Farber, 2015; Haggag et al., 2016). The notion of heterogeneity across drivers' driving behaviours is not new, but this paper contributes the key factor of reliable data, where previous research has fallen short (Camerer et al., 1997; Farber, 2015)). Lastly, there is a theoretical gap within the existing research regarding taxicab drivers: current literature limits the possible decision horizon for taxicab drivers to one day. The analysis of driving behaviour will benefit from extending the timeframe, and possibly reference-dependence is observed when the models are extended to capture a less myopic timeframe. Thus, the importance of the notion of reference-dependent preferences and its effects on labour supply decision is still not completely determined. The Stockholm taxicab market allows for the theory to be tested within a market where the drivers are allowed to freely choose their working hours, providing external validity to Farber's (2015) findings.

3. Data

Stockholm's Taxicab Market

The Stockholm taxicab market is dominated by a number of large-scale franchise companies. 60 per cent of the fares go through a call centre. Therefore, there is a strong incentive to develop advanced log systems to track the drivers. The prevailing employment arrangement is the following: drivers lend a taxicab during a certain time period (weeks or months) and in return the drivers are required to pay a part of the income earned. The taxicab is owned by a fleet company, which in turn is connected to one of the taxicab franchise companies. Drivers only earn commission based on their income and the standard rate is about 37%. The taxicab franchise company is responsible for training the drivers, sets the fares, and operates the booking centres. The data is supplied from one of the major taxicab franchise companies in the Stockholm market (1,200 taxicabs which generates about 15,000 taxicab trips a day). The company did not change their prices during or in connection to the sample period. Consequently, the standard employment arrangement makes sure that the taxicab drivers are in fact free to set their own hours.[2] The data sorting process is found in the *Appendix 1*. The data contains trip-by-trip information of 47 drivers, randomly selected from the population of drivers working at the franchise company, over the time span of 3 months (April – June 2012). The collection was done in two phases: the trip-by-trip data was collected in April of 2013 and individual driver characteristics were obtained in April 2016. Weather data is publically available from SMHI.

[2] A precautionary measure is taken: separation of the vehicle code and driver identification number excludes the possibility of multiple drivers driving one vehicle. No evidence of multiple drivers per vehicle was found.

Sample Characteristics

In DAY1 a taxi driver works on average about 11.5 hours per day, and earns a wage of approximately 95 SEK per hour. The reported average daily temperatures during the examined time period ranges from 0.2 to 18.6 degrees Celsius with an average of 10.29. The amount of daily rainfall is within the range 0 and 33.5 millimetres with a sample mean of 2.65. The two measures of experience are number of years as an authorized taxicab driver and tenure at the franchise-company. The medians of the two measures are 8 and 4 years, respectively, and both exhibit a large spread.

Table 1: Sample Characteristics

	Median	Mean	Std. Dev.
DAY1 (n=3530)			
Hours worked	11.52	11.60	3.14
Average wage (SEK)	93.67	94.89	28.78
Total revenue (SEK)	2979.00	2961.50	1101.53
Rain (mm)	0	2.65	5.99
Driving Experience			
Authorized	8	10.0628	7.2388
Tenure	4	4.7872	3.2284

Note: The table contains information on the key variables and summary statistics.

Turning to distributional figures: Figure 1 shows the distribution of shift lengths in hours (the numbers are truncated, e.g., 11.5 hours is shown as 11). The modal shift duration is in the 12th hour, and shift durations are concentrated between hours 9 and 15 (92.6 per cent). 3.7 per cent of shifts are less than 6 hours, and 2.5 per cent of shifts are 16 hours or longer. There are some shifts that are relatively long, well above 20 hours. Legislative action to prevent this has been taken, but the occurrences of extremely long shifts have been corroborated by informal interviews with taxicab drivers. However, excluding the longest shifts has no substantial impact on the results presented in Section 5. Comparing the Figure 1 to Farber (2015), the taxicab drivers supply more hours in my sample. However, the actual distribution is fairly similar.

Figure 2 shows the fractions of shifts in the sample starting on different clock hours. There are no clear discontinuous spikes. The only spikes observable are at 5:00AM and 7:00AM. The fractions of shifts starts are 17.1 per cent and 16.9 per cent respectively. The number of shifts that starts early is probably related to demand variation and that taxicab drivers are picking up passengers who are going to Arlanda

Airport. The lack of discontinuous spikes is in stark contrast to the pattern observed after a visual inspection of other available data sources. Farber's (2015) data set has two spikes, accounting for almost the whole sample (86.9 per cent), leaving only a small fraction of shifts (4.6 per cent) that start during the eight hours between 8PM and 3:59AM. The remaining shifts (8.5 per cent) start in the four hours between 10AM and 1:59PM. In addition, Farber (2015) neglects the drivers with no clear starting pattern when conducting the analysis and splits the sample into "Night" and "Day" drivers. He argues that there is a sharp difference in earnings patterns between day and night shift drivers and that the drivers are exposed to dissimilar amount of available information, which could have consequences on their labour supply decisions. I have conducted an experiment dividing the drivers in the same manner as Farber (2015), neglecting the fact that the Stockholm taxicab drivers do not work in shift. However, Figure 2 implies that this is not a plausible categorization of my sample and no result of differences in earning patterns is obtained.

Figure 1: Distribution of Shift Length

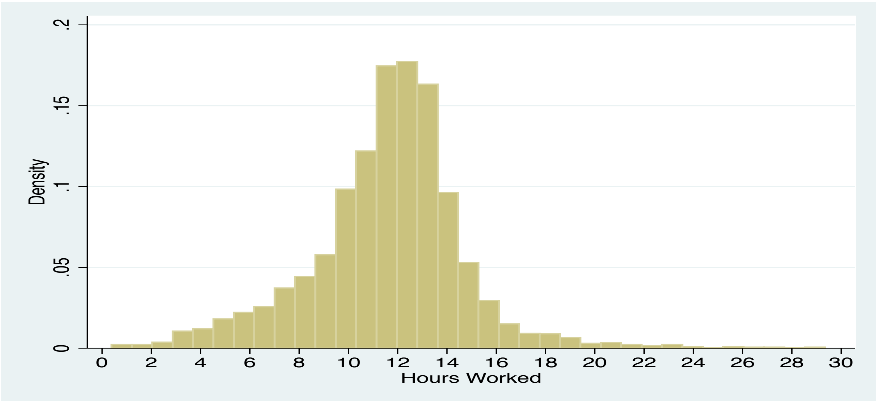


Figure 2: Distribution of Shift Start Time, By Clock Hour

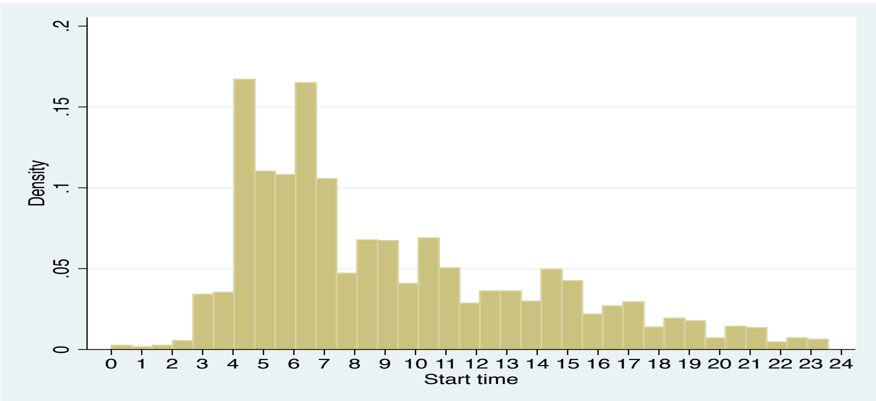
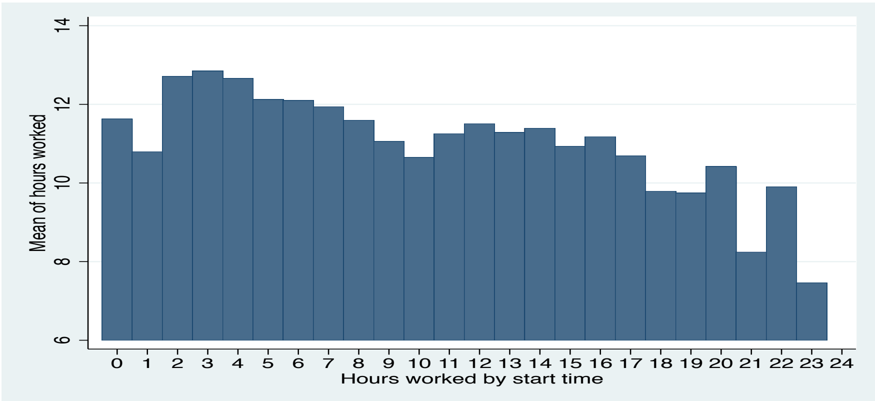


Figure 3 shows the groups of shifts starting on the same clock hours and their average durations in hours. Interestingly, shifts that start early in the morning last longer on average (more than 11 hours) than those starting later during the day (10 to 11 hours).

Shifts starting later in the evening (from 7PM on), last progressively less time. One reason for the late shift starting hours could be that drivers react to the demand fluctuations and choose to start working when people are going home after a night out.

Figure 3: Average Hours Worked During a Shift, by Clock Hour



Labour supply and earnings have distinct patterns over the week. Figure 4 plots the average shift length by day of the week. Drivers work the same amount of hours on average during Monday to Thursday (fairly close to the sample mean). On Friday, working shifts are longer, most likely due to the fact that some drivers are serving dual costumers, business people during the day and leisure costumers during the night. In contrast, on Saturdays average work hours are the second lowest, which is quite expected since compared to Friday there is no business consumer segment. Sunday is clearly the day when drivers, on average, work the least amount of hours.

Total income per shift, shown in figure 5, generally follows hours worked except for Saturdays. Drivers do not keep the entire earned income as wage, but receives 37 per cent of the daily income. The actual percentage is equal across the franchise-companies and was not changed during the sample period. Drivers earn the highest income per shift on Saturdays compared to hours of work, implying that Saturdays' average hourly earnings are higher (see Figure 6). Figure 6 is comparable to Figure 4-5 with the significant difference that Saturdays' hourly wages are high (largely driven by the high income during a shift compared to hours worked). Also, hourly wages on Sundays are high by low amount of hours worked compared to total income.

On an aggregated level, i.e. when only looking at the mean value of variables, there is no clear evidence that drivers adopt neither the neoclassical theory nor the reference-dependent model of labour supply. For example, on Sundays drivers supply few hours even though the wages are relatively high.

Figure 4: Distribution of Hours Worked, by Day of Week

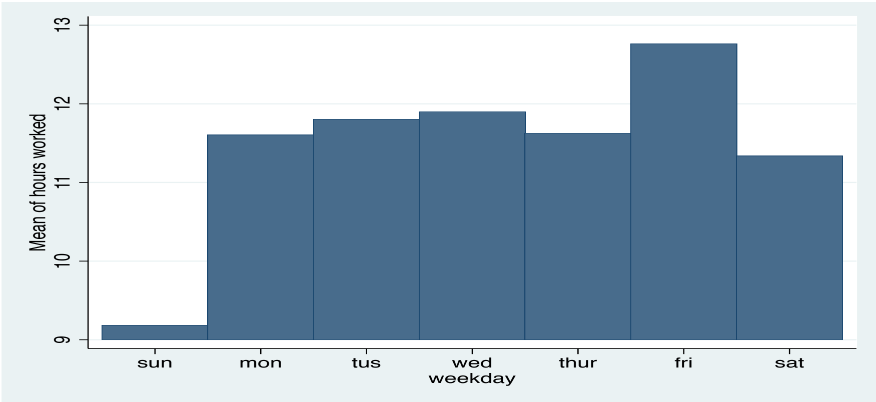


Figure 5: Distribution of Shift Earnings, by Day of Week

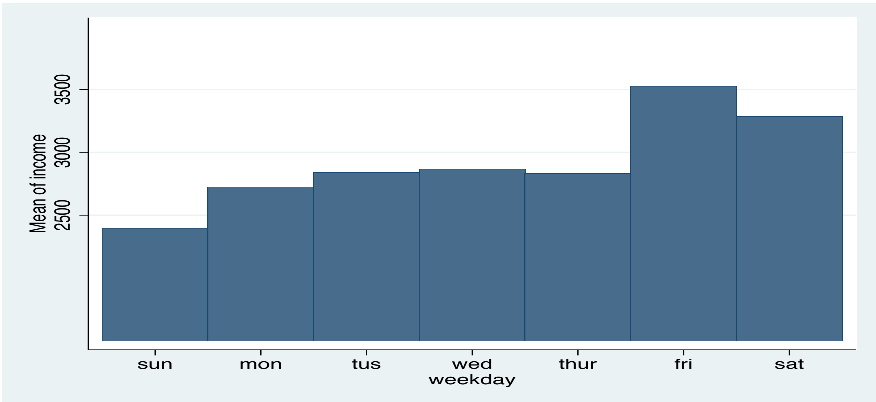
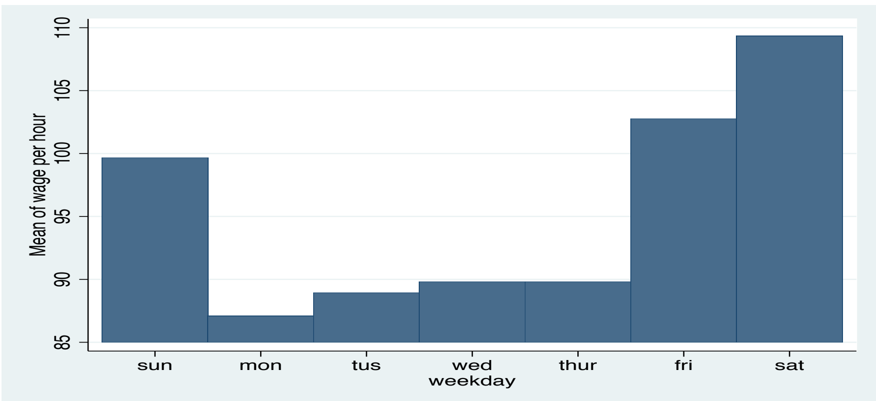


Figure 6: Average Hourly Wage, by Day of Week



More graphs are found in *Appendix 1*. Figure 9 to Figure 13 show driving decisions and outcome for the whole sample period. Interestingly, a pattern with remarkable persistency is displayed in all figures. For example, in Figure 12, there is a clear pattern showing that the mean of hourly wage is remarkably consistent with the graphical inference made from Figure 6. The remaining pictures capture the same notion but with a variation of the time frame: day of the week (see Figure 14 -18) and weeks (see Figure 19-21).

To conclude, according to Figure 2 it appears that the Stockholm taxicab market institutional setting indeed allows drivers to set their own working hours without any restriction. This is in contrast to earlier studies, where spikes in night shift starting time / day shift ending times clearly indicate a restriction in the labour supply decisions among drivers. However, even if drivers are free to work whenever they want, their behaviours seem to follow a pattern. Indeed, hours and wages temporarily increase and decrease across days, but in a systematic manner. This is in line with the only prior article discussing the subject (Farber, 2015) and consistent with the earlier findings of strongly correlated average daily hourly wages found by Jonasson and Wållgren (2013). This indicates that prior work might have been wrongly assuming a one-day at a time decision-making horizon and that the temporal aspect of labour supply decisions is important.

4. Conceptual Framework

In this section I present an overview of the empirical and theoretical set up of Farber (2015). The analysis in his paper follows three specific areas: 1) anticipated or unanticipated wage variation and how wage elasticity varies given the difference between actual wage and expected wage, 2) the discrete-choice stopping model, and 3) heterogeneous behaviour. In addition, I analyse the consequences if the time frame is extended.

Anticipated or Unanticipated Wage Variation

The wage elasticity plays an important role in the determination of which model of labour supply that governs the labour supply decision of taxicabs drivers. Previous research conducted before Farber (2015) and Crawford and Meng (2011) have simply relied on measuring the standard (Frisch) wage elasticity, drawing the inference that the obtained estimates support the notion of reference-dependence if the wage elasticity is negative. However, as previously discussed, the role of reference-dependence in explaining variation in labour supply is limited to the response to unanticipated wage variation (Farber, 2015).

First, to distinguish the anticipated wage variation from the unanticipated wage variation, Farber (2015) calculates the expected wage for each day using data on daily average hourly wage rate. The expected wage is calculated as the predicted value of log average hourly wage from an OLS regression of the average hourly log earnings on indicators for day of week, week of year, and month of the year. This gives information on how much of the variation in wage that is anticipated versus unanticipated, by decomposing the total wage variance into variance of the predicted values and the residuals. How much of the total variance of the average wage that comes from the residuals will determine the scope of the reference-dependence theory regarding of the labour supply decision. Farber (2015) argues that this decomposition takes a conservative approach when estimating what drivers anticipate, since there ought to be numerous factors, in addition to the temporal factors, that the drivers feasibly can anticipate. However, simply controlling for the time dimension, Farber (2015) gives the reference-dependence theory its best opportunity and the estimates of the magnitude of the unanticipated wage variation will be exaggerated (since it will also include some anticipated wage variation).

Second, reference-dependence is a local phenomenon. Therefore, it is important to estimate the labour supply response within the region in which the hypothesis of the theory can be tested. According to Farber (2015) there are three cases of earning patterns that can be distinguished and ways that drivers are likely to react to unanticipated wage fluctuations in a model with reference-dependent preferences: 1) drivers that earn substantially lower wages than expected will end their shifts before the target is reached, 2) drivers that earn substantially higher wages will continue driving after their targets are reached, and 3) drivers that are sufficiently close to the expected wages will drive until their targets are reached (see mathematical derivation in *Appendix 2*). The differences between the observed average daily hourly earnings and the predicted values are used as an indicator of the daily deviation of the average daily log wage from the expected wage (Farber, 2015). When expected wage is close to the actual wage, the reference-dependent theory predicts the wage elasticity to be negative. Farber (2015) separates days when the expected mean hourly wage is close to the actual mean hourly wage, as a plausible way of testing if the wage elasticity differs between days when the deviations from the expected wage are small or big. If the deviation is small, that means the wage is close to the expected wage, and this is within the region of where it is likely to observe targeting behaviour.

Following Farber (2015) and the argument above, I split the sample into three subgroups: 1) days with an absolute deviation in the bottom 25 per cent, 2) days with an absolute deviation in the second quartile, and 3) days with an absolute deviation above the median. After distinguishing the days, the wage elasticity of the drivers is estimated using OLS. The suggested model is formulated in Eq. (1).

$$\ln H_{it} = \eta \ln W_{it} + X_{it}\beta + a_i + \varepsilon_{it} \qquad (1)$$

Where H_{it} is number of hours worked during a shift, W_{it} is the mean wage during a shift and X_{it} is a vector with weather conditions and day of the week, week of the year, and month of the year dummy variables. In addition, a driver fixed effect, a_i is included to account for constant driver heterogeneity and ε_{it} is the error term.

In respect of notation, days when absolute deviation is in the bottom 25 per cent $t = j$, and days when absolute deviation is between per cent 25 and 50 per cent as $t = k$ and days when the absolute deviation is above the median as $t = l$ gives

$$\ln H_{ij} = \eta_1 \ln W_{ij} + X_{ij}\beta + a_i + \varepsilon_{ij} \qquad (1a)$$

$$\ln H_{ik} = \eta_2 \ln W_{ik} + X_{ik}\beta + a_i + \varepsilon_{ik} \qquad (1b)$$

$$\ln H_{il} = \eta_3 \ln W_{il} + X_{il}\beta + a_i + \varepsilon_{il}. \qquad (1c)$$

The empirical prediction of the reference-dependence theory is that wage elasticity estimate in Eq. (1a) is $\hat{\eta}_1 = -1$. If the model in Eq. (1 a, b, c) is correctly specified, the relative magnitudes of the estimates are predicted to be $\hat{\eta}_1 < \hat{\eta}_2 < \hat{\eta}_3$ if the drivers are reacting consistent with the reference-dependence theory. In essence, this strategy allows determining the possible scope (the fraction of the wage variation that is unanticipated) and accuracy (negative wage elasticity measurements) of the reference-dependent theory of labour supply.

Measurement error is still a practical concern even though the data is derived from an electronic log system (as observed by Farber (2015) and Jonasson and Wållgren (2013)). Following previous literature, non-overlapping drivers' hourly wages at time *t* (lnW_t) are used as an instrument for the average hourly wage of driver *i* for shifts that start on date *t* (lnW_{it}).

The identifying assumption for the IV approach is that the instrument lnW_t must be relevant and exogenous. lnW_t needs to be strongly correlated with lnW_{it}, which is arguably satisfied here, (later tested and confirmed), since the reported hourly wage of non-overlapping drivers driving the same day are highly correlated with individually reported hourly wages from drivers working the same day. Secondly, lnW_t must satisfy the exclusion restriction. In this case, lnW_t should not have an effect on $\ln H_{it}$ other than through lnW_{it}. It is not possible to test the exclusion restriction, and hence the validity of the instrument needs to be argued. Previous research argues that since one of the motivations to use an instrument is due to the potential measurement error, the exclusion restriction will be satisfied if the measurement error is uncorrelated across drivers (Camerer et al., 1997; Farber 2005; 2008; 2015). However, there is criticism regarding using the proposed instrument and Farber (2015) argues that one concern is whether lnW_t captures demand information or noy. If this information is communicated across drivers, lnW_t would have an effect on the labour supply decision of driver i. Irrespectively, it is unknown to what extent lnW_t captures the demand fluctuation conversations and a consensus is formed in the literature to use the proposed instrument. Alternative methods are used that do not rely on the IV approach (see below).

Discrete-Choice Stopping model

While estimating wage elasticity gives some information regarding models of labour supply, the interpretation of the wage elasticity is complicated. This is due to the fact that observable wage is not constant at any point during the day because of variation in demand and other factors (Farber, 2015). A possible solution is to look at each time a fare ends as a decision point for the driver: the driver can continue to work or the driver can end the shift. This is a dynamic discrete-choice problem based on a driver's comparison of the marginal utility of ending a shift versus the marginal utility of continuing working, (and this assessment is made after each fare) (Farber, 2005).

The decision for a taxicab driver to end a shift can be seen as a simple discrete-choice problem (Farber, 2005; 2008; 2015). A driver can calculate at any point τ during a shift the optimal stopping point τ^*. The optimal stopping point can be a function of multiple factors: hours worked so far during the shift, the evaluation of further earning potential during the shift as well as the fact that it is Friday. As discussed above, if daily income is an important factor, the optimal stopping point is a function of income earned so far during the shift. Thus, this mirrors the prediction of the reference-dependence theory of labour supply. A driver quits working at τ if $\tau \geq \tau^*$ so that $\tau - \tau^* \geq 0$. Following Farber (2005; 2008; 2015) the reduced form representation of the choice of ending a shift where $R_{it}(\tau) = \tau - \tau^*$ is

$$R_{it}(\tau) = \gamma_1 h_\tau + \gamma_2 y_\tau + X_{it}\beta + \mu_i + \varepsilon_{it\tau} \qquad (2)$$

where the quantity h_τ is a vector of hours worked at τ and y_τ is a vector of indicators income earned at τ. i indexes the particular driver, t in indexes time (day of week, week of year, and month of the year), μ_i is a driver fixed effect, and ε_i is the error term. X_{it} is a vector which includes fixed effects for the hour of the day, day of the week, and week of the year.

This modelling approach gives verifiable predictions for the two competing theories of labour supply. The neoclassical model implies that the probability of ending a shift after a given fare is bound to relate positively to accumulated hours (conditional on accumulated income) and be unrelated to the income level (conditional on accumulated hours). In contrast, the reference-dependence theory implies that the probability of ending a shift after a given trip relates positively to accumulating income (conditional on accumulated hours) and relates less strongly to accumulating hours (conditional on accumulated income). Thus, the prediction of the theory is that $\hat{\gamma}_2 > 0$. Eq. (2) is estimated with squared terms to allow for a non-linear relationship. In addition, Farber (2015) argues that the specification above suggests that a probit model is suitable based on the latent variable. In Section 5, results are presented from both the OLS and the probit model (the marginal effect) to allow for the precision of the estimates of the reduced model to be evaluated (the probit model's marginal effects are obtained with the other variables set at their mean value). However, the sizes of the estimates are not important when examining the results of the models. It is the relative size of $\hat{\gamma}_1$ and $\hat{\gamma}_2$ that gives explanatory power.

Heterogeneous Behaviour and a Learning Curve

Economists often assume that estimates that are obtained from the suggested model apply to all agents. Indeed, this is a common feature of the literature analysing the labour supply decision of taxicab drivers (Farber, 2015). Some drivers might exhibit reference-dependent preferences (substantial negative wage elasticities) while other drivers are optimizers (positive wage elasticities). Measuring potential differences in driving behaviours may help explaining previous findings of negative wage elasticity in studies that is relying on limited data and furthermore stress the importance of having a random sample of drivers (i.e. Camerer et al., 1997; Chou, 2002; Agarwal et al., 2015). The Stockholm data set also allows for testing the hypothesis that drivers learn on the job, since the data set contains information on years as an authorized taxicab driver and tenure at the franchise-company (see *Appendix 1* for employment regulations).

Following Farber (2015), I estimate the labour supply curve for each driver individually based on Eq. (1). The individual elasticity is then the foundation of the estimation of a density function (Kernel density function) and this provides the range of estimates for the population of Stockholm city cab drivers. From the density estimates it is easy to compute a cumulative density function estimate from the density function. If the density function is highly concentrated, it implies a conformed behaviour of the taxicab drivers within the sample (Farber, 2015). An extreme case is that the density function is uniformly distributed over a specific interval, which suggests a dispersion of labour supply models across drivers.

Camerer et al. (1997) and Farber (2015) both discuss the possibility of explaining divergence in driving behaviour among taxicab drivers with the occurrence of LBD-process. Having no data on experience, Farber (2015) suggests a simple approach where he sorts the data based on time spent driving of each driver within his sample.

The Stockholm data set allows me to replicate this strategy with two actual measures of experience (years as an authorized taxicab driver and tenure at the franchise-company). The computational strategy is to sort each driver with respect to experience and then estimate different labour supply models for each category. If there is a learning process, plotting the different estimates of wage elasticity could potentially display a learning curve in the sense that drivers with more experience are more likely to have relatively higher estimates of wage elasticity.

One Day at A Time? Extending the Time Horizon

Even though Farber's (2015) latest contribution incorporates a temporal aspect of the labour supply decision, one can argue that it is still is a very narrow timeframe (one day). Farber's (2015) model allows drivers in the model to make daily decisions at the intensive margin but never to look ahead. This view may seem myopic since drivers can arguably have longer time horizons. To capture this, the model and earlier studies will be replicated, but with the application of extending the time window, (i.e. two-day, three-day, or seven-day shift respectively). The interpretation of the discrete-choice stopping model is as straightforward as described above. In essence, the modelling approach captures the possibility of a longer timeframe and can be viewed as an empirical foundation when developing a dynamic model of labour supply decisions and taking into account passed effort levels.

Extending the time frame, the appearance of serial correlation across time periods is still a problem. Thus, the method employed can not determine whether the potentially negatively estimated wage elasticity is a confirmation of reference-dependence theory or if it capture a substitution effect, when ending a shift early on days is simply because drivers may expect higher wage the next shift. However, as Jonasson and Wållgren (2013) noted, even if correlation was very high between days, the correlation decreased when adding days together implying that extending the time window might mitigate the serial correlation across time periods. Thus adding days together might be one way of obtaining more precise estimates on wage elasticity.

Analysing the discrete-stopping model is simply a natural extension of current states of knowledge to see if the probability of ending a shift is only observed one day at a time. Or can the ending a shift also be because of the fact that accumulated hours worked has been higher for three days and the drivers are suffering from fatigue? This sort of experiment is novel within the field of labour supply decisions of taxicab drivers.

5. Results

In this section I first replicate Farber's (2015) analysis. The structure fallows the steps laid out in Section 4, and subsequently I commence with presenting the decomposition of the variance of the mean hourly wage and estimating different labour supply functions for different days allotted by absolute deviation from expected wage. Second, the result from the discrete-choice stopping model is presented. Third, heterogeneous behaviour and a potential learning behaviour are analysed and this completes the replication exercise. Lastly in this section I present the results and implications of widening the time horizon.

How Much of the Wage Variation is Unanticipated?

The regressions of average hourly wage on day of the week, week and month are found in *Appendix 3*. The fitted values of the regression are the expected wage. The results from the regression (that are displayed in the *Appendix 3*) suggest that a substantial part of the wage fluctuations are anticipated and can be explained by temporal factors. In Table 2 the average daily mean hourly wage's variance is decomposed into an anticipated variation and an unanticipated variation by separating the variance of the predicted values and the residuals respectively. Most variation is transitory anticipated variation (79 per cent). This limits the scope of the reference dependence theory to account for variation in labour supply. The remaining 21 per cent of the total variation is transitory unanticipated variation.

Table 2: Variance Decomposition of Average Daily Mean Hourly Wage

	Total Variation	Anticipated Variation	Unanticipated
$ln\,(\overline{W_t})$	0.0171	0.0136	0.0035

Note: The variance decomposition is obtained by taking the variance of the predicted values from Table 11 (see *Appendix 3*). The unanticipated variation is the variance of the residuals. Together they add up to the total variance of the average daily mean hourly wage.

The decomposition of anticipated and unanticipated wage variation is done in a simplistic way. There are elements that are not controlled for when doing the regression of the expected wage. Consequently, parts of the residual may be anticipated and the unanticipated variation may be even lower. The results in Table 2 are in line with the findings in Farber (2015) and limit the scope of reference-dependence theory to explain variation in hours supplied, since 80 per cent of the variation will be governed by the neoclassical theory and only 20 per cent of the variation has influence on the gain and losses utilities. Hence the overall assessment is that the labour supply of taxicab drivers is decided by anticipated variation in demand through day of the week and other temporal aspects.

Targeting behaviour is only expected to be observed within a certain interval of the realized wage Farber (2015). If the wage is lower than what is expected, drivers will find it optimal to stop working before the reference income level is reached. If the wage is higher than what is expected, drivers will find it optimal to continue working after the reference income level is reached. Thus, separate regressions are run for days where the absolute difference between the average log hourly earnings and expected log average hourly earnings are very small, relatively small, and large. The sample is split into three subgroups: a) days with an absolute deviation in the bottom 25 per cent (an absolute log wage deviation less than 0.0018), b) days with an absolute deviation in the second quartile (an absolute log wage deviation between 0.0018 and 0.0058), and c) days with an absolute deviation above the median (an absolute log wage deviation larger than 0.0058). Results are found in Table 3.

Table 3: Wage Elasticity, OLS of Eq. (1 a, b, and c),
Subsamples by Deviation of Average Log Daily Wage from Expected Value

Sample	Deviation Percentile	Elasticity
(a)	0-25	-0.0654*
	(N=883)	(0.0334)
(b)	25-50	0.0784**
	(N=882)	(0.0321)
(c)	50-100	0.14532***
	(N=1765)	(0.0216)

Note: 25th percentile and the median of the absolute deviation from its expected value are 0.00182 and 0.00581. Decomposition of the absolute deviation is found in *Appendix 3*. The expected value is the predicted value from the regression of average log hourly earnings on indicators for day of week, week of year and month of the year, and is also found in *Appendix 3*. Elasticities are the estimated coefficients of Eq. (1 a, b, c) which include additional control variables (weather conditions, time and driver fixed effects). Robust standard errors clustered by drivers are in parentheses and *** p<0.01, ** p<0.05, * p<0.1

If the actual wage is close to the expected wage (Sample (a)), the wage elasticity is negative with a value of $\hat{\eta}_1$ is - 0.0654 and statistically significant at the 10 per cent significant level. If the actual wage is fairly close to the expected wage (Sample (b)), the wage elasticity estimate is statistically significant and $\hat{\eta}_2 = 0.0784$. However, if the actual wage is very far from the expected wage, the elasticity becomes significantly positive and $\hat{\eta}_3$is 0.14532 (Sample (c)). Regression outputs in Table 3 show that the wage elasticity is smaller on days when the deviation is small. This is evidence of some support for the reference-dependent theory. The estimate is far from the predicted value of $\hat{\eta}_1$= -1 and the result implies that the utility curve is not very "kinked".

The IV estimates of the labour supply elasticities are found in Table 4 and the first stage result is located in *Appendix 3*. The estimates of the wage elasticity follow the same pattern as the wage elasticity estimates in Table 3. The sample is separated in the exact same manner as before. If the actual wage is close to the expected wage (Sample (a)), the wage elasticity is negative with a value of - 0.0834 and statistically significant at 5 per cent significant level. If the actual wage is fairly close to the expected wage (Sample (b)), the estimate is 0.0857 and statistically significant. However, if the actual wage is very far from the expected wage, the elasticity becomes significantly positive with a value of 0.1506 (Sample (c)). The consistency of the estimates is reassuring and standard errors increase slightly.

Table 4: Wage Elasticity, IV Regression of Eq. (1 a, b, and c),
Subsamples by Deviation of Average Log Daily Wage from Expected Value

Sample	Deviation Percentile	Elasticity All Shifts
(a)	0-25	-0.0834**
	(N=883)	(0.0411)
(b)	25-50	0.0857**
	(N=882)	(0.0371)
(c)	50-100	0.1506***
	(N=1765)	(0.0271)

Note: 25th percentile and the median across days of the absolute deviation from its expected value are 0.00182 and 0.00581. The decomposition of the residual is found in *Appendix 3*. The expected value is the predicted value from the regression of average log hourly earnings on indicators for day of week, week of the year and month of the year. Each estimated elasticity is from a separate IV regression of Eq (1 a, b , c) (first stage in *Appendix 3,* Table 12). The instrument for average log hourly wage is the average of average log hourly wage for non-overlapping drivers on the same day. The regressions include control variables (weather conditions, time and driver fixed effect). Robust standard errors clustered by driver are in parentheses and *** p<0.01, ** p<0.05, * p<0.1

The results of Table 2 – 4 are consistent with Farber's (2015) findings. However, the consistency and precision of my estimates are somewhat better and the difference between the OLS and IV approach is not as substantial as in the Farber (2015). This suggests that the measurement error may be less of a concern in my data set. The overall assessment of the first part of the replication exercise of estimating the slope of the labour supply curve is that the results are coherent with the Farber (2015), and that the labour supply decision is not governed by the reference-dependent theory.

Moreover, the results are in line with my previous work (Jonasson and Wållgren, 2013) even though the estimates presented above give the reference-dependent theory its best conditions to be evident. Given that the data used to produce the above results are free from the institutional constrains (taxicab drivers in NYC not being able to set their own working hours freely and that tipping is likely to be an substantial part of the drivers' income), and are nonetheless consistent with Farber (2015), it is not likely that the constraints do substantially influence the results in Farber (2015). The first part of the replication exercise gives external validity of Farber's (2015) findings and questions the importance of reference-dependence when analysing the labour supply decision of taxicab drivers.

The Discrete-Choice Stopping Model

I exploit the fact that I know the accumulated earnings and hours worked after each trip for each driver, and that I know after which trip the shift actually ends. Table 5 contains estimates of the Farber (2005) discrete-choice stopping model and estimates Eq. (2).

Table 5: Probability of Ending a Shift, Linear Probability Model

	1	2	3	4
Hours	0.0834***	0.0836***	0.1329***	0.1315***
	(0.0009)	(0.0009)	(0.00269)	(0.0027)
Income	0.0001***	0.0001***	-0.0000***	-0.0000***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$Hours^2$			-0.0019***	-0.0018***
			(0.0001)	(0.0001)
$Income^2$			0.0000***	0.0000***
			(0.0000)	(0.0000)
Driver fixed effect	No	Yes	No	Yes
Time fixed effect	Yes	Yes	Yes	Yes
Daily observations	3,530	3,530	3,530	3,530
Number of drivers	47	47	47	47
Adjusted R-sq	0.8016	0.8516	0.8062	0.8663
Constant	-0.5422***	-0.5424***	-0.6805***	-0.6741***
	(0.0111)	(0.0089)	(0.0149)	(0.01439)

Note: Result of the discrete-choice stopping model (estimations of Eq. (2)). Income is scaled by a factor of 100. I include squared terms to allow for differences in functional form. The regressions include fixed effects for drivers (2 and 4), and day of the week, week of the year, and month of the year (all specifications). Robust standard errors clustered by driver are in parentheses and *** p<0.01, ** p<0.05, * p<0.1

The estimates in Table 5 imply that an hours target, relative to an income target, largely explains the driving behaviours of taxicab drivers. Models 1-4 predict that if a driver works for one more hour, the probability of ending a shift, keeping income constant, increases from 8 to 13 per cent. Moreover, accumulating income during a shift is not accompanied by a large increase of the probability of ending a shift, (notice that the income measure is scaled, marginal increase is given by an increase of 100 SEK). When including additional terms to allow for a non-linear functional form, the estimates are slightly higher. When accounting for driver heterogeneity (FE), the estimates do not substantially change.

As discussed earlier, one drawback with the linear probability model is that for some observations the predicted probability of ending a shift could be below zero or above one ($P(stopping) > 1\ or\ P(stopping) < 0$) and additionally imposes that hours and income enter the decision linearly. One solution to this problem is to employ a probit model. To investigate the marginal effects, and to promote comparability to previous research, Table 5 presents probit marginal effects given categorization of the discrete-choice stopping model in accordance with Farber (2015).

Table 6: Marginal Effects of Income and Hours on Probability of Ending Shift, Probit Model

Income	(1)	Hours	(2)
300-999	0.0002	3-5	0.0026***
	(0.0001)		(0.0002)
1000-1999	0.0001	6	0.0211***
	(0.0001)		(0.0002)
2000-2999	0.0003*	7	0.0242***
	(0.0001)		(0.0004)
3000-3999	0.0004*	8	0.0398***
	(0.0002)		(0.0003)
4000-4999	0.0001***	9	0.0472***
	(0.0000)		(0.0002)
5000-5999	0.0001***	10	0.0850***
	(0.0000)		(0.0006)
>6000	0.0000***	11	0.1154***
	(0.0000)		(0.0003)
		12	0.1716***
			(0.0003)
		13	0.1869***
			(0.0005)
		14	0.1289***
			(0.0004)
		15	0.0669***
			(0.0005)
		>16	0.0037***
			(0.0002)

Note: Based on estimates from the probit model of probability of ending a shift (columns 1 and 2). The marginal effects are obtained by setting the other variables at its mean value and neglecting the squared terms. Marginal effects are obtained from Wednesday at 2:00 PM. Robust standard errors clustered by driver are in parentheses and *** p<0.01, ** p<0.05, * p<0.1

Column (1) of Table 5 contains estimates of the effects of accumulated income on an income target obtained from probit regression (neglecting the squared terms since the probit estimates allow for non-linearity by construction). As shift income accumulates (conditional on hours at the mean - column (1)), the probability of ending a shift shows no tendency to increase. There is no increase in the probability of ending a shift as income accumulates during the shift. The results are all statistically significant but clearly do not have economic significances given that the estimates are all well below 1 percentage point.

Column (3) of Table 5 contains the analogous results for accumulation of hours (conditional on income at the mean) on an hours target. The probability of stopping is sharply increasing with accumulated hours. The probability of ending a shift increases around 17 percentage points when a driver has driven for 12 hours compared with earlier during the shift.

Graphical representations of the natural survival probabilities are found in *Appendix 3* (see Figure 22 – 23). Figure 22, portraying the natural survival rate for ending a shift based on income, is slightly more linear compared to Figure 23 (which shows survival rate for ending a shift based on hours). Conforming to Köszegi and Rabin (2006) and Crawford and Meng (2011), Farber (2015) suggests an unsophisticated way of looking at dual targets by computing an interaction term between h_τ and y_τ and include the term when estimating Eq. (2). If the interaction term is included the results in Table 5 do not change and the interaction term is not statistically significant. However, I do not obtain the probit estimates due to the computational complexity when performing marginal effects of an interaction term.

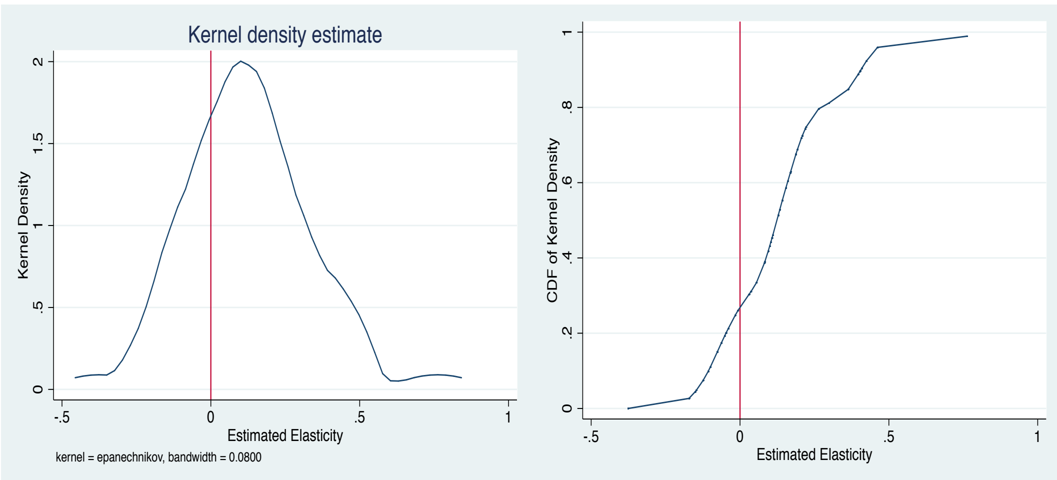
The results presented in Table 5 – 6 are inconsistent with the reference-dependent theory of labour supply, which stipulates that accumulated income is the driving force of the decision to end a shift. In contrast, the results indicate, in accordance with the neoclassical theory, that the decisive factor influencing the probability to end a shift is accumulated hours worked during a shift. However, the fact that that the results indicate an hours target is consistent with both theories of labour supply. Keeping accumulated income constant, the neoclassical theory predicts a substantial increase in the probability of ending a shift after each passing hour since the driver is not making any money. Likewise, the notion of reference-dependence could be expanded to incorporate a target in respect to hours. Farber (2015) notices that it is difficult to disentangle the prediction of the two theories with respect to an hours target, and concludes that if reference-dependence theory is limited to an hours target, not much insight has been gained from the theory since this is in line with the neoclassical theory's prediction.

To sum up, the results suggest a limited role of reference-dependence and are also consistent with the results found in Table 2 – 4, Farber's (2005:2015) results of the discrete-choice stopping model, and the estimation of a naive stopping-targeting model presented in Jonasson and Wållgren (2013) obtained from aggregated data.

Heterogeneous Behaviour

In this section, I return to estimating the relationship between log hours worked and the log wage and the estimation of separate labour supply models across drivers. The left panel contains kernel density estimates of the distribution of drivers' wage elasticities obtained from separate IV regressions (with the main specification Eq. (1)). The right panel is the cumulative distribution function of wage elasticity implied by the kernel density. Figure 7 clearly demonstrates significant differences across the driving behaviours observed in the sample. Noticeably, the median driver has positive wage elasticity estimates, but the dispersion is large. This highlights the importance of technique and sample size. Shortcomings with obtaining a random sample with few drivers can arguably explain results in previous studies (especially Camerer et al. (1997) and Chou (2002)). The Stockholm data set is based on 47 individuals, which does not allow for definite evidence to be presented. However, the conclusion of heterogeneous behaviours across drivers is in line with Farber (2015) and Haggag et al. (2016).

Figure 7: Kernel Density Estimates of Distribution of Estimates of Elasticities by Individual Drivers

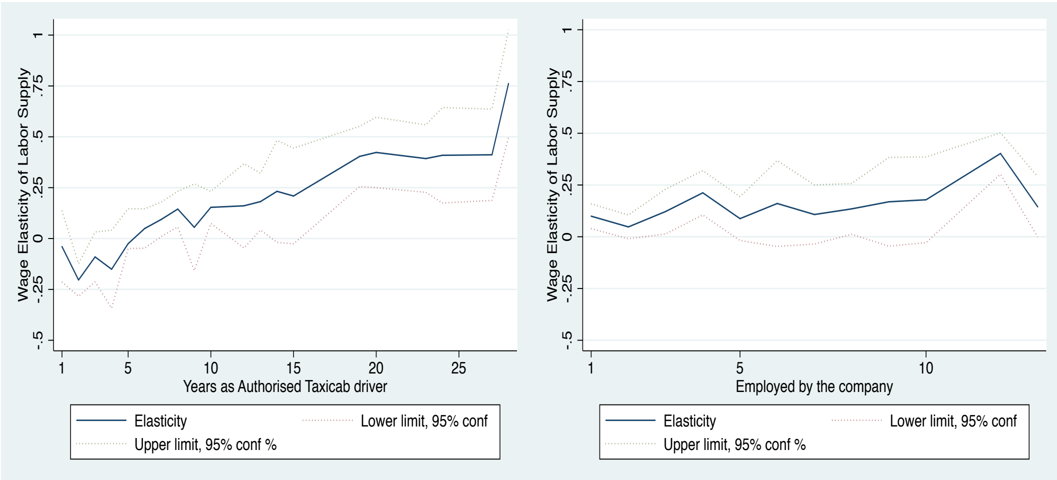


Note: Kernel Density Estimate and implied cumulative distribution function of IV estimates of wage elasticity by individual driver.

Even if the sample size is limited, I can estimate potential learning curves for the drivers. The data is sorted based on two different measures of experience and then the slope is estimated from the relationship between log hours worked and the log wage. The first measure of experience is the number of years working as an authorized taxicab driver. The second measure is for how long time the taxicab driver has been driving for the franchise-company. The correlation between the two measures is 0.25. Intuitively one could argue that years as an authorized driver should have a larger impact on optimization behaviour than the alternative measure (tenure at the franchise-company), since the first measure is a better proxy of true experience. The

IV regression estimates are displayed in Figure 8, with the dashed line being the 95 per cent limit of the confidence interval.

Figure 8: Wage Elasticity of Labour Supply, IV Estimates by Experience



Note: Wage Elasticity Estimate by Experience. In the left panel years as an authorized taxicab driver is used as a proxy for experience. In the right panel tenure within the franchise-company is used as a proxy for experience.

The left panel of Figure 8 indicates a learning pattern, and the right panel the estimates are not statistically different from zero (except for one year). It is hard to make inference from the above figure since the data set contains so few drivers. Nevertheless, the figure is consistent with the notion in Haggag et al. (2016) and Farber (2015) and the claims made by Camerer et al. (1997) of a potential learning curve.

Time Horizon – One, Two, Three or Seven Days at a Time

I now return to the question of wage elasticity estimates and the possibility to get more precise estimates of wage elasticity (for one-day time horizon see Jonasson and Wållgren (2013) Table 5 – 6 with estimates of wage elasticity of 0.13-0.15). In Table 7 the result of the two consecutive days are merged to account for the serial correlation. In the middle of the table, the indicators of time and driver fixed effects are found respectively. The regressions include a dummy variable to control for potential demand fluctuations that might effect the labour supply decision: whether it is a high temperature on any of the two days (above 10 degrees Celsius), whether it rains during any of the two the days, and it also controls for weekdays (in pairs) and week of the year.

Table 7: OLS Hours worked Eq. (1), Two-Day Shift

	1	2	3	4	5	6
Log hourly wage	0.1862**	0.214***	0.1767***	0.204***	0.1457**	0.1701**
	(0.0357)	(0.0367)	(0.0357)	(0.0367)	(0.0356)	(0.0366)
Rain			0.00668	0.00658	0.0109	0.0108
			(0.0174)	(0.0174)	(0.0183)	(0.0183)
High temperature					0.0161	0.0159
					(0.0237)	(0.0236)
Driver Fixed Effect	No	Yes	No	Yes	No	Yes
Time Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,765	1,765	1,765	1,765	1,765	1,765
Adjusted R-squared	0.003	0.005	0.007	0.012	0.021	0.035
Constant	2.468***	2.316***	2.489***	2.344***	2.611***	2.481***
	(0.198)	(0.202)	(0.197)	(0.202)	(0.196)	(0.201)
Number of drivers	47	47	47	47	47	47

Note: The table contains OLS regression results estimating Eq. (1), including control variables and estimates with fixed effects (suppressed dummy variables are day of the week and week of the year). Robust standard errors clustered by driver in parentheses and *** p<0.01, ** p<0.05, * p<0.1

The resulting estimated wage elasticity of the six models range from 0.1457 to 0.214 and are all statistically different from zero, with p-values<0.05. Both control variables are insignificant and the fixed effect wage elasticity estimates are slightly higher when including the fixed effect (comparing estimates on equation 1, 3, and 5 to 2, 4 and 6 in Table 7).

When controlling for measurement error, the results are all statistically different from zero (see Table 8 below and *Appendix 4,* Table 16 for first stage). The estimates are clearly higher, which is to be expected if the OLS regression suffers from attenuation bias. The estimated wage elasticity of the six models ranges from 0.251 to 0.381 with p-values<0.01, and all estimates are statistically different from zero. The control variables are insignificant and the fixed effect estimates are slightly higher than those of the standard model. The IV estimates are somewhat larger than the OLS estimates, but the difference is still considerably less than in previous research (Camerer et al., 1997; Farber, 2015)

Table 8: IV Hours worked Eq. (1), Two-Day Shift

	1	2	3	4	5	6
Log hourly wage	0.335***	0.381***	0.310***	0.354***	0.251***	0.292***
	(0.0505)	(0.0523)	(0.0502)	(0.0521)	(0.0506)	(0.0524)
Rain			0.00826	0.00807	0.0123	0.0120
			(0.0157)	(0.0156)	(0.0164)	(0.0163)
High temperature					0.0156	0.0152
					(0.0212)	(0.0211)
Driver Fixed effect	No	Yes	No	Yes	No	Yes
Time Fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,765	1,765	1,765	1,765	1,765	1,765
Constant	1.163***	0.918***	1.268***	1.030***	1.540***	1.324***
	(0.278)	(0.288)	(0.276)	(0.286)	(0.277)	(0.287)
Number of drivers	47	47	47	47	47	47

Note: The table contains IV regression results estimating Eq. (1) with a longer time horizon. The first stage can be found in *Appendix 4*. The regressions include control variables and fixed effects (suppressed dummy variables are week day and week of the year). Robust standard errors clustered on individual drivers in parentheses and *** p<0.01, ** p<0.05, * p<0.1

Additional regressions on extending the time horizon even further are found in *Appendix 4,* Figure 24 – 27. The wage elasticity is now more bound towards zero. However, all estimates (for a three-day and seven-day time-horizon respectively) are statistically different from zero and positive. The precision of the estimates is reduced, largely due to the drop in the number of observations when merging the days. Sensitivity analyses are carried out shifting the starting date to see if coupling of different dates has an impact on the results, (i.e. Tuesday – Wednesday instead of Monday – Tuesday etc.), and the conclusion is that this do not change the result. In addition to supplying more precise estimates of wage elasticity, this section could also be viewed as evidence against reference-dependence since none of the estimates of the wage elasticity is below zero. This is the minimal requirement of the targeting theory, which is clearly not satisfied.

As discussed earlier, the trouble of interpreting the wage elasticities is not present when examining the estimates of the discrete-choice stopping model. Results from when extending the time-horizon are found in Table 9.

Table 9: Probability of Ending a Shift, Linear Probability Model, Two-Day Shifts

	1	2	3	4
Hours	0.0419***	0.0419***	0.0639***	0.0635***
	(0.0009)	(0.0007)	(0.0019)	(0.0019)
Income	0.0000***	0.0001***	-0.0000	-0.0000***
	0.0000	0.0001	(0.0000)	(0.0000)
$Hours^2$			-0.0004***	-0.0004***
			(0.0001)	(0.0002)
$Income^2$			0.0000***	0.0000***
			(0.0000)	(0.0000)
Driver fixed effect	No	Yes	No	Yes
Time fixed effect	Yes	Yes	Yes	Yes
Observations	1,765	1,765	1,765	1,765
Number of drivers	47	47	47	47
Adjusted R-sq	0.8911	0.8911	0.9027	0.9027
Constant	-0.5190***	-0.5172***	-0.5492***	-0.5436***
	(0.0134)	(0.0124)	(0.0200)	(0.0196)

Note: Result of the discrete-choice stopping model (estimations of Eq. (1)). Squared terms are included to allow for differences in functional form. Robust standard errors clustered by driver are in parentheses and *** p<0.01, ** p<0.05, * p<0.1

By extending the Farber (2015) model and examining multiple days, the estimates in Table 9 imply that the driving behaviours of taxicab drivers are largely driven by hours targeting, relative to an income targeting. Thus, the results from Table 4 are robust when expanding the time-horizon. Models $1-4$ predict that if a driver works for one more hour (income held constant), the probability of ending a shift increases by 4 to 6 per cent (neglecting the squared terms). Similarly to Table 5, accumulating income during a shift is not accompanied by a substantial increase in the probability of ending a shift. The estimations are accompanied by a larger fraction of uncertainty and the size of the effect is slightly smaller when incorporating that hours worked is increasing given a two-days span compared to a one-day time horizon. The results from Table 5 are robust for the variation of the time horizon. Furthermore, the estimates are consistent in various specifications and including squared terms do not substantially affect the results. However, the marginal effect of accumulated hours is substantial and well in the region of estimates from previous research. The estimates imply that income targeting is not the decisive factor when examining the choice of ending a shift among taxicab drivers.

Table 10: Marginal Effects of Income and Hours on Probability of Ending a Shift, Probit Model, Two-Day Shift

Income	(1)	Hours	(2)
600-1,999	0.0000***	6-10	0.0009***
	(0.0000)		(0.0003)
2,000-3,999	0.0001***	11-12	0.0241***
	(0.0001)		(0.0006)
4,000-5,999	0.0002***	13-14	0.0353***
	(0.0000)		(0.0004)
6,000-7,999	0.0001***	15-16	0.0162***
	(0.0000)		(0.0004)
8,000-9,999	0.0001***	17-18	0.0229***
	(0.0000)		(0.0002)
>10,000	0.0000***	19-20	0.0369***
	(0.0000)		(0.0005)
		21-22	0.0656**
			(0.0002)
		23-24	0.0968***
			(0.0002)
		25-26	0.0888***
			(0.0002)
		27-28	0.0518***
			(0.0005)
		29-30	0.0209***
			(0.0004)
		>30	0.0019***
			(0.0002)

Note: Based on estimates probit model for the probability of ending a shift (columns 1 and 2). Setting the other variables at their mean values and neglecting the squared terms I obtain the marginal effects. Marginal effects are obtained from Wednesday at 2:00 PM. Robust standard errors clustered by driver are in parentheses and *** p<0.01, ** p<0.05, * p<0.1

Column (1) of Table 10 contains estimates of the effects of accumulated income on an income target obtained from probit regression. As shift income accumulates (conditional on hours at the mean - column (1)), the probability of ending a shift shows no tendency to increase. In addition, there is no economically relevant increase in the probability of ending a shift as income accumulates during the shift.

Column (3) of Table 10 contains the analogous effects for accumulation of hours (conditional on income at the mean) on the probability of ending a shift. The probability of ending a shift is sharply increasing with accumulated hours. The probability of quitting increases by 10 percentage points from early in the shift. Graphical representations of the survival probabilities are found in *Appendix 4* (see Figure 23 – 24).

To conclude, the results from extending the time horizon are in line with the neoclassical theory of labour supply. Not in any specification negative labour supply elasticity estimates are found, and the probability of ending a shift is arguably independent from accumulated income. However, an hourly target is supported, not only within a two-day time horizon, but a three-day or a weekly target can also be inferred from the results in this section (see *Appendix 4*).

6. Findings

This section presents a summary of the findings presented in Section 5 followed by a discussion part which examines the open question of how to model demand, if taxicab drivers are a representative group of worker, and how to incorporate the participation decision.

The result of this paper can be summarized as a confirmation of Farber (2015), leading to the conclusion that reference-dependence preferences do not play a significant role in governing the labour supply decision of Stockholm City taxicab drivers.

When examining the data, the decomposition of the variation shows that only 20 per cent of the total variation in wage is due to unanticipated wage variation. The estimate of wage elasticity when the expected wage is close to actual wage is found to be around -0.06, which implies some reference-dependence when in the region of gains and losses. However, the given the limited scope of the variation the theory can explain this implies that reference-dependence is not a significant factor when taxicab drivers in Stockholm make their labour supply decisions.

No evidence of reference-dependence is found in the discrete-choice stopping model. In various specifications the results suggest no reliance on an income target. The results instead confirm the notion of an hourly target, which is inconsistent with the idea of reference-dependence, but consistent with the neoclassical theory of labour supply. The probit estimation confirms the findings of linear probability models.

The heterogeneity across drivers' behaviours is investigated. The results suggest that there is a large range of differences in driving behaviours across taxicab drivers in the sample. While the majority of drivers have positive wage elasticity estimates there are drivers with significantly negative wage elasticity measures. One possible explanation for the variety could be that drivers learn to be rational when working and absorbing experience. The results show some indications of a learning pattern of taxicab drivers, revealing that drivers with more experience optimize their earnings and inexperienced drivers seem to follow a targeting model. However, the key implication of the analysis is the importance of data quality and sample size when conducting the micro studies of behavioural patterns.

Lastly, the above findings are not sensitive to temporal changes when expanding the time horizon. Instead, the wage elasticity is positive for various model specifications and never negative. This is clearly not supporting the reference-dependence theory of labour supply. The result suggests that drivers make labour supply decisions more than one day at a time.

Discussion

The identification strategy allows this paper to analyse the labour supply decisions of taxicab drivers in multiple dimensions, and the results are in general consistent with the neoclassical theory of labour supply. However, some factors might influence this assessment and below follows a few extensions regarding the modelling of taxicab drivers' labour supply decisions.

The question how to model demand is not completely answered in this paper nor in previous work. Farber (2015) acknowledges that to only use weather conditions as a proxy for demand may not be sufficient. Adverse weather conditions may have two opposite effects. First, it is likely to increase demand given that people will be more reluctant to wait for public transport or walk to the planned destination. On the other hand, there is the danger of traffic congestion and this is likely to decrease the demand for taxicabs (Farber, 2015). In addition, weather may have the supply side effect that drivers find it harder to drive and consequently reduce their hours supplied.

Investigating pizza vendors, Saia (2016) argues that taxicab demand fluctuation cannot be captured by transitory variability in weather conditions since the effect is not prone to be in any distinct direction. Instead Saia (2016) exploits the relationship between Google searches for pizza vendors in Bologna and weather condition and finds support for the notion that if it is bad weather on a given day, the demand for pizza goes up. This allows the paper to use exogenous variation in demand and analyse the labour supply decision using an instrumental variable approach. Saia (2016) argues that the result in the paper is perfectly in line with the standard neoclassical model of labour supply, and when using exogenous demand shifters no trace of targeting can be found.

In the taxicab setting, the use of proxy targets (Crawford and Meng, 2011) and expected wage (Farber, 2015) incorporate demand fluctuation, but the fluctuations are not exogenous by construction. However, Saia (2016) shows that the frequency of Google searches is not correlated with adverse weather conditions during Farber's (2015) sample period in NYC, which implies that weather fluctuations cannot be used in the taxicab market as a source of exogenous variation in demand. Hence, there is still a question how to model demand and what implications this may have on the results of previous research.

Camerer et al. (1997) argues that it is indeed likely that taxicab drivers are not representative of the working population. This is in fact true for many occupations where the workers have flexible hours and variable wages (for example farmers, fishermen, and small-business owners). The workers in these occupations self-select themselves into work with low and uncertain income and long hours, which suggests that this subgroup of worker may systematically differ from the general working population. Consequently, there is a tension between pursuing the ideal occupation, which allows the theory to be tested, and the external validity of the results. Given this trade-off, I find that taxicab drivers are a far better test subgroup than many of the workers chosen by other scholars.

Lobster fishermen and pizza vendors are two occupations that have received increasing attention during the recent years. In neither of the occupations the workers are completely free to set their own working hours. Lobster fishermen work in groups on boats far off the coast and hence when analysing the labour supply decision, the best scenario is that you obtain the estimate of the aggregate mean response of the group, but most likely you will receive the discrete decision of the captain whether or not to end a shift. The captain's choice does not necessarily reflect his own individual preferences with respect to transitory wage fluctuations, but instead involves a more sophisticated analysis of the cost of keeping the boat out fishing with a full crew, fishing quotas, and financial goals decided by investors. The same is likely to be true with respect to pizza vendors, who also work in groups and in a working environment with fixed opening hours that require staff to be present. Indeed, this is not a problem with taxicab drivers who can determine when to drive on their own, without taking other working companions' preferences into account and without a large participation cost (unlike the lobster fishermen and pizza vendors).

Choosing a subgroup of the working population will cause problems regarding the external validity of the results. However, these weaknesses can be justified if the subgroup has features allowing a specific theory to be tested. From Camerer et al. (1997) to Farber (2015) taxicab drivers are still seen as the ideal testing ground for reference-dependence versus the neoclassical theory because of the fact that scholars have argued that taxicab drivers are allowed to choose their own working hours. My findings strengthen the empirical work of Farber (2015), showing that the institutional setting in NYC arguably does not substantially drive his results.

Future Research

One problem with the validity of the above findings and indeed with all studies conducted on taxicab drivers is self-selection into participation. The implication of neglecting the participation decision is that if daily wage fluctuations affect both hours worked and the participation decision in the same direction, it can lead to negative correlation between the wage and the error term in Eq. (1) and create negative bias to the wage elasticity estimate (Heckman, 1979). In general, the wage elasticity in this paper is found to be non-zero and positive, hence the bias is not a problem for the validity of the findings. However, the results in Table 2 and 3 might be exaggerated given that some of the estimates are negative.

This paper does not incorporate the participation decision when conducting the analysis because of the failure to find an appropriate instrument for the Heckman selection model (Tobit type II) but acknowledges the possible effects it may have on the results. The results presented in this paper follow the convention within the literature of taxicab drivers. However, my data allows for an examination of the fraction of workers participating separated on day of the week, and it is found in *Appendix 5*, Table 24. The participation rate is surprisingly constant during the week with one exception: Sundays. There is a large drop of participating drivers on Sundays when on average only half of the drivers in the sample work. Nevertheless, the numbers presented in Table 24 are somewhat reassuring since we do not observe much dispersion across days. A natural extension for future research is to follow Stafford (2013) and take the participation decision into account when examining the wage elasticity.

7. Conclusion

My paper is the first to give external validity to Farber's (2015) findings. My results suggest that only a small fraction of wage variation (about 1/5) is unanticipated implying that reference-dependence plays a limited role in determining the labour supply decisions of taxicab drivers. In addition, the result is confirmed by the fact the probability of ending a shift is positively related to accumulated hours (conditional on accumulated income) while being seemingly unrelated to accumulated income (conditional on accumulated hours). The results are obtained without the institutional constraints associated with the NYC taxicab market (where taxicab drivers work in shifts and tipping is a large share of the income and is not reported).

To my knowledge, this paper is the first to analyse LBD behaviours of taxicab drivers with actual information on the tenure of the drivers. Some evidence of LBD behaviour is found, but the result should be viewed with caution (because of limited sample size). Lastly, when extending the time horizon, some evidence for a longer planning horizon is found and precise wage elasticity estimates are positive and ranging from 0.20-0.35. This implies that there is room for theoretical and empirical developments modelling taxicab drivers' labour supply decisions with a longer time frame.

8. References

Altonji, J. G. (1986). Intertemporal Substitution in Labour Supply: Evidence from Micro Data. *Journal of Political Economy*, 94, 176–215.

Abeler, J., Falk, A., Goette, L., & Huffman, D. (2011). Reference Points and Effort Provision. *American Economic Review*, 101(2), 470-92.

Agarwal, S., Diao, M., Pan, J., & Sing, T. S. (2015). Labour Supply Decisions of Singaporean Cab Drivers. *Working paper* March 2015. (Retrieved April 15, 2016 from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2338476).

Barberis, N. C. (2013). Thirty Years of Prospect Theory in Economics: A Review and Assessment. *Journal of Economic Perspectives*, 27(1), 173–196.

Browning, M., Deaton, A., & Irish, M. (1985). A Profitable Approach to Labour Supply and Commodity Demands over the Life-Cycle. *Econometrica*, LIII, 503–543.

Camerer, C. (2000). Prospect Theory in the Wild: Evidence from the Field in Choices, Values, and Frames. (pp. 288-300). Kahneman & Tversky (Eds). Cambridge UK: Cambridge University Press.

Camerer, C., Babcock, L., Loewenstein, G., & Thaler, R. (1997). Labor Supply of New York City Cab Drivers: One Day at a Time. *Quarterly Journal of Economics*, 112(2), 407–441.

Chang, T., & Gross, T. (2014). How many pears would a pear packer pack if a pear packer could pack pears at quasi-exogenously varying piece rates? *Journal of Economic Behaviour and Organization,* 99, 1-17.

Chetty, R., Guren, A., Manoli, D., & Weber, A. (2011). Are micro and macro labour supply elasticities consistent? A review of evidence on the intensive and extensive margins. *American Economic Review*, 101(3), 471-475.

Chou, Y. K. (2002). Testing Alternative Models of Labour supply: Evidence from Taxi Drivers in Singapore. *Singapore Economic Review*, 47(1), 17–42.

Crawford, V. P., & Meng, J. (2011). New York City Cab Drivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income. *American Economic Review*, 101(5), 1912–1932.

Doran, K. (2014). Are Long-Term Wage Elasticities of Labor Supply More Negative than Short-Term Ones? *Economic Letters* 122, 208-210.

Dupas, P., Robinson, J., & Savedra, S. (2016) The Daily Grind: Cash Needs, Labour Supply and Self-Control. *Working paper* revised March 2016 (Retrieved April 15, 2016 from http://web.stanford.edu/~pdupas/DRS_TheDailyGrind.pdf.).

Farber, H. S. (2005). Is Tomorrow Another Day? The Labour Supply of New York City Cabdrivers. *Journal of Political Economy*, 113(1), 46–82.

Farber, H. S. (2008). Reference-Dependent Preferences and Labour Supply: The Case of New York City Taxi Drivers. *American Economic Review*, 98(3), 1069–1082.

Farber, H. S. (2015). Why You Can't Find a Taxi in the Rain and other Labour Supply Lessons from Cab Drivers. *Quarterly Journal of Economics*, 130(4), 1975-2026.

Fehr, E., & Goette, L. (2007). Do Workers Work More if Wages Are High? Evidence from a Randomized Field Experiment. *American Economic Review*, 97(1), 298–317.

French, S. T., & Stafford, T. M. (2015). Experience and the Wage Elasticity of Labour Supply. *Working paper* revised August 2015 (Retrieved March 3, 2016 from http://research.economics.unsw.edu.au/tstafford/Research_files/Stafford_EXP.pdf).

Haggag, K., McManus, B., & Paci, G. (2016). Learning by Driving: Productivity Improvements by New York City Taxi Drivers. Accepted at *American Economic Journal: Applied Economics* 2016.

Haggag, K., McManus, B., & Paci, G. (2014). Default Tips. *American Economic Journal: Applied Economics,* 6(3), 1-19.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error *Econometrica*, 47(1), 153-161.

Jonanson, D., & Wållgren, J. (2013). Labor supply of Stockholm cab drivers: revisiting the evidence. (Retrieved November 23, 2015 from http://arc.hhs.se/download.aspx?mediumid=1957).

Kahneman, D., & Tversky A. (1979). Prospect Theory: An Analysis of Decision Under Risk. *Econometrica*, 47(2), 263–292.

Kahneman, D., & Tversky, A. 1991. Loss Aversion in Riskless Choice: A Reference-Dependent Model. *Quarterly Journal of Economics,* 106(4), 1039–1061.

Keane, M. P. (2011). Labour supply and taxes: A survey. *Journal of Economic Literature*, *49*(4), 961-1075.

Kimball, M. S., & Shapiro, M. D. (2008). Labour supply: Are the income and substitution effects both large or both small? *National Bureau of Economic Research,* (No. w14208, 2008).

Kőszegi, B., & Rabin, M. (2006). A Model of Reference-Dependent Preferences. *Quarterly Journal of Economics*, 121(4), 1133–1165.

Kőszegi, B., & Rabin, M. (2007). Reference-dependent risk attitudes. *American Economic Review*, 97(4), 1047-1073.

Kőszegi, B., & Rabin, M. (2009). Reference-dependent consumption plans. *American Economic Review*, 99 (3), 909-936.

Lucas, R. E. Jr., & Rapping, L. A. (1969). Real Wages, Employment, and Inflation. *Journal of Political Economy*, 77 (5), 721–754.

Mankiw, G. N., Rotemberg, J. J., & Summers, L. H. (1986). Intertemporal Substitution in Macroeconomics. *Quarterly Journal of Economics*, 100(1), 225-251.

Nguyen, Q., & Leung, P. (2009). Revenue Targeting in Fisheries: The Case of Hawaii Longline Fishery. *MPRA Paper* (No. 13846, March 7, 2009). (Retrieved March 14, 2016 from http://mpra.ub.uni-muenchen.de/13846).

Oettinger, G. S. (1999). An Empirical Analysis of the Daily Labour Supply of Stadium Vendors. *Journal of Political Economy*, 107(2), 360–392.

Read, D., Loewenstein, G., & Rabin, M. (1999). Choice Bracketing. *Journal of Risk and Uncertainty*, 19, 171–197.

Saia, A. (2016). Wishing for Bad Weather: Demand Shocks and the Labour Supply Behaviour of Bologna Pizza Vendors. *Working paper* revised March 2016. (Retrieved April 22, 2016 from https://www.dropbox.com/s/6q6ypxnkx7yhk1q/saia_march16.pdf?dl=0).

Stafford, T. (2013). Labour Supply of Fishermen: An Empirical Analysis. *Australian Agricultural and Resource Economics Society, 56th Conference.* Freemantle, Australia. 7-10 February 2012. (Retrieved January 3, 2016 from http://faculty.smu.edu/millimet/classes/eco7321/papers/stafford%202013.pdf).

Thompson, P. (2010). Learning by Doing. *Handbook of Economics of Technical Change*. Bronwyn & Rosenberg (Eds). Chapter 11. North-Holland/Elsevier.

Thomas, P. (2012). The Relationship Between Unit Cost and Cumulative Quantity and the Evidence for Organizational Learning-by-Doing. *The Journal of Economic Perspectives,* 26(3), 203– 224.

9. Data

Weather data is publicly available at SMHI: (Retrieved Mars 4, 2013 from http://opendata-download-metobs.smhi.se/explore/)

Appendix 1: Data Composition and Sample Characteristics

Data Set

This paper is based on a set of data collected at two different points in time. The main part of the data set is obtained in Mars 2013 and includes trip-sheets of 47 taxicab drivers during a time span of 3 months, from the 1st of April to the 30th of June 2012. The drivers in the sample do not share their taxi vehicles with other drivers, which make them able to choose working hours that reflect their personal preferences. The collection of data was conducted by the franchise company, ensuring that the drivers were not sharing taxicabs. The data set includes trip-by-trip information, daily income, hours worked, and breaks (which is only observed if a driver removes his or her identification card from the software). Additional information was gathered in April 2016, including for how many years the driver has been actively working for the franchise company and for how many years the driver has been able to work in a taxicab profession (authorized with a taxicab license). This data set was eventually merged with a data set containing daily weather conditions (rainfall and temperature), collected from Sweden's Meteorological and Hydrological Institute (SMHI).

In essence, the screening process of the data is based on Jonasson and Wållgren's (2013) work and follows the guidelines formulated by Farber (2005; 2015) to confirm the internal consistency of the data. The first phase is to make sure that no shift starts before the previous shift has ended and to reveal other similar spurious features of the data. The data exhibited no internal inconsistency, which is anticipated since the data is collected from an electronic log system. Observations being abnormally long (5 observations) or short (2 observations) were excluded, which did not alter the result in any substantial way. The outliers are explained by software malfunction or drivers keeping the taximeter on when using the car off-duty (Jonasson and Wållgren, 2013).

The sample contains observations of drivers working multiple shifts within one single day and also short breaks during a shift. In DAY1 all breaks and gaps between logged shifts are excluded. 94 observations are altered in this manner, leaving 3530 observations in the set. DAY1 is used for the main analysis in section 5.

Sample Characteristics

Figure 9: Active Drivers on a Given Day

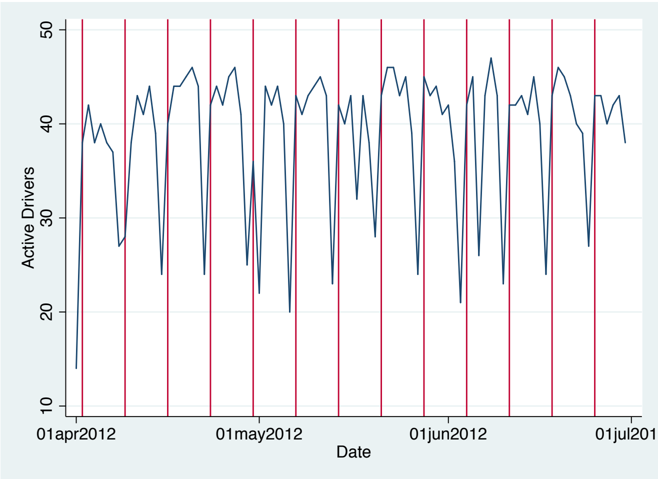


Figure 10: Conditional and Unconditional Mean of Hours Worked

Figure 11: Conditional and Unconditional Median of Hours Worked

Figure 12: Conditional and Unconditional Mean Wage per Hour

Figure 13: Conditional and Unconditional Median of Wage per Hour

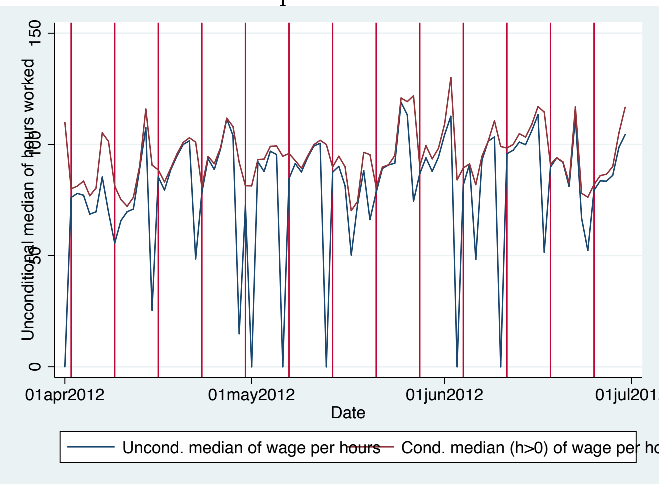


Figure 14: Active Drivers, by Weekday

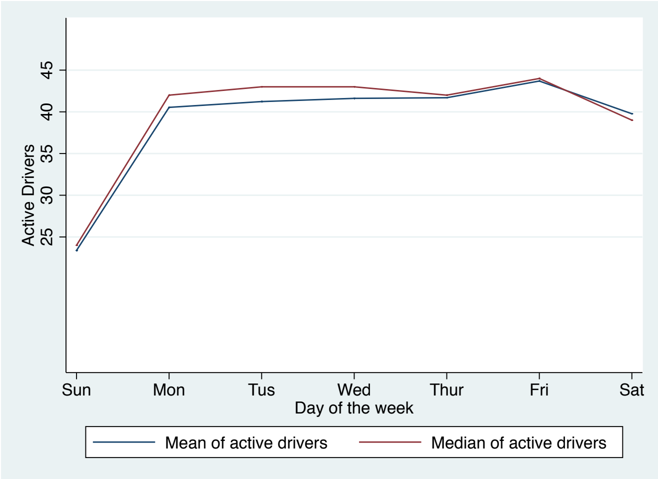


Figure 15: Unconditional and Conditional Mean of Hours Worked, by Weekday

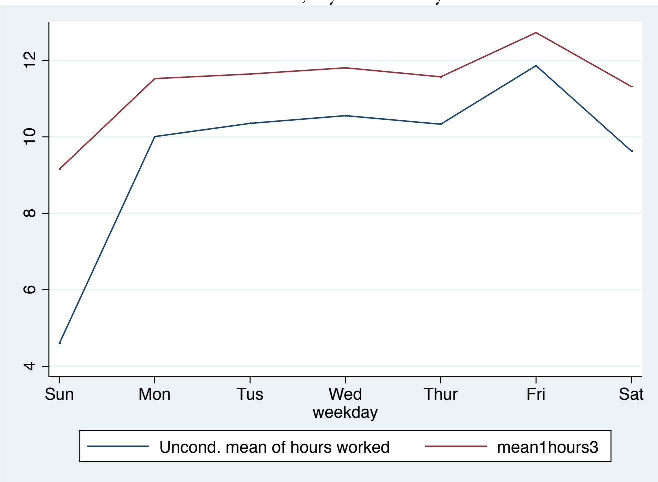


Figure 16: Unconditional and Conditional Median of Hours worked, by Weekday

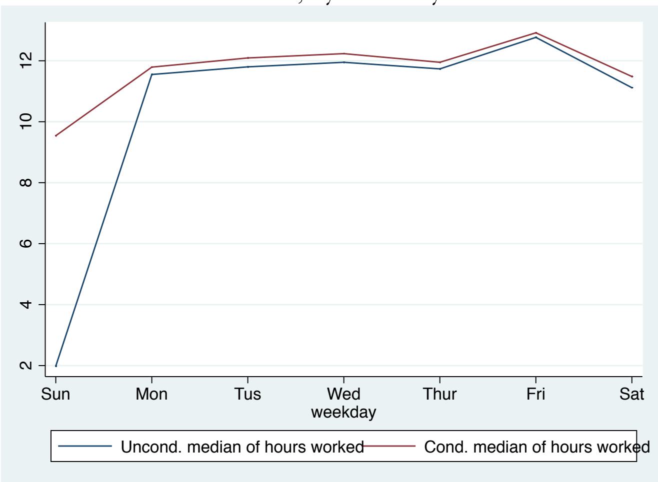


Figure 17: Unconditional and Conditional Mean Wage per Hour, by Weekday

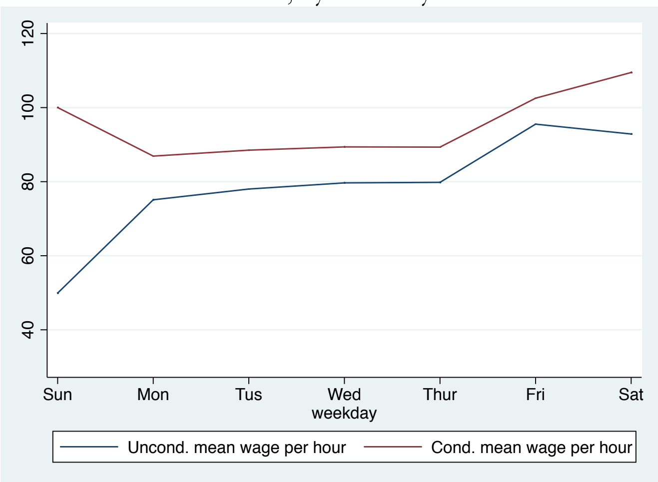


Figure 18: Unconditional and Conditional Median Wage per Hour, by Weekday

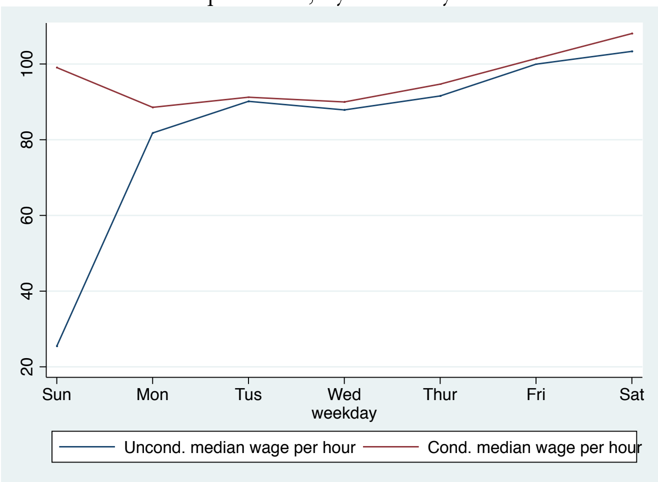


Figure 19: Mean and Median of Active Drivers, by Week

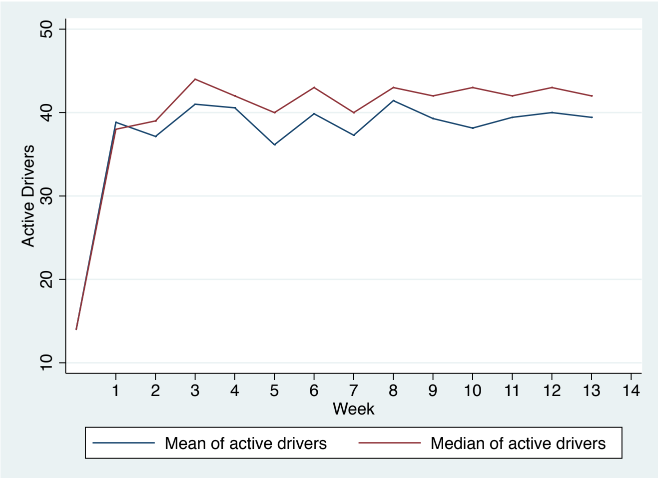
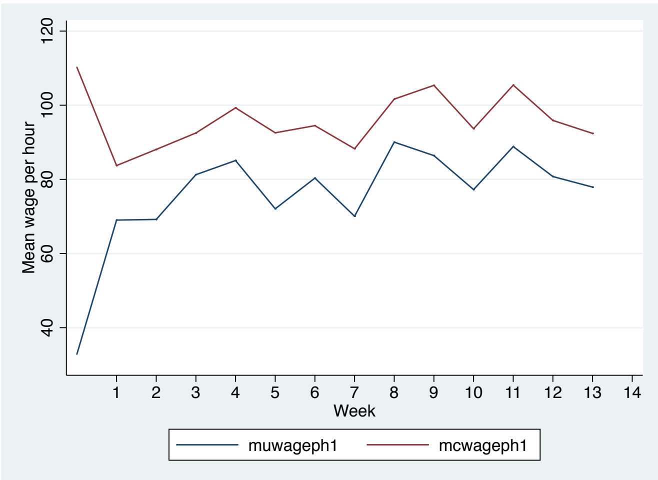


Figure 20: Unconditional and Conditional Mean of Hours Worked, by Week

Figure 21: Unconditional and Conditional Mean Wage per Hour, by Week



Appendix 2: Mathematical Derivation

Following Farber (2015), I set up a simple model of labour supply with reference-dependent preferences. Agents receiving wage rate W extract utility from income ($Y = Wh$) and disutility from hours worked (h). Agents have kinked utility function at some reference level of income (T):

$$U = \begin{cases} U(Y, h) = (1+\alpha)(Y-T) - \frac{\theta}{1+\nu} h^{1+\nu} & Y < T \\ U(Y, h) = (1-\alpha)(Y-T) - \frac{\theta}{1+\nu} h^{1+\nu} & Y \geq T \end{cases} \qquad (1A-2A)$$

where the parameter $\alpha > 0$ controls the change in marginal utility at the reference point, θ is a parameter indicating the disutility of hours, and ν is a parameter related to the elasticity of labour supply. Farber (2015) notes that Eq. (1A – 2A) follows the model of Köszegi and Rabin (2006), specifying neoclassical utility function ($(Y - T) - \frac{\theta}{1+\nu} h^{1+\nu}$) which is extended with a "gain-loss" component ($\pm\alpha(Y-T)$).

A strict neoclassical model is the special case when $\alpha = 0$, implying that there is no "gain-loss" utility (then Eq. (1A) is equal to Eq. (2A)). The labour supply function in this case is $h = (\frac{W}{\theta})^{\frac{1}{\nu}}$, which implies the elasticity of labour supply to be $\frac{1}{\nu}$.

Maximizing the utility function in Eq. (1A – 2A) with respect to hours of work gives three distinctive labour supply functions depending on the wage all derived in a model with reference-dependent preferences.

1. For sufficiently low wages ($W < W^*$), the reference point is irrelevant because the hours required to reach the target at this moderate wage yield excessive disutility of hours worked. The optimal behaviour is to end the shift within the region of high marginal utility section (Eq. (1A)), which is before the target is reached. In this region, the labour supply function is neoclassical:

   $$h = (\frac{(1+\alpha)W}{\theta})^{\frac{1}{\nu}} \qquad (3A)$$

   with elasticity of labour supply $\frac{1}{\nu} > 0$

2. For intermediate wage levels ($W^* < W < W^{**}$), it is optimal to end a shift when the reference income level is reached. The reason is that the wage is sufficiently high to motivate working when marginal utility is high ($Y < T$) but the wage rate is inadequate to stimulate working when marginal utility is low ($Y \geq T$). In this region, the agent is a "income-targeters" in the original Camerer et al. (1997) meaning with the labour supply function

$$h = \frac{T}{W} \tag{4A}$$

and elasticity of labour supply is -1.

3. For sufficiently high wages ($W > W^{**}$), the reference point is irrelevant due to the fact that the wage is sufficiently high so that the optimal behavioural response is to continue working at the low marginal utility section (Eq. (2A)). In this region, the labour supply function is neoclassical.

$$h = (\frac{(1-\alpha)W}{\theta})^{\frac{1}{\nu}} \tag{5A}$$

with elasticity of labour supply $\frac{1}{\nu} > 0$.

To summarize, a model with reference-dependent preferences do not necessarily have to produce negative estimates of wage elasticity. It is only in certain regions around the target that it is relevant, and the prediction of the theory is negative wage elasticity.

Appendix 3: Anticipated vs Unanticipated Wage Variation

Table 11: The Average Daily Mean Hourly ln(Wage) Regressed on Temporal Factors

Monday	-0.1297***
	(0.0074)
Tuesday	-0.1096***
	(0.0074)
Wednesday	-0.1001***
	(0.0074)
Thursday	-0.1027***
	(0.0074)
Friday	0.0218***
	(0.0073)
Saturday	0.0925***
	(0.00743)
Week (absorb)	0.0554***
	(0.0084)
Month (absorb)	0.0012
	(0.0213)
Observations	3530
Adjusted R-squared	0.7545
Constant	4.5947***
	(0.0059)
Number of Drivers	47

Note: The table contains regression of the average daily mean hourly wage on day of the week, week of the year and month. The predicted value of the regression is the anticipated wage.

Table 12: Absolute Deviation from its Expected Values

	Per centile
25%	0.0012
50%	0.0058
99%	0.1815

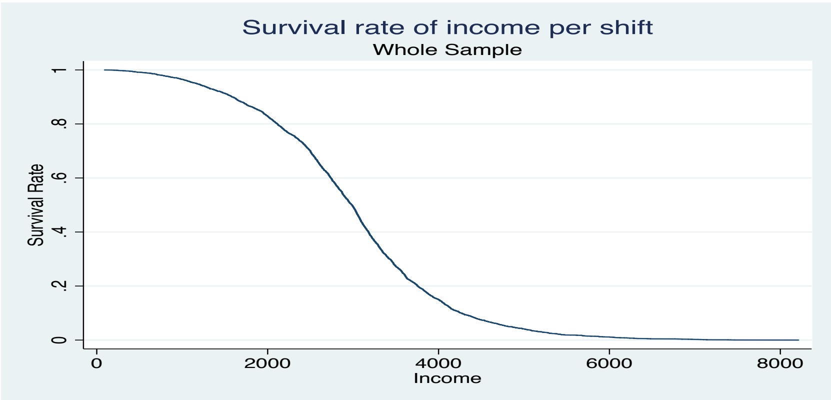
Note: The table includes the absolute deviation from the difference between predicted values in Table 11 and the actual values, hence the residuals.

Table 13: First Stage for Table 4

Sample	Other Drivers Measures
0-25	1.0576***
(N=883)	(0.02315)
25-50	1.0348***
(N=882)	(0.0236)
50-100	1.2737***
(N=1765)	(0.02117)

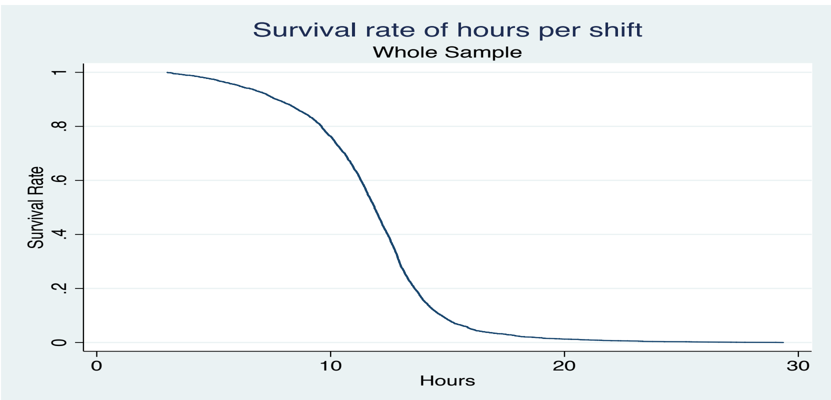
Note: The table includes first stage estimation of lnW_{it} regressed on other drivers reported lnW_t working the same day, while not working overlapping hours. In addition the first stage include the controls from the second stage as their own instrument.

Figure 22: Survival Rate of Income Per Shift



Note: The figure displays the unconditional probability of ending a shift when a certain cumulative income has been earned during a shift.

Figure 23: Survival Rate of Hours Worked



Note: The figure displays the unconditional probability of ending a shift after cumulative hours worked during a shift

Appendix 4: Time Horizon

Table 14: First Stage for Table 8, Two-Day Shift

	1	2	3	4	5	6
Other Drivers Wage	0.9653***	0.9378***	0.9743***	1.0432 ***	1.0467***	1.0609***
	(0.0185)	(0.01836)	(0.0186)	(0.0185)	(0.0188)	(0.0186)
Observations	1,765	1,765	1,765	1,765	1,765	1,765

Note: First stage for estimation in Table 7 (Section 5). The reported wage of other drivers driving the same day during a two-day time window is used as an instrument for individual wage. The control variables are used as an instrument on themselves, not reported above.

Table 15: OLS Eq. (1), Three-Day Shift

	1	2	3	4	5	6
Log hourly wage	0.0553	0.0863*	0.0581	0.0931**	0.0308	0.0654
	(0.0444)	(0.0459)	(0.0441)	(0.0456)	(0.0442)	(0.0458)
Rain			0.0737***	0.0736***	0.0984***	0.0976***
			(0.0178)	(0.0177)	(0.0186)	(0.0185)
High temperature					0.101***	0.0981***
					(0.0232)	(0.0231)
Driver fixed effect	No	Yes	No	Yes	No	Yes
Time fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,167	1,167	1,167	1,167	1,167	1,167
Adjusted R-squared	0.003	0.004	0.007	0.016	0.021	0.029
Constant	3.038***	2.867***	2.988***	2.796***	3.107***	2.918***
	(0.246)	(0.253)	(0.245)	(0.252)	(0.245)	(0.252)
Number of drivers	47	47	47	47	47	47

Note: The table contains OLS regression results estimating Eq. (1) given a three-day shift, including control variables and estimates with fixed effects. Standard errors in parentheses and *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 16: IV Log Hours Worked Equation, Three-Day Shift

	1	2	3	4	5	6
Log hourly wage	0.358***	0.407***	0.326***	0.372***	0.280***	0.324***
	(0.0744)	(0.0773)	(0.0731)	(0.0759)	(0.0743)	(0.0773)
Rain			0.0812***	0.0821***	0.100***	0.100***
			(0.0165)	(0.0164)	(0.0172)	(0.0171)
High temperature					0.0784***	0.0751***
					(0.0218)	(0.0218)
Driver fixed effect	No	Yes	No	Yes	No	Yes
Time fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,167	1,167	1,167	1,167	1,167	1,167
Constant	1.421***	1.158***	1.563***	1.312***	1.791***	1.553***
	(0.410)	(0.426)	(0.403)	(0.418)	(0.408)	(0.425)
Number of drivers	47	47	47	47	47	47
First Stage						
Other drivers wage	0.5225***	0.4896***	0.5313***	0.4939***	0.5299***	0.4885***
	(0.0202)	(0.0196)	(0.0203)	(0.0196)	(0.0207)	(0.0199)

Note: The table contains IV regression results when estimating Eq. (1) for a three-day shift including control variables and estimates with fixed effects. Standard errors in parentheses and *** p<0.01, ** p<0.05, * p<0.1

Table 17: OLS Log Hours Worked Equation, Seven-Day Shift

	1	2	3	4	5	6
Log hourly wage	0.0200	0.136	0.0199	0.137	-0.0380	0.0780
	(0.176)	(0.205)	(0.175)	(0.204)	(0.176)	(0.208)
Rain			0.00863	0.00614	0.0158	0.0131
			(0.0186)	(0.0186)	(0.0188)	(0.0188)
High temperature					0.0719	0.0678
					(0.0252)	(0.0250)
Driver fixed effect	No	Yes	No	Yes	No	Yes
Time fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	503	503	503	503	503	503
Constant	4.090***	3.453***	4.085***	3.444***	4.390***	3.754***
	(0.974)	(1.133)	(0.967)	(1.127)	(0.971)	(1.146)
Number of drivers	47	47	47	47	47	47

Note: The table contains OLS regression results of estimating Eq. (1) for a seven-day shift including control variables and estimates with fixed effects. Standard errors in parentheses and *** p<0.01, ** p<0.05, * p<0.1

Table 18: Wage Elasticity, IV Eq. (1), Seven-Day Shift

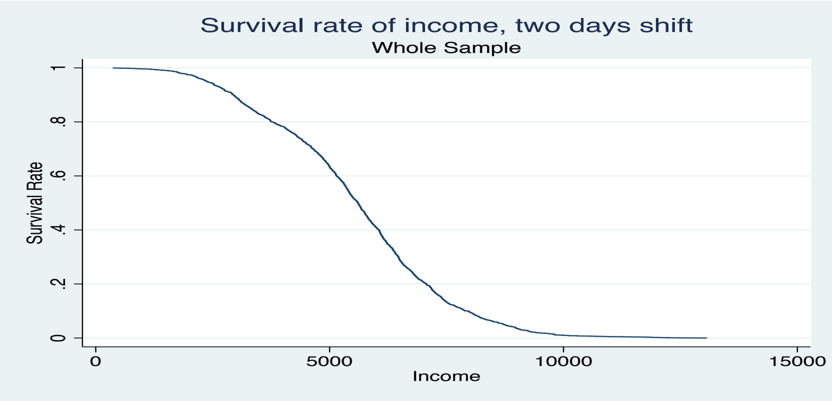
	1	2	3	4	5	6
Log hourly wage	0.152**	0.263***	0.145**	0.261***	0.121*	0.242***
	(0.0687)	(0.0744)	(0.0689)	(0.0749)	(0.0686)	(0.0748)
Rain			0.00806	0.00553	0.0141	0.0112
			(0.0182)	(0.0180)	(0.0183)	(0.0181)
High temperature					0.0685***	0.0643***
					(0.0247)	(0.0244)
Driver fixed effect	No	Yes	No	Yes	No	Yes
Time fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	503	503	503	503	503	503
Adjusted R-squared	0.003	0.004	0.007	0.016	0.021	0.029
Constant	3.361***	2.747***	3.392***	2.757***	3.514***	2.849***
	(0.380)	(0.411)	(0.381)	(0.413)	(0.379)	(0.412)
Number of drivers	47	47	47	47	47	47
First Stage						
Other drivers wage	0.240***	0.214***	0.242***	0.216***	0.2384***	0.2119***
	(0.0219)	(0.0217)	(0.0218)	(0.0215)	(0.0220)	(0.0217)

Note: The table contains IV regression results estimating Eq. (1) for seven-day shift including control variables and estimates with fixed effects. Standard errors in parentheses and *** p<0.01, ** p<0.05, * p<0.1

Figure 24: Survival Rate of Hours, Two-Day Shift

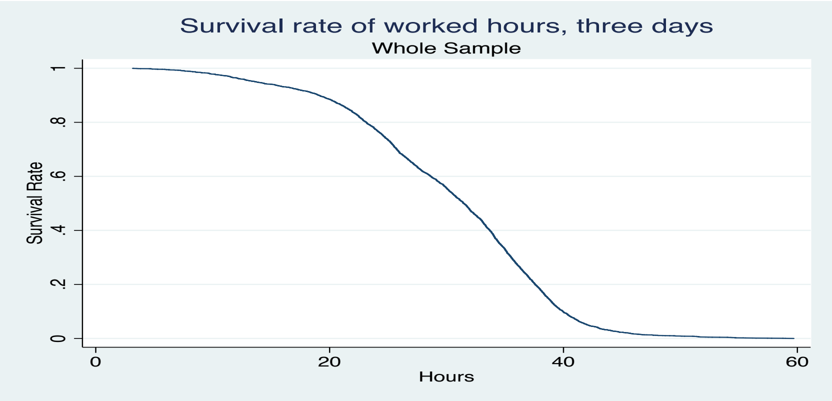
Note: The figure displays the unconditional probability of ending a shift after cumulative hours worked during a two-day shift.

Figure 25: Survival Rate of Income, Two-Day Shift



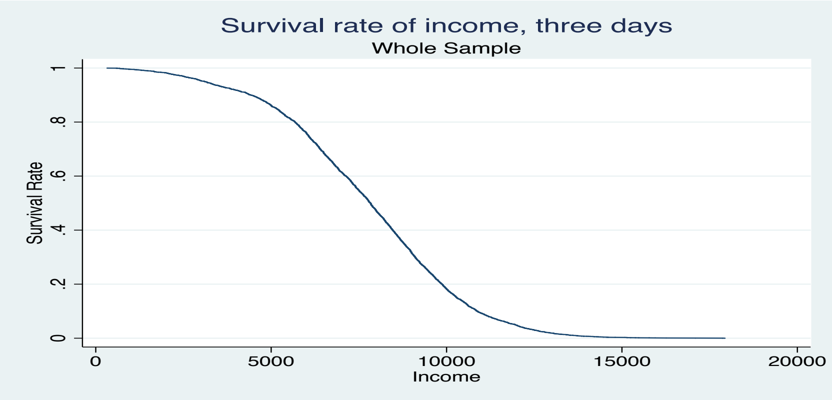
Note: The figure displays the unconditional probability of ending a shift when a certain cumulative income has been earned during a two-day shift.

Figure 26: Survival Rate of Hours, Three-Day Shift



Note: The figure displays the unconditional probability of ending a shift after cumulative hours worked during a three-day shift.

Figure 27: Survival Rate of Income, Three-Day Shift



Note: The figure displays the unconditional probability of ending a shift when a certain cumulative income has been earned during three-day shift.

Table 19:Probability of Ending a Shift, Linear Probability Model, Three-day Shift

Hours Target				
	1	2	3	4
Hours	0.0322***	0.0321***	0.0261***	0.0257***
	(0.0003)	(0.0004)	(0.0014)	(0.0015)
Income	0.0000	0.0000	-0.0000***	-0.0001***
	0.0000	0.0000	(0.0000)	(0.0000)
$Hours^2$			-0.0001***	-0.0001***
			(0.0001)	(0.0000)
$Income^2$			0.0000***	0.0000***
			(0.0000)	(0.0000)
Driver fixed effect	No	Yes	No	Yes
Time fixed effect	Yes	Yes	Yes	Yes
Observations	1,167	1,167	1,167	1,167
Number of drivers	47	47	47	47
Adjusted R-sq	0.9305	0.9399	0.9509	0.9510
Constant	-0.4820***	-0.4831***	-0.2678***	-0.2579***
	(0.0079)	(0.0079)	(0.0148)	(0.0150)

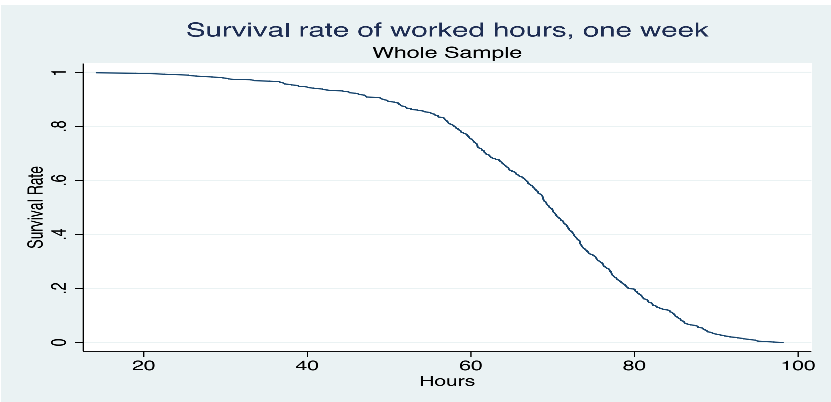
Note: Result of the discrete-choice stopping model (estimations of Eq. (2)), with a three-day shift. I include squared terms to allow for differences in functional form. Robust standard errors clustered by driver are in parentheses and *** p<0.01, ** p<0.05, * p<0.1

Table 20: Marginal Effects of Income and Hours on Probability of Ending Shift, Probit Model, Three-Day Shift

Income	(1)	Hours	(2)
1,000-2,499	0.0001***	10-14	0.0082***
	(0.0001)		(0.0001)
2,500-4,999	0.0002***	15-19	0.0089***
	(0.0001)		(0.0001)
5,000-7,499	0.0002***	20-24	0.0269***
	(0.0000)		(0.0002)
7,500-9,999	0.0003***	25-29	0.0462***
	(0.0001)		(0.0004)
10,000-12,999	0.0001***	30-34	0.0489***
	(0.0001)		(0.0001)
>13,000	0.0001***	35-39	0.0500***
	(0.0001)		(0.0001)
		40-44	0.1022**
			(0.0003)
		>45	0.0867***
			(0.0002)

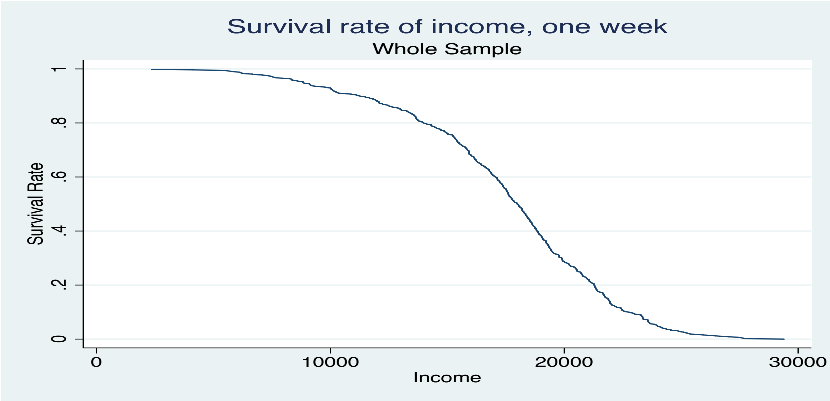
Note: Based on estimates of two probit models (columns 1 and 2). The marginal effects in each model are obtained holding the other variables at its mean values. Both models additionally include sets of fixed effects for drivers, hour of the day by day of the week. Robust standard errors clustered by driver are in parentheses and *** p<0.01, ** p<0.05, * p<0.1

Figure 28: Survival Rate of Hours, Seven-Day Shift



Note: The figure displays the unconditional probability of ending a shift after cumulative hours worked during a one-week shift.

Figure 29: Survival Rate of Income, Seven-Day Shift



Note: The figure displays the unconditional probability of ending a shift when a certain cumulative income has been earned during a one-week shift.

Table 21: Probability of Ending a Shift, Linear Probability Model, Seven-Day Shift

	1	2	3	4
Hours	0.0119***	0.0112***	0.0212***	0.0208***
	(0.0006)	(0.0006)	(0.0014)	(0.0015)
Income	0.0000***	0.0000***	-0.0000***	-0.0001***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$Hours^2$			-0.0001***	-0.0001***
			(0.0000)	(0.0000)
$Income^2$			0.0000***	0.0000***
			(0.0000)	(0.0000)
Driver fixed effect	No	Yes	No	Yes
Time fixed effect	Yes	Yes	Yes	Yes
Observations	503	503	503	503
Number of drivers	47	47	47	47
Adjusted R-sq	0.7635	0.7542	0.8608	0.8923
Constant	-0.4121***	-0.4074***	-0.1212***	-0.1263***
	(0.0230)	(0.0218)	(0.0235)	(0.0241)

Note: Result of the discrete-choice stopping model (estimations of Eq. (1)) with seven-day shift. I include squared terms to allow for difference in functional form. Robust standard errors clustered by driver are in parentheses and *** p<0.01, ** p<0.05, * p<0.1

Appendix 5: Findings

Table 22: Participation Rate – Summary Statistics

	Rate	Std.
Whole Sample	0.8279	0.3774
Monday	0.8625	0.3446
Tuesday	0.8772	0.3284
Wednesday	0.8871	0.3167
Thursday	0.8887	0.3147
Friday	0.9296	0.2559
Saturday	0.8494	0.3579
Sunday	0.5008	0.5004

Note: Participation rate of drivers in the sample taxicab drivers.