Title
Essays on Labor Supply and Uncertainty

Author
Leah-Martin, Vincent

Publication Date
2017

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Essays on Labor Supply and Uncertainty

A dissertation submitted in partial satisfaction of the
requirements for the degree of Doctor of Philosophy

in

Economics

by

Vincent Andrew Leah-Martin

Committee in charge:

Professor James Andreoni, Chair
Professor Julie Berry Cullen
Professor Gordon Dahl
Professor Craig McKenzie
Professor Marta Serra-Garcia
Professor Charles Sprenger

2017

The Dissertation of Vincent Andrew Leah-Martin is approved and is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2017

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

For Chapter 1, I would like to thank Jim Andreoni, Charlie Sprenger, and Gordon Dahl for their ongoing support and guidance. I am also extremely grateful for comments and critiques from Vincent Crawford, Henry Farber, Julie Cullen, Joel Sobel, Mark Machina, and Isabel Trevino. I would like to thank my colleagues, in particular David Coyne and Isla Globus-Harris for their insights and help with editing this paper. This paper benefited from excellent research assistance provided by Ellen Kim. Lastly, I am very grateful to Andreas Walters for input related to the taxi industry and its underlying data.

For Chapter 2, I am extremely grateful to Glassdoor for supplying the earnings and company review data used in this research. In particular, Andrew Chamberlain was instrumental in facilitating acquisition of the data and providing helpful comments and feedback throughout the analysis. This paper has also greatly benefits from helpful comments and suggestions made by Gordon Dahl and David Coyne.

For Chapter 3, I am very grateful to Charles Sprenger and Marta Serra-Garcia for early comments on the experimental design used in this paper. Additionally, James Andreoni and Mitch Downey provided extremely helpful suggestions for the analysis used. I would also like to thank Amanda Brainerd for helping to coordinate running individual subjects independently of one another. Lastly, I would like to thank the UC San Diego Department of Economics for providing funding for this project.

Chapter 1, in full, has been submitted for publication of the material as it may appear in the Journal of Economic Behavior & Organization, 2016, Leah-Martin, Vincent. The dissertation author was the sole investigator and author of this paper.

Chapter 2, in part is currently being prepared for submission for publication of the material. Leah-Martin, Vincent. The dissertation author was the sole investigator and author of this material.

Chapter 3, in part is currently being prepared for submission for publication of the material. Leah-Martin, Vincent. The dissertation author was the sole investigator and author of this material.

VITA

2012	Bachelor of Arts, University of California, San Diego
2011–2017	Teaching Assistant, Department of Economics University of California, San Diego
2013	Master of Arts, University of California, San Diego
2017	Doctor of Philosophy, University of California, San Diego

FIELDS OF STUDY

Major Field: Economics

Studies in Labor Economics
Professors Julie Berry Cullen and Gordon Dahl

Studies in Experimental & Behavioral Economics
Professors James Andreoni, Craig McKenzie, Marta Serra-Garcia, and Charles Sprenger

ABSTRACT OF THE DISSERTATION

Essays on Labor Supply and Uncertainty

by

Vincent Andrew Leah-Martin

Doctor of Philosophy in Economics

University of California, San Diego, 2017

Professor James Andreoni, Chair

The content of this dissertation focuses on the link between earnings risk and the preferences of workers in the labor market. The research papers which comprise the chapters of the dissertation all examine worker responses to short-run earnings variation. Chapter 1 addresses the question of whether or not behavior consistent with having risk preferences around a daily income target can explain observed behavior in taxi drivers. Chapter 2 examines hows the outcomes of a wage distribution can influence job satisfaction. Lastly, chapter 3 utilizes an experiment to establish a link between estimated risk attitudes and labor supply outcomes.

Chapter 1

When to Quit: Narrow Bracketing and Reference Dependence in Taxi Drivers

1.1 Introduction

The labor supply of taxi drivers is perhaps one of the most studied and controversial topics in research on labor supply. A very rich and highly regarded literature has examined taxi driver decisions primarily using trip-level data for individual drivers. The focus on taxi drivers stems from their labor supply environment being ideal to test implications of competing labor supply models.

Two models of labor supply have stood out to be the main focus of previous literature. The first is the standard, or Neo-Classical, model of labor supply. In this model, workers who face varying wages intertemporally substitute by working more to capitalize on days when wages are high and working less when wages are lower. This model implies that drivers would have a positive wage elasticity for short-term labor supply and that drivers are risk neutral over short time periods.

On the other hand is a model of driver labor supply based on reference dependence, specifically, loss aversion. In a model of labor supply under loss aversion, drivers have short-term income targets referred to as reference points. A driver's utility is linear in this model except for a kink at the reference point. Thus, the reference point in this model generates short-term risk aversion within a given time-frame, typically modeled as a day. This model implies a negative wage elasticity over short time frames.

The time frame over which a worker evaluates his earnings is typically referred to as a “bracket”. Evaluating outcomes over a short time-frame is a phenomenon known as “narrow bracketing”. Reference dependent preferences are often associated with narrow bracketing (Read et al., 1999). An individual who narrowly brackets their choices only evaluates outcomes within a short time-frame. In the context of labor supply, this could mean decisions and utility are evaluated within the context of a day’s labor supply.[1] Combining risk preferences, bracketing, and reference dependence in a model of decision-making yields additional intuition for interpreting a reference point, as a short-term goal, while mitigating calibrational criticisms of utility curvature (see Rabin (2000)).

Understanding which model better explains driver behavior and, more generally, behavior under flexible labor supply has several important implications. First, for theory, having risk preferences over short-term outcomes would violate a life-cycle model of labor supply and suggests a limited role for intertemporal substitution. From a practical perspective, risk preferences over short-term earnings have policy implications as they would imply markets might be made more efficient by the introduction of insurance for low, but highly variable, wage workers. Lastly, such preferences might also be used to form more preferred labor contracts which account for the nature of worker’s preferences.

Taxi drivers operate in an ideal setting to test for implications of narrowly bracketed risk preferences in labor supply. Drivers have a high degree of flexibility in their work schedules and face large degrees of variation in their hourly earnings. Additionally, taxi firms and regulatory agencies compile large sets of data on taxi trips allowing researchers access to dynamic daily labor supply decisions for thousands of drivers. For these reasons, taxi drivers have been widely studied by economists, with most interest

[1] The bracket need not specifically be a day though a daily bracket is what I will discuss in this paper.

focused on driver labor supply responses to wage changes across days (Camerer et al., 1997; Chou, 2002) and explaining a robust negative wage elasticity of hours worked. This result has been controversially explained by reference dependence with an extensive literature devoted to confirming or rejecting this hypothesis (Farber, 2005; Fehr and Goette, 2007; Farber, 2008; Crawford and Meng, 2011; Farber, 2015).

Thus far, the literature has focused exclusively on identifying evidence of loss aversion, a particular kind of reference dependent preferences under which risk aversion exists only around the reference point and utility is linear aside from a kink at the reference point. I extend on the debate over taxi driver reference dependence by testing for risk attitudes which change relative to the reference point, rather than existing only in a small neighborhood around the reference point. This manner of reference dependence is modeled by Prospect theory and results in the well-known S-shaped reference dependence.[2] Under such preferences, driver labor supply would be sensitive to changes in the distribution of earnings throughout the day with the response being dependent on both the level of earnings and distance from the reference point as opposed to only the latter.

To test for this manner of reference dependent preferences, I follow the approach of Farber (2005) which utilizes the relationship between cumulative earned income within a shift and hazard of stopping to make inferences about the shape of a driver's utility for wealth within a day. This model is identified off of variation in a driver's earnings throughout the day with a positive hazard rate in daily income being indicative of risk aversion. I test for and find evidence of a very particular kind of reference dependence in which an individual is risk loving below a reference point and risk averse beyond the reference point. This form of reference dependence is uniquely identified

[2]While I refer to this form of reference dependence as Prospect theory throughout the paper, I will neither address nor incorporate probability weighting in my discussion. For the purposes of this work, Prospect theory refers only to the S-shape of preferences around a reference point.

with this methodology by a non-monotonic relationship between hazard of stopping and cumulative income. This non-monotonic hazard rate within a day also yields the clearest evidence to date of narrow bracketing in labor supply decisions.

My findings highlight the connection between risk preferences and labor supply with wage uncertainty beyond loss aversion, the focus of previous research in this area. Additionally, my findings yield evidence that workers evaluate their earnings within narrow time frames implying that workers react to short-term volatility in their earnings.

The paper is organized as follows: Section 1 provides a summary of competing models of daily labor supply and an overview of the literature on reference dependence in labor supply. Section 2 models a driver's decision to stop working which yields testable implications for various forms of reference dependence. Section 3 is a description of the data I use in my analysis and compares and contrasts the San Francisco and New York taxi markets. Section 4 presents my empirical analysis of a driver's stopping decision as an empirical test of reference dependence in driver labor supply. Lastly, Section 5 provides a discussion of my results in the context of previous work on reference dependence and flexible labor supply.

1.2 Context & Background

1.2.1 Reference Dependent Preferences

Reference dependent preferences refer to an agent evaluating the value of an object relative to some reference point. Such preferences were first discussed in Friedman and Savage (1948) as a means to explain individuals who would simultaneously demonstrate risk averse behavior by purchasing insurance and risk loving behavior by buying lottery tickets.

Reference dependent preferences can be modeled as a variety of discontinuities in a utility function. A given discontinuity is called a reference point. The discontinuity may be in the utility function itself, producing a jump in utility at a certain point. Alternatively, the discontinuity may be in the first derivative which yields a phenomena known as loss aversion. Lastly, the discontinuity may be in the second derivative which yields the widely known, S-shaped value function. Figure 1.1 provides a visualization of various reference dependent utility functions compared to a standard (linear) utility function.

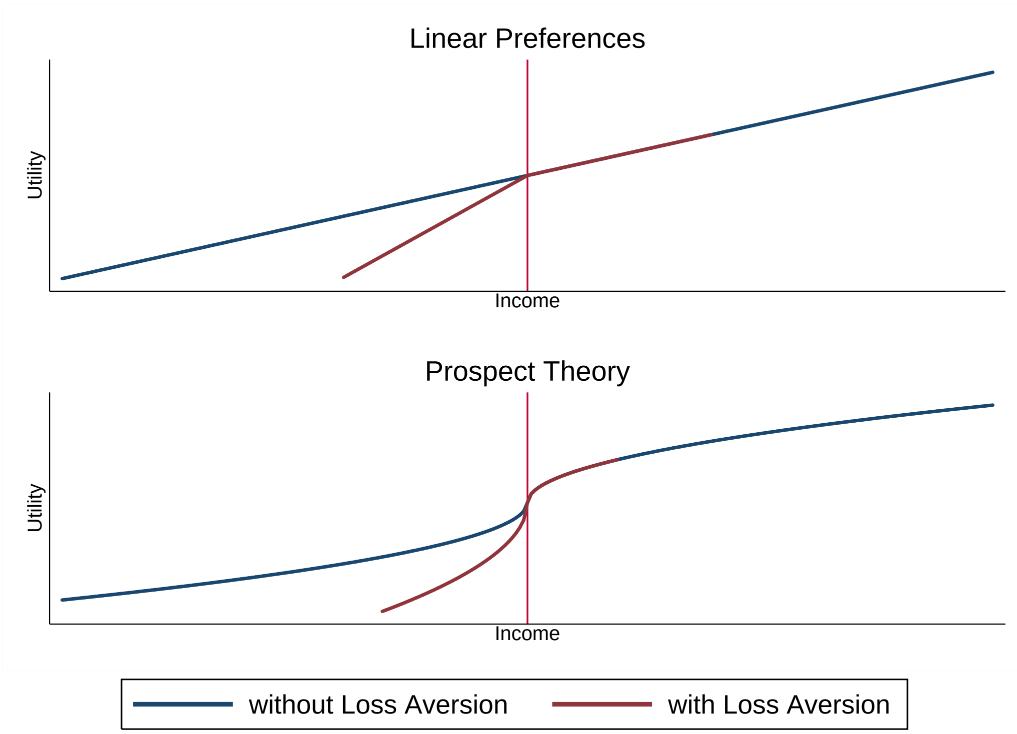

Figure 1.1. Reference Dependence examples with and without loss aversion

Note: Here the reference point is denoted by the red vertical line. The inclusion of loss aversion increase the marginal utility of income in the range of income below the reference point, referred to as the "loss region".

Reference dependence has been used to explain laboratory observations in scenarios with choices under uncertainty. Such observations sparked the development of Prospect Theory (Kahneman and Tversky, 1979). In a prospect theoretic model, agents are risk loving over dollar amounts below their reference point (the loss region) and risk averse over dollar amounts above their reference point (the gains region). If an agent's base wealth in a decision is below their reference point this implies a shift in an agent's behavior around and after reaching a reference point. The key mathematical feature of this model is that for a reference point r, $u''(x) \geq 0$ for $x < r$ and $u''(x) \leq 0$ for $x > r$. This is the final form of reference dependence discussed above which produces an S-shaped value function.

A popular version of reference dependence in behavioral economics is loss aversion. Loss aversion, as outlined in Tversky and Kahneman (1991), is the assumption that an agent has a kink at their reference point caused by a discontinuity in the first derivative. This manner of reference dependence results in a discontinuous decrease in marginal utility at the reference point. This can be modeled with or without an additional discontinuity in the second derivative around the reference point.

One critical drawback of reference dependence theories with respect to their application is that most offer no insights as to where the reference point is. Even in laboratory experiments, it is very difficult to detect where a reference point may actually lie which severely limits the tractability of reference dependence for predictions of behavior in the field. A major advancement in postulating where a reference point may plausibly be came from Koszegi and Rabin (2006) whose model of reference dependence endogenizes the reference point as an agent's rational expectation. Other more well-defined objects such as purchasing prices (Shefrin and Statman, 1985) or whole numbers (Allen et al., 2014) have also been suggested as reference points in some settings.

A key element of reference dependence in labor supply is narrow choice bracketing. Choice bracketing, as outlined by Tversky and Kahneman (1981) and Read et al. (1999), is the phenomenon of decision-makers evaluating their choices within a narrow scope rather than holistically. In the context of labor supply, this means that rather than making labor supply decisions accounting for work on future days, weeks, months, etc., a decision-maker evaluates his decision of how much labor to supply only within a narrow time frame, or bracket. For instance, a driver may have a target to make $1000 in a week. Income earned in the preceding or subsequent weeks is not integrated to the driver's utility within the week. In this example, the one week period in which the driver evaluates his wealth relative to his $1000 reference point is the bracket in which the driver is making decisions. Subsequent weeks would be their own bracket and wealth in those weeks would be evaluated independently from the week considered in this example.

Evaluating decisions in this manner is a necessary though not sufficient condition for reference dependence within a day's labor supply. However, even in the absence of reference dependence, narrow bracketing has important implications for labor supply. For example, decision-makers who bracket narrowly might violate smoothing behavior predicted by lifecycle models of labor supply (Browning and Crossley, 2001).

1.2.2 Competing Models of Labor Supply

Neo-Classical Model

In a standard, Neo-Classical, model of labor supply, workers do not have risk preferences over short-term outcomes such as daily income. Instead, these dynamic life-cycle models of labor supply predict that workers respond to wage variation by intertemporally substituting between days. When a worker observes lower wages on a particular day, he responds by working fewer hours on that day. On the other hand, when a worker observes higher wages, he responds by working more hours.

This model of labor supply predicts a positive labor supply response to higher average wages during a working day. Additionally, a model based on intertemporal substitution suggests that dynamic labor supply decisions throughout the day should primarily be based on observed wages and cumulative hours worked, not cumulative daily income (Farber, 2005). I discuss this last point in more detail in Section 2 when I describe a model of a driver's dynamic labor supply decision within a day.

Loss Aversion

A loss aversion model, applied to short-term labor supply, implies that drivers have income targets. These income targets are modeled as a driver's reference point. The discontinuous jump in marginal utility at the income target generates localized risk aversion around the reference income level. Combining loss aversion preferences with narrow bracketing results in workers having daily income targets which reset between days.

The discontinuity in marginal utility from a loss aversion model generates starkly different predictions for daily labor supply than the Neo-Classical model. If average wages are high, then a driver will reach his reference income faster and thus advance to his gains region in which his marginal utility of income is lower than it was in the loss region. The lower marginal utility of income results in a driver being more likely to stop working.

In contrast to the Neo-Classical model, a loss aversion model predicts a negative labor supply response to higher average wages during a working day. Furthermore, the level of cumulative earnings within a day plays a more significant role in the labor supply decision. Local risk aversion generated by the reference point in this model implies that a driver should be more likely to stop as he approaches his reference point. While this is not deterministic, drivers should be more likely to quit at or beyond their reference point relative to earlier in the shift regardless of hours worked or other factors.

Prospect Theory

Prospect Theoretic preferences, while also a form of reference dependence, yield different predictions and implications for daily labor supply. Under Prospect theoretic preferences, marginal utility at the beginning of the day is relatively low but is increasing as cumulative earnings build. This implies that early earnings are disproportionately important for shift survival. Unlike a model of loss aversion, increasing marginal utility up to the reference point results in a driver being less likely to stop the closer he is to the reference point. In the gains region, beyond the reference point, diminishing marginal utility results in a driver being more likely to stop working as his income for the day increases.

This particular model of reference dependence, unlike the Neo-Classical and loss aversion models, does not yield clear predictions for wage responses. Increasing sensitivity in the loss region implies a positive response to higher observed wages for incomes within that region but diminishing sensitivity beyond the reference point implies the opposite wage response once the driver has sufficient earnings for the day. With these preferences, average elasticities could be positive or negative but would be expected to be small or 0.

The intuition for Prospect theoretic preferences is somewhat different than for loss aversion. While the reference point in both cases can be interpreted as an income target, the responses to that target are very different in the two models. In the loss aversion model, a driver's disutility from being a given distance below his income target is greater than his utility from being that distance above his income target. In the Prospect theoretic model, each additional dollar that brings a driver closer to his income target makes him relatively happier than the last.[3]

1.2.3 Evidence

Camerer, Babcock, Loewenstein, and Thaler (1997) were the first to introduce reference dependence to the labor supply literature in their study of wage elasticities of

[3] Allen et al. (2014) provides a useful analogy one may consider for the intuition behind this phenomenon. In their context, runners have reference points in finishing times. As a runner approaches the finish line, he exerts more effort if he is close to his finishing time goal in order to complete his race goal. The runner does not stop if his time goal has passed but exerts less effort to finish.

hours worked by New York City cab drivers. In their seminal paper, they found that the average wage elasticity for a small sample of drivers was negative, consistent with the loss aversion model described above.

Despite early replications (Chou, 2002; Farber, 2005), the negative elasticity result generated skepticism for econometric reasons. The econometric model used to estimate these elasticities was subject to a well-documented downward bias in the estimated results known as division bias. Discussed in detail by Borjas (1980), division bias causes the coefficient interpreted as an elasticity to be biased downwards in the presence of measurement error or model misspecification. To correct for this, Camerer et al. (1997) instrumented for the wage of a given driver with the median wages of all other drivers working at the same time, arguing that such an instrument would be uncorrelated with biasing measurement error. Even with the IV approach to correct for division bias, estimates of the wage elasticity were negative.

Given the econometric issues with estimating an elasticity, Farber (2005) presented an alternative approach to testing for reference dependence, changes in hazard rates. Under a loss aversion model, there should be a spike in the hazard of stopping rate close to the reference point. However, in Farber's analysis, cumulative shift income was uncorrelated with hazard of stopping with the inclusion of driver fixed effects, a result consistent with Neo-Classical labor supply. Farber (2008) extended on this work by attempting to directly estimate driver reference points but found that point estimates of the reference point were either unstable from day to day or implausibly large. Moreover, stopping behavior appeared too smooth to be indicative of loss aversion.

Crawford and Meng (2011) extended this line of research by endogenizing the reference point as the driver's expectation of earnings, given their past earnings, on a particular day of the week. Crawford and Meng (2011) posited that taxi drivers are loss averse in both hours and earnings. At each reference point, hazard of stopping would thus spike. Under these conditions, a driver would be most likely to quit after reaching his second reference point. In this model, earnings should be correlated with stopping conditional on reaching the reference point in hours first. To test for this, Crawford and Meng (2011) split the original Farber trip sample into two groups, one in which the first hour of a shift was below the average earnings for that driver and the other where the first hour's earnings were above average. They found that in the former sample, income earned was positively correlated with stopping. However, in the latter sample, hours, not income was positively correlated with stopping even accounting for driver fixed effects.

This form of endogenizing the reference point based on a driver's observed average earnings or hours is an implementation of the expectations-based reference dependent model outlined in Koszegi and Rabin (2006) (henceforth referred to as KR). The implementation of the KR model has several advantages. First, it allows for weighting between standard consumption utility (as in the Neo-Classical model) and gain-loss utility (as with reference dependence). Additionally, it provides insight into what a reference point is, allowing the reference point for a driver to be estimated separately from other variables.[4] In a structural estimation, assuming that a driver has constant marginal utility of income[5] within a day, Crawford and Meng (2011) estimate a plausible and robust weighting on KR gain-loss utility; the conclusion of the Crawford and Meng analysis was that drivers behave consistently with being loss averse with reference points in both income and hours worked.

Most recently in the literature on taxi driver labor supply, Farber (2015) examines a large subset of a dataset similar to the one I use. With a sample of all trips taken by approximately 15% of all drivers[6] in New York City between 2009 and 2013, Farber

---

[4]Crawford and Meng (2011) use several variations of calculating a KR reference point ranging from an econometrically estimated expectation of earnings and hours to taking simple averages.

[5]In the context of a decision under uncertainty, this is equivalent to saying the driver is risk neutral with a von Neumann-Morgenstern utility function.

[6]Farber's sample of drivers accounts for 8,802 drivers and 116,177,329 trips.

estimates a positive average wage elasticity of hours worked, with the majority of drivers in his sample having a positive individually estimated wage elasticity. Using this large dataset of taxi trips, Farber also estimates a hazard of stopping model similar to the one estimated by Farber (2005). The stopping model estimated yields an increasing probability of stopping in hours worked but no clear monotonic increase in the probability of stopping with respect to cumulative income earned in the shift.

With the majority of drivers exhibiting a positive wage elasticity, and without a clear, monotonically increasing probability of stopping in income, the conclusion of Farber (2015) is in favor of at most a limited role for loss aversion in driver labor supply. Building on the models of these previous papers, I explore the potential role for reference dependent preferences beyond loss aversion in a driver's daily labor supply.

1.3 Model

I now outline a model of the driver's decision to pick up another fare or not. The model is heavily based on the one originally presented in Farber (2005). I make no assumption about the presence of reference dependence and place only minimal structure on the functional form of the utility function. A driver's utility is a function of two objects, hours worked and income earned in that day. I separate each of these variables into two components: the quantity that the driver has already realized and the additional quantity the driver expects to realize in the future. The former is treated by the driver as a parameter and is taken as given in the driver's decision rule. The utility function I write as:

$$V(I,H) = u_I(I) + u_H(H)$$

Where H and I are the hours worked and income earned at the decision point respectively. u_I and u_H are separate functions which represent a driver's preferences over income and hours respectively. I make the following assumptions on the structure of the utility function which include:

Assumption 1: V is twice differentiable in all of its arguments almost everywhere.

Assumption 2: $\frac{\partial V}{\partial H} < 0$ a.e.

Assumption 3: $\frac{\partial V}{\partial I} > 0$ a.e.

Assumption 1 is self-explanatory. Assumptions 2 and 3 simply state that the driver's utility is decreasing in hours worked and increasing in income and that utility is smooth in these variables. Note that I am also assuming that utility from hours worked and income are additively separable, an assumption made by Crawford and Meng (2011) which allows me to evaluate the derivative of utility for one object without concern for the level of the other.

Following the precedent set by Farber (2005) and followed by Crawford and Meng (2011), a driver will quit working if the utility from stopping at his current level of income and hours worked, (I_t, H_t) is greater than his expected utility of working the next trip $(E[I_{t+1}], E[H_{t+1}])$ where $E[I_{t+1}] = I_t + E[i_{t+1}]$ and $E[H_{t+1}] = H_t + E[h_{t+1}]$, with i_{t+1} and h_{t+1} being the additional income and hours worked from trip $t+1$. The stopping rule can then be written as follows:

$$V(I_t, H_t) - EV(I_{t+1}, H_{t+1}) > \varepsilon_t \tag{1.1}$$

ε_t allows for unobservable stochastic elements to influence the decision to stop work after trip t. I assume only that $E[\varepsilon_t | H_t, I_t] = 0$. This model does not account for the driver's expectations beyond trip $t+1$ though all results I obtain are without loss of generality.[7]

In the data, previous researchers and I observe the hours worked and earnings at the end of a trip, in this model, I_t and H_t respectively. What is actually estimated empirically is a probability of stopping conditional on these variables. The driver's decision rule allows me to write the probability of stopping as:

$$Pr(\text{last trip of shift}) = Pr(\varepsilon < V(I_t, H_t) - EV(I_{t+1}, H_{t+1}))$$

$$= F(V(I_t, H_t) - EV(I_{t+1}, H_{t+1})) \tag{1.2}$$

Where F here is the cdf for the random variable ε_t. I would like to interpret the meaning of the observable estimations (probability of stopping) for the utility function. In particular, I and other researchers which have done such estimations observe a change in the probability of stopping as the level of income a driver has at the end of his last trip increases. Thus, I am interested in how equation 1.2 changes when I_t, the initial income at which the driver is making his decision to stop or continue, changes.

With this in mind, I differentiate equation (1) with respect to I_t:

$$\frac{\partial Pr(\text{last trip of shift})}{\partial I_t} = \frac{\partial F(V(I_t, H_t) - EV(I_{t+1}, H_{t+1}))}{\partial I_t}$$

Since F is a cdf it is monotonically increasing in its argument and so this derivative will be positive if and only if:

$$\frac{\partial V(I_t, H_t)}{\partial I_t} > \frac{\partial EV(I_{t+1}, H_{t+1})}{\partial I_t} \tag{1.3}$$

[7] Accounting for the expectations of trips beyond $t+1$ requires a simple application of an Envelope Theorem for arbitrary choice sets (Milgrom and Segal, 2002), treating earned income and time worked as a parameter and holding expectations only over future income and hours to be worked.

Note that $i_{t+1} > 0$ (a driver cannot have less income from taking additional trips) and so here I argue that the utility level of the right-hand side of the inequality is at a higher income level than the left-hand side.

The intuition for the exercise here is simple: consider a driver who has just dropped off a fare and is contemplating working more trips or stopping and going home. The driver has well-defined expectations for how much money he will have and how many hours he will have to work if he decides to take trip $t+1$. From this information, he makes the decision to continue or stop working. Now, suppose that same driver contemplating the same decision has earned one additional dollar from the same shift. Is the fortuitous driver, who has slightly more money in his pocket than the other driver but has worked the same number of hours, more or less likely to stop working? The answer depends on the shape of his utility function.

First, consider a driver who is completely risk neutral. In this case, $\frac{\partial V}{\partial I} = c$ for some positive constant c. Then, equation 1.3 is always an equality and thus probability of stopping is constant in income. This would be consistent with drivers not evaluating income within narrow brackets as in the Neo-Classical model.

Next, consider a driver who is either loss averse or has diminishing marginal utility of income. In the case of diminishing marginal utility of income, it is globally true that the marginal utility at any given higher level of income is lower and thus hazard of stopping should be monotonically increasing over all income ranges. With loss aversion, on the other hand, utility is linear in income and thus there is constant marginal utility of income except when the income level is close to the reference point. Close enough to the reference point, marginal utility at a higher income level is lower resulting in an increase in the hazard rate over income levels below but close to the reference point. For income levels sufficiently far from the reference point, hazard of stopping is constant in income, thus there would be a spike in the hazard rate near the reference point.

Lastly, I will discuss a case which has been heretofore unexamined in the labor supply literature, a reference point with respect to the second derivative within a day. This is similar to the models proposed by Friedman and Savage (1948) and, more popularly, Kahneman and Tversky (1979). In this case, marginal utility of income is increasing up until the reference point and decreasing after the reference point. It is straightforward to see that if $\frac{\partial^2 V}{\partial^2 I} > 0$ then the probability of stopping would be *decreasing* in income until the reference point, after which the probability of stopping would be increasing in income.

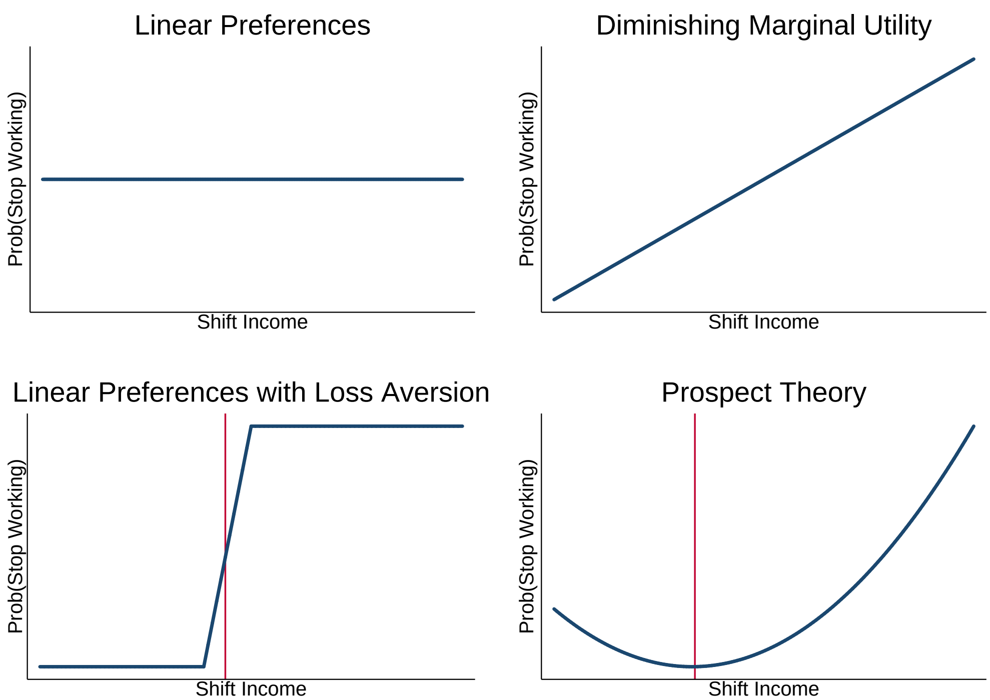

Figure 1.2. Example potential relationships between hazard of stopping and shift income under each model

Note: For applicable models, the reference point is represented by a red horizontal line.

With these models in mind, I now have a way to identify reference dependent preferences from estimates of the probability of ending a shift. If the probability of stopping here is uncorrelated with income levels within the day, then I will reject the hypothesis that drivers evaluate income within narrow brackets. If hazard rates are either increasing in income globally or spike at an income level then I would have evidence of narrow bracketing and either risk preferences or loss aversion in daily labor supply. Finally, I could observe a decreasing probability of stopping over a certain income range followed by an increasing probability of stopping. If this were the case, this would be the strongest evidence to date that taxi drivers are reference dependent and evaluate income within narrow brackets. This would also provide a rough estimate of what the reference point is without an assumption about its existence.

1.4 Data

1.4.1 Description of Data

The dataset of San Francisco taxi drivers was obtained from the origination company with assistance and facilitation of an SFMTA (San Francisco Municipal Transportation Agency) Taxi division employee. Each observation is a trip made by a cab driver operating with a mid-sized fleet. The raw data include 1,634,982 trips from a mid-sized fleet spanning August, September, and October of the years 2010 through 2013 as well as July 2013. Each trip includes the fare, type of payment, distance traveled, number of passengers, start location, end location[8], driver identifiers, cab number, and start/stop times for the trip.

Taxi drivers with this particular fleet operate under one of two contract structures. The most common contract is for a driver to be an employee of the company with his own license to drive a taxi. These drivers lease their cabs for 10 hour periods during which, a driver keeps all of his revenue but must pay a “gate fee” for the vehicle. The

[8]After obtaining the data, SFMTA policy was revised to no longer allow the distribution of GPS location data used in these analysis.

driver must also pay for the gas used in the shift. If a driver returns a cab after the 10 hour period, he must pay a fine for a late return. The alternative contract for a driver is to be a medallion holder. Medallion holders own their own taxi medallions but are attached to a fleet. These drivers have their own choice of shift and vehicle and do not pay a gate fee. The MTA was kind enough to provide me with information on which drivers in the data are medallion holders. Individual driver information includes year of birth, medallion status, and the year the driver was licensed to operate a taxi was provided by the SFMTA for matching driver identifiers.

Trip data from New York City was collected by the New York City Taxi and Limousine Commission (TLC) for 173,179,759 rides initiated in New York City between January 1, 2013 and December 31, 2013. These data are published online by the TLC. For each trip, the taxi driver's hack license, the cab's medallion, and the cab company's vendor ID were recorded, along with information about the ride's length, cost, and GPS location. These data are a portion of the data used by Farber (2015), whose data spans 2009 to 2013. Unlike Farber, I will be conducting my analysis on the entirety of the 2013 New York data rather than a random sample of drivers.

While the New York dataset is much larger than the San Francisco data, it is not as rich. For instance, I do not have any information about the individual drivers in the New York data nor do I have information about the fees drivers must pay for their vehicle. As New York data have already been researched extensively, I will analyze the New York data for the purpose of comparison with the San Francisco data and completeness.

Tip amount is also included in both datasets but only if the tip was paid via credit card. As such, I do not use this amount in calculating driver revenue. Additionally, trips are not broken down by shift and so for the purposes of analysis, I infer the shift by a six[9] hour or more period between the end of a trip for a driver and the beginning of the

[9]Earlier versions of this paper separated shifts by a four hour or more period between the end of a trip

driver's next trip. Appendix 2 details the data cleaning procedure I use for both datasets.

1.4.2 Comparing New York and San Francisco

The first question to arise when comparing behavior across two cities is how the market environments differ between the two cities. I will begin to answer this question by looking at differences between the two cities in fundamental observables such as average trip duration, distance, and earnings. The summary statistics for trips in the cleaned data for each city can be found in Table 1.1.

Table 1.1. Trip-Level Summary Statistics

	San Francisco	New York City
Trip duration (minutes)	10.99	12.97
	(8.30)	(19.70)
Trip earnings	15.92	12.88
	(14.68)	(10.06)
Passengers	1.09	1.71
	(0.71)	(1.37)
Time between trips without passenger	15.69	11.64
	(24.63)	(20.96)
Observations	1,550,080	153,716,233

Note: Standard deviation in parentheses. Trip earnings are in real 2014 dollars.

Despite being very different cities, the trip-level data for San Francisco and New York City are remarkably similar. Trips in San Francisco are, on average, slightly shorter than trips in New York City though average trip earnings are higher in San Francisco. This can be explained by taxi rates being slightly higher in San Francisco where the standard rate set by the SFMTA is a \$3.50 initial charge and \$0.55 per $\frac{1}{5}$ mile. In contrast, New York City's standard rate set by the TLC which starts at \$2.50 at a rate of \$0.50 per $\frac{1}{5}$ mile.

for a driver and the beginning of the next trip. The change to a six-hour gap was made to make results more directly comparable to Farber (2015) but changes nothing materially about the results.

It is difficult to interpret the variance of these statistics across the two cities due to the New York City data having roughly 100 times the observations of the San Francisco data. Another caution that should be taken when looking at these summary statistics is that data for San Francisco are limited to a single fleet whereas New York City data are a census of fleets. With that said, all results for San Francisco are underlaid by the assumption that this is a representative fleet for San Francisco. Discussions with agents of the San Francisco MTA support this assumption.

One thing of note is that there appears to be significantly more variation in trip length in New York City compared to San Francisco. This is partly explained by trips for the San Francisco data covering a wide range over the Bay Area. Using GPS data attached to most trips in the San Francisco dataset, I plot the drop-off and pick-up locations for the San Francisco data. Figure 1.3 shows all drop-off locations for the San Francisco data and Figure 1.4 shows all pick-up locations. While the vast majority of trips are clustered in San Francisco and Oakland, taxi trips (both pickups and drop-offs) do appear to radiate as far south as San Jose.

Despite the abundance of pick-ups in the outskirts of the Bay Area (relative to San Francisco), trip origins in the outskirts of the Bay Area are much more sparse. This highlights the relevance of controlling for drop-off location in my later analysis. A trip that terminates far away from San Francisco itself is likely to result in a longer waiting time between trips either due to the cab looking in an area with lower demand for taxi services or due to time to move to a higher demand location without carrying a fare.

Table 1.2 gives summary statistics for each city at the shift level. The shift summary statistics tell essentially the same story as the trip-level statistics. Both cities appear to be very similar in the composition of shifts: with about the same percentage of AM and PM shifts in each city. The key difference appears to be that San Francisco drivers work less than New York City drivers but bring in slightly more money on average.

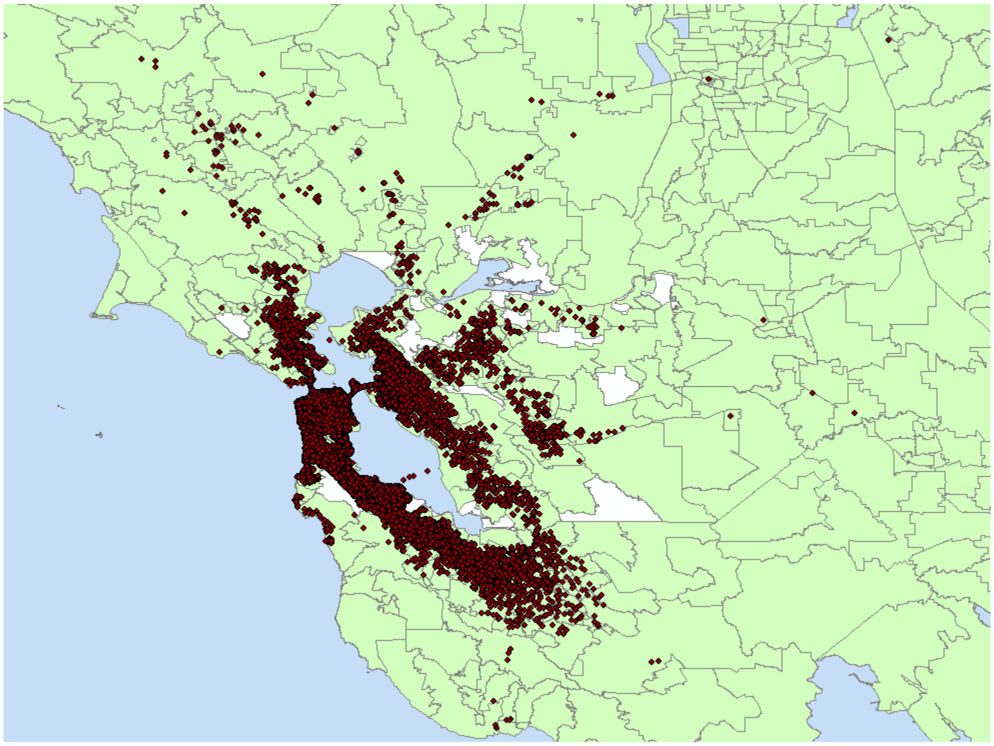

Figure 1.3. San Francisco trip drop-offs

It is unclear whether or not taxi drivers in San Francisco spend more time looking for trips or if they spend more time on break relative to their New York counterparts.

Perhaps most surprising is that despite differences in productivity measures, average shift earnings for both cities are remarkably similar (though subject to high variance). Nothing about income targeting reference dependence can be directly inferred from this, but it lends credence to the notion that preferences for labor supply across the two cities are comparable.

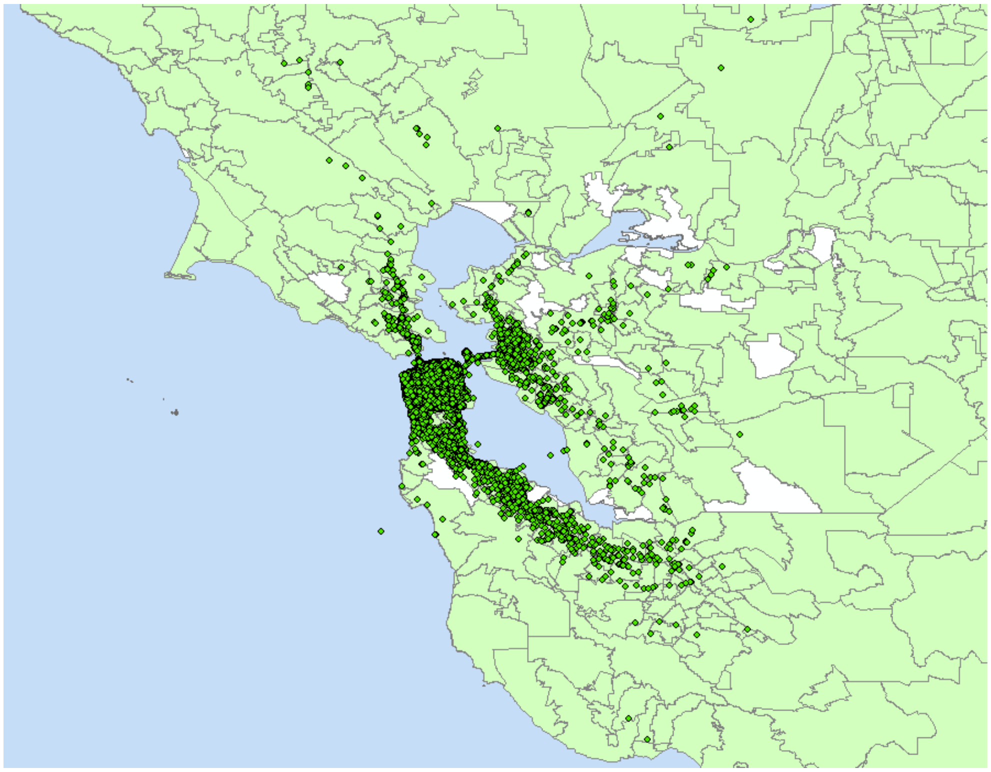

Figure 1.4. San Francisco trip origins

Table 1.2. Shift-Level Summary Statistics

	San Francisco	New York City
PM shifts	44.56%	43.25%
	—	—
Shift length (hours)	8.11	8.74
	(2.62)	(2.64)
Percent of shift spent with passenger	42.10%	52.51%
	(13.49%)	(12.14%)
Shift earnings	290.46	277.81
	(117.56)	(93.79)
Number of trips	18.25	21.58
	(9.46)	(8.57)
Observations	84,943	7,123,924

Note: Standard deviation in parentheses. A PM shift is defined based on the definition of a PM shift used by the SFMTA and includes all shifts beginning after 4pm and before 4am. Shift earnings are in real 2014 dollars.

1.5 Analysis

My analysis of reference dependence in taxi drivers is divided into three parts. First, I estimate changes in the average hazard of stopping as cumulative shift income increases. Next, I will repeat the aforementioned estimation for each individual driver in the data and examine the potential and extent of heterogeneity in driver preferences over labor supply. Lastly, I estimate a labor supply function implied by using elasticities calculated from wages averaged over all active drivers in a given hour.

1.5.1 Hazard of Stopping

As outlined in Section 2, a reference point in sensitivity, like that of a prospect theoretic value function, can be identified from a switching point in the effect of cumulative income on hazard of stopping. To identify such a switching point, from decreasing probability of stopping to increasing in income, the model I estimate will have to allow for such a switch. The hazard of stopping models estimated in Farber (2005) and Crawford and Meng (2011) included only a linear term and so would not capture such an effect. Farber (2015)'s hazard of stopping estimation compared average hazard of stopping between income bins. However, it is possible that using income bins and estimating an average difference between bins could mask the effect I am searching for, a possibility I will address later. As such, my primary identification strategy for the non-linear effect will involve the use of spline regression to allow for hazard of stopping to take a mostly flexible non-linear shape.[10] Equation 1.4 will be the baseline estimating equation.

$$Pr\left(\text{trip } t = \text{last trip of shift}\right) = f(I_t) + g(H_t) + \gamma X_t + \varepsilon_t \tag{1.4}$$

In Equation 1.4, $f(I_t)$ denotes a continuous function to be estimated by linear

[10]The only constraint on the shape of the estimated relationship is that it is continuous and linear between knots.

splines which represents the relationship between cumulative shift income and a driver's hazard of stopping. I place the knots used for estimating the linear pieces of this function in \$50 increments. $g(H_t)$ denotes a continuous function to be estimated by linear splines which represents the relationship between shift hours and a driver's hazard of stopping. I estimate this model with a linear and non-linear $g()$. In the non-linear estimation of the shape of $g()$, I set knots in one-hour increments. Lastly, X_t is a vector of additional controls for trip t and the associated vector of coefficients for the controls is γ. These additional controls include controls for the hour of the day, day of the week, month, year, driver fixed effects, and destination fixed effects by ZIP code.

Figure 1.5 depicts the shape of the relationship between shift income and stopping hazard estimated by the linear splines with hours included in the estimating equation linearly. For both San Francisco and New York City, the relationship is clearly non-monotonic. Indeed, this relationship is consistent with what we would expect if the representative driver had prospect theoretic preferences with a reference point of approximately \$200 to \$250. A reference point in this range does not appear to be unreasonable given the costs to drivers of taking out the cab for a shift and typical earnings.

Of some note is that for relatively large shift incomes (greater than \$450) there appears to be a dip in the hazard of stopping followed by a spike. These shifts represent only a small fraction of the data and likely represent shifts in which the driver had a large positive shock. In such shifts is when income affects hazard, in an absolute sense, the most but is not particularly helpful in understanding driver behavior on typical days. The spline, however, indicates that hazard of stopping does being to increase in ranges that would be attained by drivers on typical days.

While hours are controlled for linearly in the specifications depicted by Figure 1.5, it is not necessarily clear that the relationship between hours and hazard of stopping is linear. Indeed, Crawford and Meng (2011) postulates that drivers could have hours

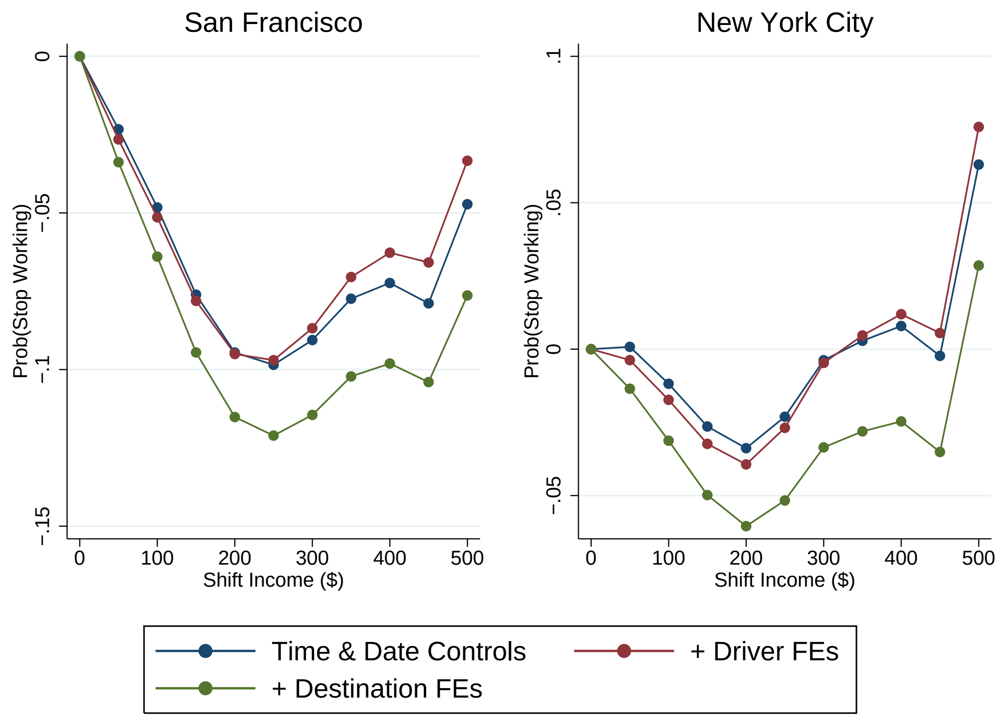

Figure 1.5. Linear spline estimation of hazard of stopping in income controlling for hours linearly

Note: Probability is depicted relative to a shift income of $0. Estimation results may be found in Tables A.1 and A.2.

targets which would imply that the hazard of stopping may not be increasing linearly in hours but may spike. To account for this possibility, I estimate the model described by Equation 1.4 with splines for both income and hours.

Figure 1.6 depicts the hazard rate in with respect to shift income with splined hours. Including hours with splines instead of linearly in the estimating equation changes the shape of the relationship between hours and income. For both cities, the hazard rate is still decreasing for the first $100 earned in a shift. However, the rate of change in the hazard rate flattens much earlier and remains flat until the extreme earnings ranges ($450+).

This shape indicates that beyond a certain range of earnings, income no longer

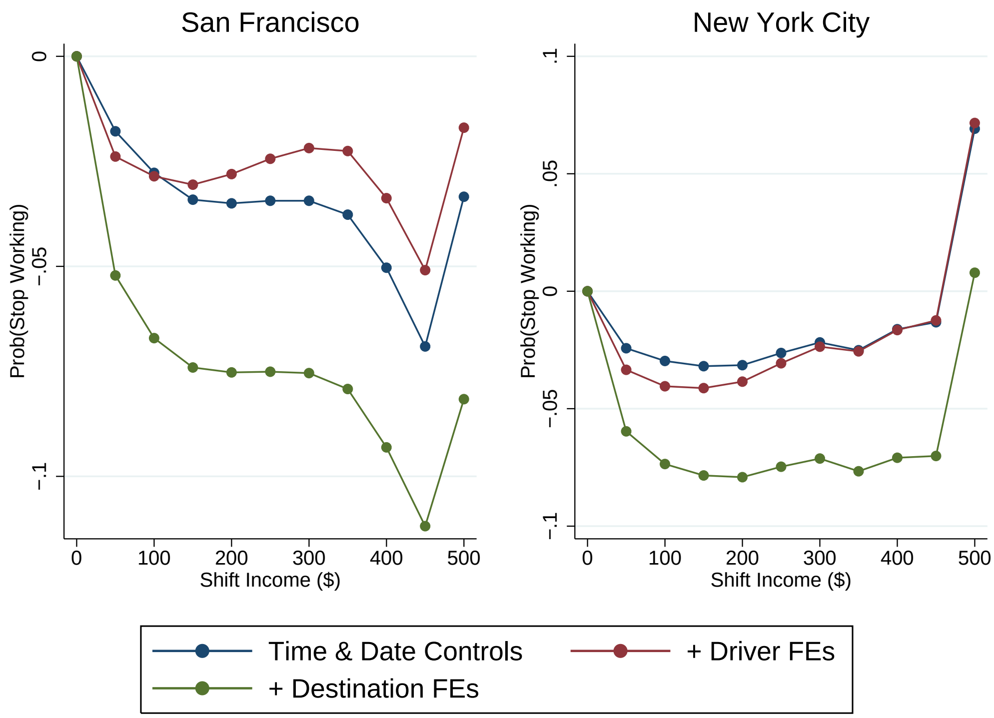

Figure 1.6. Linear spline estimation of hazard of stopping in income controlling for hours with linear splines

Probability is depicted relative to a shift income of \$0. Estimation results may be found in Tables A.3 and A.4.

plays a significant role in a driver's hazard of stopping. Nevertheless, the result that a driver is more likely to stop at lower income levels holds indicating that something that could be interpreted as a "loss region" in which the driver has increasing marginal utility of wealth or is risk loving exists for a representative driver. That this area appears to be consistent with the typical gate fees in San Francisco strengthens the argument that this range of income is of particular importance since ending a shift in this range of income would certainly result in a driver incurring a loss for the day.

Crawford and Meng (2011) postulates that drivers may have reference points in hours worked as well as income. Figure 1.7 explores this possibility by examining the estimated shape of $g()$, the relationship between hours worked and hazard of stopping

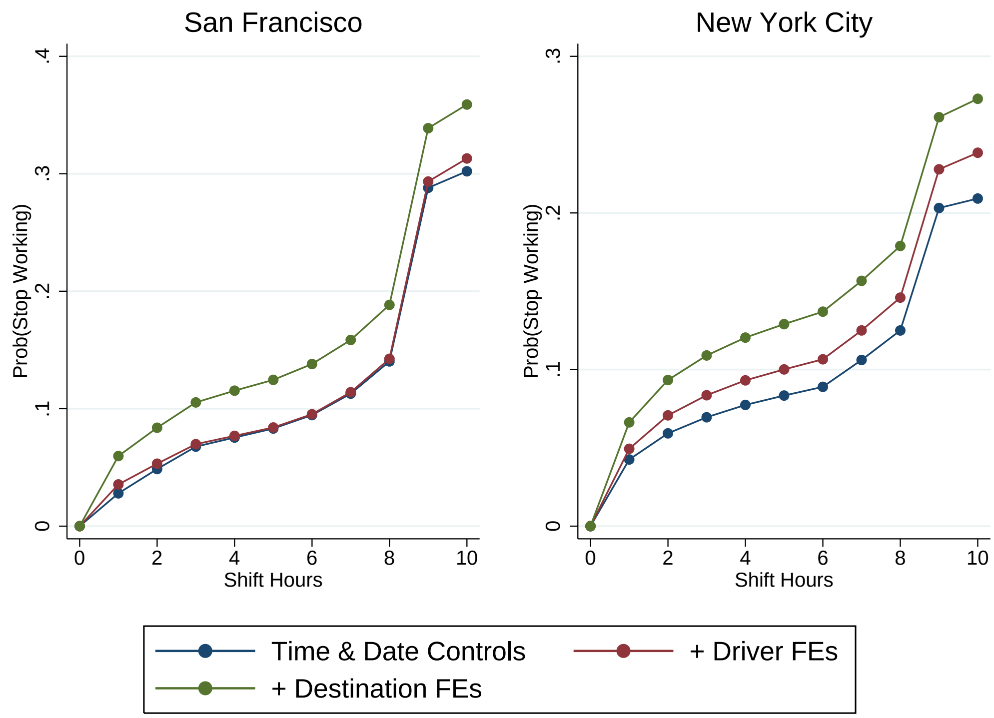

Figure 1.7. Linear spline estimation of hazard of stopping in hours worked

Note: Probability is depicted relative to 0 hours worked. Estimation results may be found in Tables A.3 and A.4.

from the same estimations depicted by Figure 1.6. Hazard of stopping is strictly monotonically increasing in hours. However, at the 8-hour mark, there is a notable jump in the hazard rate. This jump is consistent for both cities. Given that drivers are allowed their vehicles for 10 hours in San Francisco and 12 hours in New York City, it is unlikely that this spike in the hazard rate at 8 hours is due to institutional constraints.

This paper is not the first paper to have the potential to identify non-linear effects in drivers' hazard rate. Indeed, as mentioned before, Farber (2015) uses income and hours bins and finds no non-linear effects on the scale which I have presented here. To explore this further, I replicate the Farber (2015) analysis using the binned income and hours model as opposed to splines. I also estimate a binned income and hours model

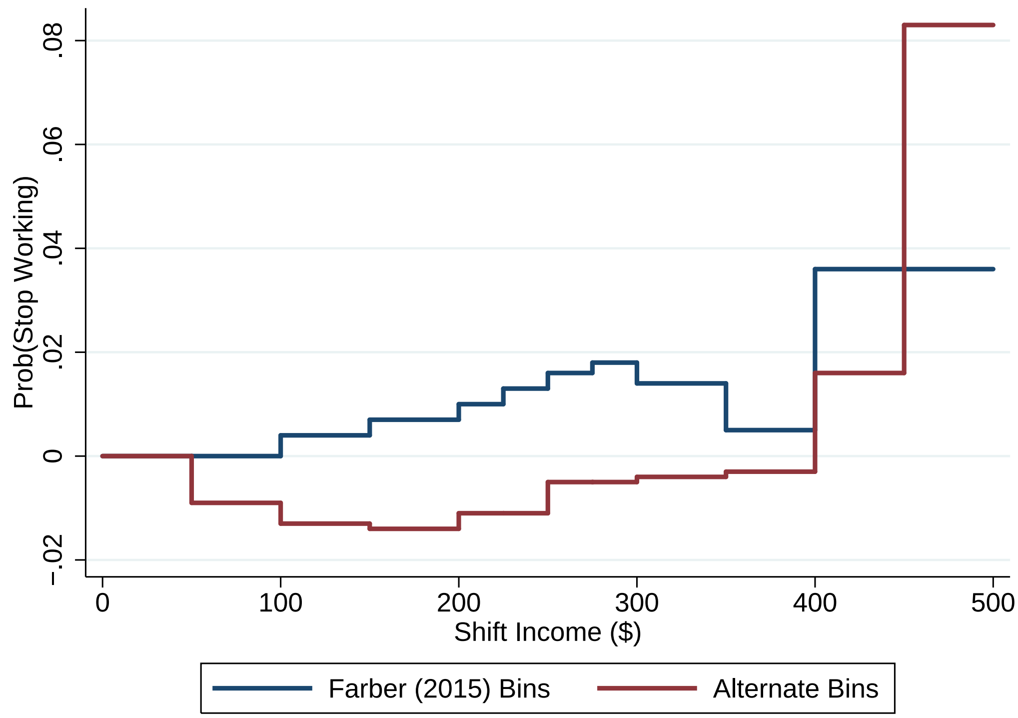

Figure 1.8. Comparison of hazard estimates for income using various bins with New York City data

Note: Bins for income and hours from Farber (2015). Alternate bins spaced evenly every $50 for income. Probabilities are shown relative to the reference bin for a given selection of bins. Estimation results for San Francisco may be found in Tables A.4 and A.6. Estimation results for New York City may be found in Tables A.5 and A.7 (pictured).

using an alternative set of income and hours bins.

Figure 1.8 shows how research selection of the binned set can yield deceptive results. Using the set of bins as in Farber (2015), I find no substantial non-linear effects and, using a very similar dataset, I replicate Farber's hazard of stopping results. However, when I rearrange income and hours bins to be evenly spaced, the binned model produces results consistent with those from the spline model. This can be explained by the selection of the reference bin. With splines and evenly spaced bins, probabilities are being taken in comparison to drivers with a cumulative shift income between $0 and $50. With Farber's bins, on the other hand, while there are smaller bins (and thus higher resolution) in the $200 to $300 range, hazard rates are being compared to all trips which result in a cumulative shift income from $0 to $100 which is the range in which splines detected substantially diminishing hazard rates.[11]

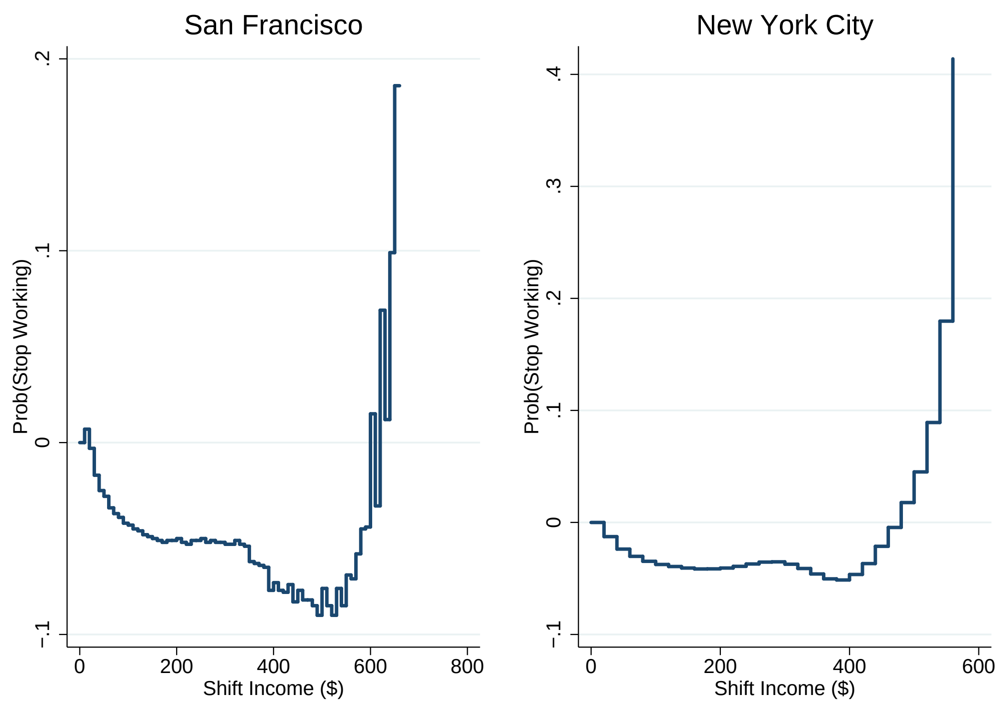

Figure 1.9. Hazard rates using small bins for income controlling for hours using small bins

Note: For San Francisco, income is binned in $10 increments and hours binned in 15-minute increments. For New York City, due to the amount of data and computational constraints, income is binned in $20 increments and hours are binned in 30-minute increments. The estimated model here is identical to that outlined in Equation 1.4 and includes controls for warm temperature, cold temperature, rain, hour of the day, day of the week, month, year, driver, and destination by ZIP code.

In an effort to reconcile modeling differences between this paper and previous papers, I attempt to estimate the effects of income (and hours) on hazard of stopping as non-parametrically as possible. To this end, I leverage the large datasets for both

[11]It should be noted that Farber (2015) was not looking for non-linear effects in the hazard rate, but rather for a spike in the hazard rate in regions in which it would be reasonable to assume a reference point would be.

cities and estimate Equation 1.4 using bins for income and hours which are as small as is computationally feasible. Figure 1.9 shows the results of using extremely small bins are in line with the results depicted by Figure 1.6 using linear splines for both income and hours.

1.5.2 Driver Heterogeneity

The richness of both the San Francisco and New York City datasets allows for analyses beyond population averages. In particular, I estimate Equation 1.5 for each individual driver. Since, for each individual driver, I have substantially less data than the aggregate and in the interest of summarizing the results clearly, I use a second degree polynomial for the $f()$ and $g()$ functions.

$$Pr\,(\text{trip } t = \text{last trip of shift}) = f(I_t) + g(H_t) + \beta_1 \hat{w}_{\tau(t)} + \beta_2 \hat{w}_{\tau(t)+1} + \gamma X_t + \varepsilon_t \quad (1.5)$$

Figure 1.10 provides histograms of individually estimated reference points obtained from the estimated coefficients of the polynomial expansion of income, $f(I_t) = \rho_1 I_t + \rho_2 I_t^2$ in equation 1.5. The reference point for a driver d is estimated by $R_d = \frac{-\rho_1}{2\rho_2}$ provided $\rho_1 < 0$ and $\rho_2 > 0$. For each city, approximately two-thirds of drivers in the sample have an estimated reference point.

Estimated reference points are tightly grouped around the \$200 level. While the majority of drivers do appear to have estimable reference points in line with a Prospect theoretic value function, there does appear to be some level of heterogeneity in driver preferences as one-third of drivers do not appear to have such a reference point. To explore this heterogeneity, I bin drivers based on their marginal hazard of stopping with respect to income in Figure 1.11. Aside from a Prospect theoretic value function, the

Figure 1.10. Individually estimated driver reference points

Note: Reference points only shown for polynomial estimates which produce a U-shape consistent with Prospect Theory. Top 10% of estimated reference points omitted for each city.

next most common response to increasing income within a day is consistent with drivers having diminishing marginal utility of income within a day.[12] As discussed before, this model does not distinguish between diminishing sensitivity and linear utility with a loss aversion parameter. A small fraction of drivers I estimate as behaving consistently with increasing sensitivity (risk loving) within a day or are not identified.

In addition to exploring heterogeneity, I would also like to examine the extent to which individual drivers can capitalize on aggregate expectations of hourly earnings. To this end, I estimate expected earnings w_τ in an hour τ by regressing the hourly earnings for each active driver in an hour on observable characteristics such as weather

[12]This category includes drivers who's hazard of stopping is increasing up until $500 and potentially decreasing afterwards. Less than 5% of all shifts earn more than $500.

Figure 1.11. Implied shape of utility of income within a day

and time. Estimating results for expected hourly earnings are tabulated in Table A.8 for San Francisco and A.9 for New York City. Consistent with a similar analysis performed by Farber (2015), a great deal of variation in the average wage across drivers can be explained by observables.

Figure 1.12 shows the distribution of the coefficients β_1 and β_2 for hazard of stopping on the expected wage in the hour a trip ends and the next hour. When individual drivers' responses to the expected wage are tallied in this manner, there is no clear relationship between the calculated expected wage and hazard of stopping. Indeed, the distributions appear to just be noise centered approximately at 0.

Additionally, there seems to be no meaningful relationship between a driver behaving consistently with reference dependence and their response to expected wage changes. Drivers with an estimable reference point are significantly less likely to respond

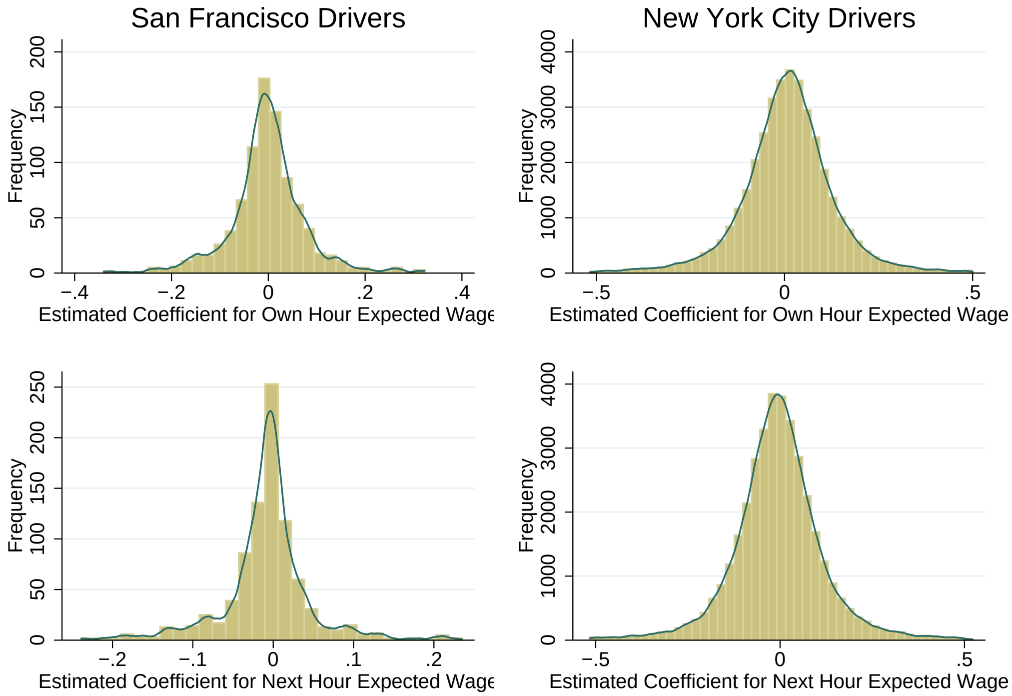

Figure 1.12. Individually estimated driver responses to expected wages

Note: Top and bottom 1% of coefficients for each city are omitted.

to a higher expected wage in the hour a trip ends by being more likely to quit (Mann-Whitney rank sum test, $p < .001$). There is no difference (Mann-Whitney rank sum test, $p = .796$) in drivers' propensity to stop working and their response to the expected wage in the next hour. This lack of a relationship between expected wage changes in an hour and hazard of stopping could potentially be a reflection of an inability to accurately forecast hourly wages. It is entirely possible that a driver on the ground has information or knowledge that is not reflected in even the most careful ex-post econometric models. With this in mind, I examine the relationship between estimated reference dependence and a more traditional measure of driver responses to wages, a shift-level wage elasticity.

As in previous papers, I estimate the wage elasticity from the model described by equation 1.6. Here, β is the main coefficient of interest and is interpreted as the wage elasticity. H_s denotes the hours worked by a driver in shift s and W_s the average wage for that shift as calculated by $\frac{I_s}{H_s}$. W_s is instrumented for by the average shift wage of every other driver who worked that day to compensate for division bias.[13]

$$ln(H_s) = \beta ln(W_s) + \gamma X + \varepsilon_s \tag{1.6}$$

The distribution of estimated driver elasticities for each city is shown in Figure 1.13. As noted by Farber (2015), there is substantial variation in the estimated driver wage elasticity. The median elasticity in San Francisco is small and negative. The median elasticity for New York City, on the other hand, is positive.[14]

A long debate, stemming from the original Camerer et al. (1997) paper, has existed in the literature over the sign of the wage elasticity of drivers and the role of reference dependence in labor supply. It is typically argued that reference dependence and the traditional model of labor supply are incompatible as the negative wage elasticity is best explained by reference dependence rather than the income effect dominating the substitution effect. However, I find that drivers whom I identify as reference dependent in fact have significantly higher estimated wage elasticities in both San Francisco and New York (Mann-Whitney rank sum test, $p < 0.001$).[15]

It is difficult to draw generalizations about driver labor supply from these distributions. However, if the variation in the relationship between wage and propensity to quit can be explained by other factors, then driver responses to a wage may still yield useful predictions for labor supply. To examine this possibility, I use individual data on

[13]Borjas (1980) gives a more detailed description of division bias and Camerer et al. (1997) gives the first example of a similar instrument being used to correct for it with taxi data.

[14]For San Francisco, the median elasticity is -.26 with an elasticity estimated for 871 drivers. For New York City, the median elasticity is .07 with an elasticity estimated for 32,866 drivers.

[15]In San Francisco the median elasticity of a non-reference dependent driver is -.56 compared to a median elasticity of -.12 for a driver identified as reference dependent. For New York City the median non-reference dependent wage elasticity is .03 compared to a median elasticity of .08 for reference dependent drivers.

Figure 1.13. Individually estimated driver wage elasticities

Note: Top and bottom 5% of estimates are omitted.

drivers in the San Francisco data and relate their individual characteristics to their labor supply characteristics.

Farber (2015) explores the relationship between estimated elasticities and experience level. However, the New York City data do not actually provide data on driver experience and so experience is inferred by the amount of time a driver is seen in the data. Using data on when a driver first received his "A-Card", a license to operate a taxi, I can examine the relation between a San Francisco driver's actual experience driving, his propensity towards reference dependence, and his estimated elasticity.

Figure 1.14 plots a driver's estimated reference point against his years of experience. There is no theoretical reason to believe a driver's reference point should be correlated with his years of experience driving which indeed seems to be the case.

Figure 1.14. Scatter plot of years of experience vs. estimated reference point

However, we might expect that a driver who has been driving for longer would have a more well-defined reference point. This is especially true in the Koszegi and Rabin (2006) model since a more experienced driver would have more experience from which to form his earning expectation. Figure 1.15 plots the empirical CDFs for years of experience by whether or not a reference point was estimated for a driver. While there is no clear relationship between the two distributions in the first-order stochastic dominance sense, it does appear that less experienced drivers tend to be the ones who are reference dependent.

Next, I evaluate the claim that drivers who have more experience, exhibit higher estimated wage elasticities. For comparison to Farber (2015), I first look at driver experience as the number of shifts a driver has driven rather than the amount of time he has been licensed to operate a taxi. Figure 1.16 shows the CDFs for the number of shifts observed in the data conditional on the estimated elasticity being positive or negative. Like in Farber's analysis, drivers with a positive wage elasticity are do appear to have

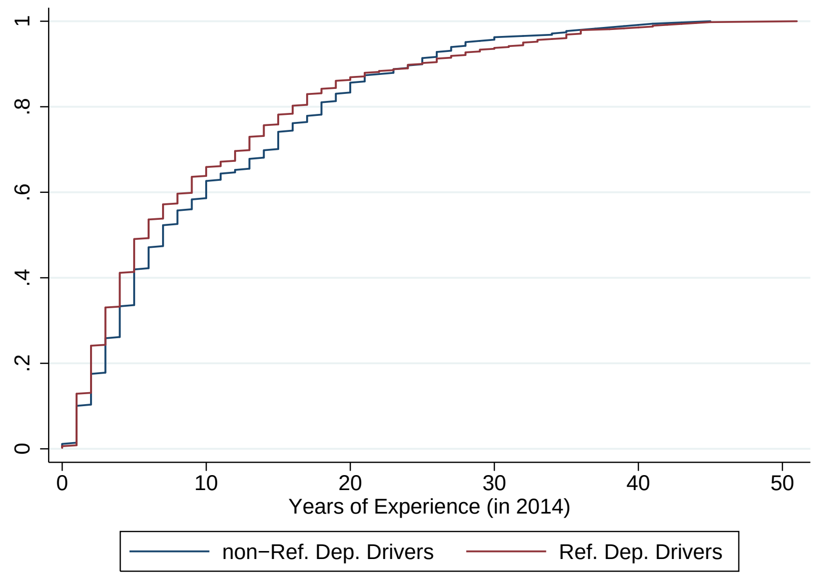

Figure 1.15. CDFs of driver years of experience by reference dependence identifiability

more shifts driven in the data. However, the opposite is true for San Francisco.

It is possible that shifts driven is simply not a good proxy for a driver's actual experience and thus I turn to San Francisco where I have data on a driver's actual experience. Figure 1.17 plots years of experience against the wage elasticity estimated for that driver. There is no discernible relationship in the San Francisco drivers between wage elasticity and driver experience. In fact, if any relationship at all is to be discerned, it appears that more experienced drivers in San Francisco have a smaller wage elasticity.[16]

Finally, I examine the relationship between labor supply characteristics and contract structure. The vast majority of drivers in the San Francisco data do now own their own medallion.[17] I categorize drivers as those who hold a medallion of any kind versus those who do not own any medallion. I find that drivers who own their own

[16]Comparing drivers with 0-5 years of experience to drivers with more than 10 years of experience, the later group has a significantly (Mann-Whitney rank sum test, $p < 0.001$) lower wage elasticity. The median wage elasticity for drivers with 0-5 years of experience is -.099, whereas the median wage elasticity of drivers with more than 10 years of experience is -.369.

[17]817 drivers in the sample do not have a medallion. 77 drivers have a "Post-K" medallion which comes with restrictions on ownership such as non-transferability and carries a driving requirement.

Figure 1.16. CDFs of total shifts driven by sign of estimated wage elasticity

Figure 1.17. Scatter plot of years of experience vs. estimated wage elasticity

medallion have higher estimated reference points conditional on being identified as reference dependent (Mann-Whitney rank sum test, $p < 0.05$).[18] However, drivers who have a medallion are no more likely to be identified as reference dependent than those without, for both populations, about 58% of drivers have an identifiable reference point.

There is no relationship between a driver's contract structure and his estimated wage elasticity (Mann-Whitney rank sum test, $p = 0.686$). Given the degree of variation in the wage elasticity, lack of clear correlation between expected wage and hazard of stopping, and no discernible relationship between either of the aforementioned variables and driver characteristics, I find it hard to believe that taxi labor supply is better modeled by the traditional Neo-Classical labor supply model than as a choice under uncertainty with or without reference dependence. I discuss this further as well as implications for future research in Section 5.

1.5.3 Predicting Labor Supply from Wages

To this point, I have outlined evidence that most drivers' preferences over daily labor supply have the following features: 1) the decision to stop working is narrowly bracketed within a day, 2) daily income is evaluated relative to a reference point, and 3) there is increasing sensitivity to changes in income below the reference point and decreasing sensitivity to changes in income above the reference point. While most of the literature on taxi driver labor supply has focused heavily on the first two points, the third point is, to the best of my knowledge, novel to the labor literature on flexible labor supply. The third point is perhaps also the most relevant for understanding driver labor supply since it implies that a driver's risk preferences play a role in his labor supply decision.

Evidence that risk preferences play a role in driver labor supply, while examined in only a limited capacity in the literature, should not be particularly surprising. A driver

[18] The median estimated reference point for a non-medallion driver is \$195 as opposed to a median estimated reference point for Post-K medallion holders of \$265.

does not face a fixed wage. A driver may be probabilistically sophisticated and face a conditional distribution of wages for a given hour, day of the week, etc. Both myself and Farber (2015) find that the average wage within an hour varies little from an estimated expected wage.[19] However, given any degree of uncertainty in wages, a driver's labor supply will depend on his preferences over risk (for more detail on this, see Chiu and Eeckhoudt (2010)).

Given that drivers face uncertainty in their wages and that they appear to be non-risk neutral within the bracket of daily income, I now attempt to evaluate the predictive power of simply an expected wage, without taking into account uncertainty on the driver's part, for labor supply. To do this, I estimate parameters for the labor supply function below:

$$h_s = k w_s^{\hat{\varepsilon}_d} + \gamma X + \eta \tag{1.7}$$

The labor supply function used assumes a constant wage elasticity, as is assumed in estimating a meaningful wage elasticity. Here, h_s is the total hours worked in shift s. $k w_s^{\hat{\varepsilon}_d}$ is the functional form of a labor supply function with constant wage elasticity. In this function, w_s is the expected or average wage for shift s, depending on specification. For a driver d, I use their estimated wage elasticity, $\hat{\varepsilon}_d$ from Figure 1.13 for the labor supply function of that driver. Thus, I am estimating the average constant term k for the population of drivers. Though theory would suggest no other parameters or a constant is needed, I estimate this model with and without a constant and including fixed effects for whether the shift was an AM or PM shift, the day of the week, month, year, and driver.

Table A.10 in the Appendix contains the results of estimating the expected wage,

[19] Farber's methodology for this analysis differs somewhat from mine. I construct an average wage in an hour from all observed earnings and regress that wage on observables. Farber, on the other hand, regresses average hourly log wage on a set of year indicators and then regresses the residuals of this first stage on a set of observables to decompose "anticipated" wage variation from unanticipated wage variation.

Table 1.3. Estimated Labor Supply Parameters (SF)

VARIABLES	(1)	(2)	(3)
k	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
PM Shift			-0.843
			(0.103)
Constant		8.209	8.409
		(0.046)	(0.099)
Observations	79,052	79,052	79,052
R-squared	0.005	0.004	0.215
RMSE	8.548	2.506	2.236
Additional Fixed Effects	No	No	Yes

Note: Robust standard errors clustered by driver in parentheses.

Table 1.4. Estimated Labor Supply Parameters (NYC)

VARIABLES	(1)	(2)	(3)
k	0.078	-0.005	-0.006
	(0.002)	(0.000)	(0.002)
PM Shift			-1.416
			(0.018)
Constant		8.810	8.886
		(0.008)	(0.021)
Observations	6,568,133	6,568,133	6,568,133
R-squared	0.059	0.003	0.345
RMSE	8.863	2.544	2.067
Additional Fixed Effects	No	No	Yes

Note: Robust standard errors clustered by driver in parentheses.

w_s, for a shift in each city. Tables 1.3 and 1.4 contain the estimated coefficient k for the labor supply function given the previously estimated elasticity for a given driver. For both cities, k is essentially zero.[20] Using the estimated wage elasticity in a corresponding labor supply function yields no predictive power for how many hours a driver will work. If any explanatory power can be gleaned from this model, it is derived from fixed effects for time of day, day of week, etc. and not from the expected wage for the shift. This result holds if the actual average wage for the shift is used instead of the expected average wage for a shift. Results for this specification can be found in the Appendix in Tables A.11 and A.12.

This result, or lack of, is in line with a calculated average hourly wage across an entire shift having an inconsistent effect on a driver labor supply. Rather than responding to an ex-post calculated average wage, the bulk of evidence I find points to drivers supplying labor dynamically using a process driven by their risk preferences. A model to explain flexible labor supply in this environment must account for the stochastic nature of wages beyond treating the first moment of such a distribution similarly to a fixed wage.

1.6 Discussion

Previous papers exploring the daily labor supply of taxi drivers have focused on explaining estimated negative wage elasticities with a model of loss aversion. This has produced mixed evidence of loss aversion in this setting. This paper builds on the previous literature's models and methodology to further explain the processes behind a driver's decision to stop working. In estimations of a model which is ex-ante agnostic to the presence of reference dependence, I identify an average effect of cumulative income on shift stopping hazard which is consistent with a previously unexamined variation of reference dependence, Prospect Theory.

The intuition for Prospect Theory reference dependence is somewhat different

[20] Estimates are only for drivers with elasticities between the 5th and 95th percentile of estimated elasticities which correspond to the drivers used in Figure 1.13. The result of a k being estimated as close to 0 is robust to specifications using median regression.

from that of loss aversion. Under loss aversion, drivers receive more disutility from being a given distance below their reference point than they receive utility from being that same distance above the reference point. The reference point is taken to be a driver's income target. In a Prospect Theory model without loss aversion, drivers also have an income target but they receive more utility per dollar earned approaching their reference point and less utility for dollars earned beyond their reference point. In a stopping sense, this means drivers will be less likely to quit if they are close to their income target.

A driver's work environment is inherently filled with uncertainty. A driver does not know ex-ante how much he will make from his next fare or how long his next fare will take. A driver must also search for his next fare which requires uncompensated time and effort. In this sense, a driver's wage is subject to substantial variation. My analysis confirms previous results that a substantial portion of the average wage in an hour is explained by observables. However, in an individual level analysis, I show that this high degree of anticipated variation does not appear to translate to consistent responses from drivers to an average wage. Moreover, hazard rates are consistent with drivers having differential risk preferences around an income target.

The results presented in this paper tie together many, seemingly contradictory, results from previous papers. Linear estimations of how cumulative income correlates with the hazard rate, as done in Farber (2005) and Crawford and Meng (2011), would not detect the non-linearities I estimate. However, splitting the sample by early earnings, as in Crawford and Meng (2011), would produce positive linear hazard rates in income given the underlying non-linearities I estimate. Moreover, non-linear hazard rates in income are smooth, as in Farber (2008).

Many of the results presented are in line with Farber (2015), in particular, positive average wage elasticities for New York City. However, these positive elasticities do not generalize across cities. Additionally, a Prospect Theory model of reference dependence could produce either positive or negative average wage elasticities since wage responses throughout the day would be variable. Thus, while estimates of the elasticity may be positive or negative, these estimates would not yield strong predictors of driver labor supply as I show in my analysis. Lastly, the binned hazard of stopping analysis conducted by Farber (2015) does yield strong evidence of a non-monotonic relationship between income and hazard of stopping. I replicate Farber's findings using his binned model but show that non-parametric estimation of a continuous function reveals non-monotonicities not detected by the bins used by Farber (2015) but consistent with Prospect Theory.

This particular model of reference dependence has numerous practical applications. Aside from the direct implications for the relationship between shift survival and daily income, a better understanding of driver preferences could lead to better contracts or policy in this industry. In particular, this model has surprising predictions for how drivers will respond on the intensive margin to reductions in their entry fee for the day suggesting that charging any entry fee may be a sub-optimal contract compared to a sharecropping model. More work is needed to understand how driver preferences can be used to optimally incentivize driver behavior from either a firm or policymaker's perspective.

Acknowledgements

Chapter 1, in full, has been submitted for publication of the material as it may appear in the Journal of Economic Behavior & Organization, 2016, Leah-Martin, Vincent. The dissertation author was the sole investigator and author of this paper.

Chapter 2

Relative Compensation and Job Satisfaction

2.1 Introduction

This paper seeks to build on economists' understanding of how compensation and other employment factors affect a worker's job satisfaction. In particular, I closely examine the role of an employee's compensation relative to similar workers as a determinant of job satisfaction and outlook. To this end, I use data obtained from Glassdoor, a job review website which obtains user's compensation and reviews of their current or former jobs and makes the results available to other users. Using these surveys along with H1-B visa data, I am able to determine how a worker's compensation compares to peers along a variety of dimensions including job-specific peers, geographic peers, and industry-specific peers. I then establish a relationship between worker job satisfaction and compensation relative to a variety of reference groups.

Previous literature has extensively documented the importance of subjective job satisfaction ratings for economic and business relevant outcomes (Freeman, 1978). In particular, job satisfaction has been widely shown to be closely linked to employee retention in both the private and public sectors (Clark, 2001; Lévy-Garboua et al., 2007; Böckerman and Ilmakunnas, 2009; Green, 2010; Turkyilmaz et al., 2011). Internal research conducted using Glassdoor data shows that an index of publicly traded companies with high employee satisfaction outperform stock market averages (Chamberlain, 2015). Finally, job satisfaction and worker performance has been widely linked in the industrial psychology literature though, as outlined by a survey of such literature conducted by Judge et al. (2001), the exact mechanisms behind this relationship are not clear.

The link between income and subjective happiness, while intuitive, is at best, empirically murky. Macro-level analyses have found little or no correlation between income and reported happiness within countries (Easterlin, 1974) though more recent studies using more comprehensive datasets do indicate the two are positively correlated (Sacks et al., 2010). Kahneman et al. (2006) suggest that individuals overstate the importance of income to their lifetime happiness and that elation from raises are, ultimately, short-lived.

To account for seemingly little evidence of a direct link between absolute income and happiness, a large body of research has focused on the link between relative income and happiness (Brickman and Campbell, 1971; Clark et al., 2008). At the core of this theory is that an individual's utility is not based on his absolute income, but rather is based on the individual's income's standing relative to the incomes of his peers or others in a "reference group". However, an important limitation to this research is that the reference group is not necessarily a well-defined group and could potentially be as broadly defined as an individual's countrymen or as narrowly as his immediate neighbors.

Clark and Oswald (1996) find evidence that supports the notion that relative income, rather than absolute income, is a key determinant in reported job satisfaction among a panel of British workers. Ferrer-i Carbonell (2005), using German panel data, finds that comparison income to a reference group is a key element in overall happiness and that comparisons play a stronger role for individuals at the bottom of the reference group's income distribution. Brown et al. (2008) use a similar panel dataset of British workers and find that ordinal rank within a comparison group plays a larger role in job satisfaction and quits than relative wages in an absolute sense. Moreover, they find evidence that further moments of the wage distribution play important roles in determining employee satisfaction with higher wage skewness inducing more quits. Pfeifer and Schneck (2012) suggest that relative wages can also play a significant role not only in a worker's propensity to quit his job, but also can be highly indicative of relative earnings at future jobs.

Conceptually, having a lower wage relative to co-workers may have either a positive or negative effect on job satisfaction. Having lower wages than co-workers, on one hand, may be construed as an indication that future earnings in a position will be higher, a phenomenon referred to in past literature as a signaling effect. On the other hand, workers may be jealous of their co-workers' higher compensation inducing a negative status effect. Clark et al. (2009) suggest the former has a stronger effect in an examination of Danish panel data and administrative records while Gao and Smyth (2010) find the latter to have a stronger effect on employee satisfaction using multiple metrics of satisfaction from Chinese firms. Bygren (2004), on the other hand, argues that pay satisfaction in particular is more highly correlated with pay relative to larger, more general reference groups such as within occupation or nation.[1]

The importance in understanding how workers respond to changes in relative compensation is highlighted by a growing literature on within and across firm inequality. Song et al. (2015), for instance, find that most recent wage inequality growth in the United States can be attributed to differences in wages between employers rather than within employers. This suggests that the effects of relative income on happiness would be stronger if a worker is evaluating his earnings relative to a wider reference group. Barth et al. (2016) support this notion, finding that variance in earnings across firms from the

[1]It is important to note that while Clark et al. (2009) and Gao and Smyth (2010) examine overall job satisfaction, Bygren (2004) restricts his analysis to pay satisfaction.

1970s-2000s was a key driver in earnings variation amongst workers.

Despite a plethora of research in understanding how relative compensation affects worker job satisfaction, there are important limitations with previous work. In most cases, research has been done using public survey panel data which offers little guarantee that employees actually know the results of their standing in the distribution of compensation relative to a given reference group, be it nationally or their co-workers. Studies that have not relied on survey data across numerous jobs, on the other hand, have typically been limited in scope to only one field or employer. Lastly, it is unclear what reference group is most natural for a worker to consider his income in relation to.

By using Glassdoor's user generated data, I hope to be able to extend this rich line of research on relative compensation while addressing many of the data limitations of previous studies. Glassdoor.com provides users with a way to directly compare their compensation to others at their employer, within jobs across employers, and across differing geographic locations. Due to the nature of what users use Glassdoor for, I can be relatively confident that 1) users are comparing their earnings to some other group and 2) users have substantial information at their disposal about pay distributions for their job relative to a variety of reference groups.

Since Glassdoor is not limited to collecting user information for a particular sector, field, or job type, my data contain a wide variety of occupations and a much larger sample size than that of previous work. Additionally, by combing Glassdoor and H1-B visa data, I am able to construct a more accurate picture of the distribution of compensation for a given job.

The paper is organized as follows: section 2 describes the data, its generating process in detail, and provides summary statistics. Section 3 contains my core analysis of both absolute and relative income effects on job satisfaction. Section 4 discusses my results in the context of previous findings.

2.2 Data

The data obtained from Glassdoor consist of three parts: 1) a dataset of Glassdoor user reported compensation data, 2) a dataset of H1-B visa salary data, and 3) a dataset of user generated job satisfaction surveys.

The Glassdoor user reported compensation dataset contains 2,918,870 observations, each being compensation broken down by base pay, cash bonus, stock bonus, profit sharing, sales commission, and tips for a user's job in a reported year. These data also contain information such as the name of the employer, job title, reported years of experience, metro, state, and, in some cases, the individual's birth year, education level, and gender. There are 2,734,708 unique users reporting in these data, indicating that some users have reported compensation for multiple jobs. The data span 147,235 unique job titles across 250,755 unique employers.

The data of salaries from H1-B visas are similar though not as detailed. The H1-B visa data Glassdoor collects represents first offers by employers to H1-B visa candidates. These data provide no information on the individuals for whom the visas are being obtained. Additionally, all reported compensation in these data is base pay. I use these data in order to construct more accurate representations of the pay distribution for a given job and geographic area. Within these data are 1,952,923 reports for 142,359 unique job titles across 97,433 unique employers.

Lastly, the job satisfaction data are reported by users in a separate survey from the compensation survey. The survey contains all of the employer information available in the other datasets as well as gender, birth year, and education for users who volunteered the information. There are several measures of satisfaction surveyed including: overall rating (1-5 scale), career opportunities (1-5 scale), compensation and benefits (1-5 scale), senior leadership (1-5 scale), work-life balance (1-5 scale), cultural values (1-5 scale), whether the worker would recommend the job to a friend, the worker's outlook on the business trajectory, and whether or not the worker approves of the CEO. These data consist of 3,292,447 reviews from 3,065,203 unique users and span 275,101 unique jobs across 279,261 employers.

To construct my working dataset, I create a panel from all three of the available Glassdoor datasets described above. I use Glassdoor user IDs and job IDs to match salary and job reviews which results in 480,224 observations which I will be using for my analysis. I also use salary-only observations to construct wage distributions and worker compensation rankings within various reference groups.

2.2.1 Summary Statistics

Figure 2.1 shows the density of total real annual compensation in 2016 dollars for the entire panel. This figure is calculated by annualizing and summing all forms of compensation and then converting that amount to 2016 dollars. The majority of reported compensation lies in the $60,000 to $100,000 range with a second mode at the lower tail. This distribution is observed over a wide variety of job types and sectors and so limited information can be inferred from it. Nevertheless, it appears that Glassdoor users have a higher average income than the population.

Since these are self-reported data from a group of users who presumably are interested in comparing their earnings with others in their field, any results drawn from these data should be taken with caution and there is undoubtedly selection into participating on Glassdoor. Table 2.1 and Figure 2.2 look at the dimensions of selection in somewhat more detail. The workers who participated in Glassdoor surveys appear to be younger, predominately college educated, and regularly employed. The majority of workers in the data have only been at their job for a few years.

As previously mentioned, much of the literature on worker satisfaction has shown

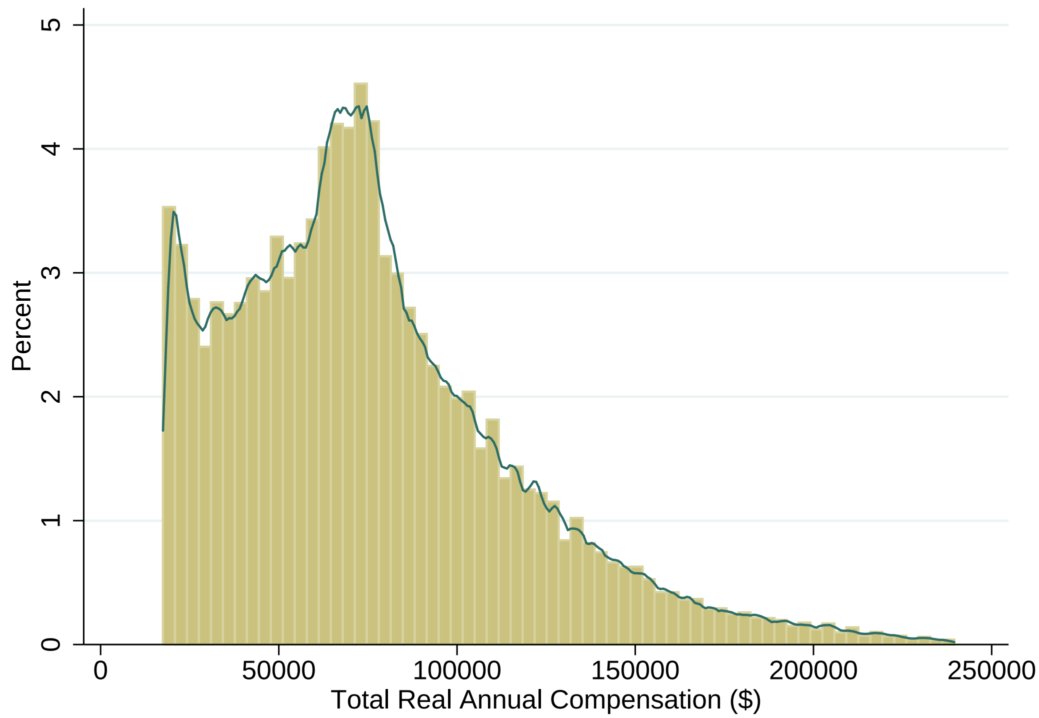

Figure 2.1. Distribution of calculated total real annual compensation

Note: Top 1% and workers earning less than $18,500 per year in total real compensation are omitted. The annual real compensation is take by scaling up part time workers' pay to their equivalent full time annual pay. Inflation is adjusted for by converting to 2016 dollars. Compensation includes the total of annualized base pay, cash bonus, stock, commisission, and tips.

that poor job satisfaction is highly linked with quits. To the extent that I am interested in this behavior, selection into participation on Glassdoor may offer some advantages in that these individuals are likely to be closer to the margin of switching jobs or are at least interested in relative compensation and conditions in their industry. Since I will not be analyzing data that is not subject to some degree of selection along the discussed dimensions, I will discuss my analysis in a way that is carefully benchmarked by other studies which have not had these selection issues. Nevertheless, given the amount of participation on Glassdoor, it is clear that these data represent a non-trivial, if select, portion of the workforce.

Table 2.1. Summary Statistics

	N	mean	sd	min	max
DEMOGRAPHIC VARIABLES					
Total Real Annual Compensation ($)	4,645,991	75,972	39,955	17,503	239,506
Age (years)	1,080,540	32.67	10.62	18	116
Length of Employment (years)	480,244	3.176	4.658	0	20
SATISFACTION VARIABLES					
Overall Rating (1-5)	480,244	3.217	1.298	1	5
Career Opportunities (1-5)	461,719	3.019	1.325	1	5
Compensation and Benefits (1-5)	462,217	3.163	1.247	1	5
Senior Leadership (1-5)	458,001	2.852	1.402	1	5
Work-Life Balance (1-5)	462,298	3.228	1.352	1	5
Cultural Values (1-5)	406,155	3.243	1.439	1	5

Note: Not all variables are reported for each observation. I omit data for individuals less than 18 years of age.

Figure 2.3 shows the distribution of satisfaction metrics measured by Glassdoor.[2] Perhaps more notable here is that for nearly every metric there is a higher variance of reviews than found in previous studies.[3] Glassdoor claims that these reviews offer a more accurate representation of feelings about a job than most review services such as those offered by Yelp for businesses or Amazon for retail products, which have heavily skewed review distributions. The lack of skew in Glassdoor's job reviews is attributed to Glassdoor's "give to get" model in which users must submit a review or salary for their job in order to access information on other user's reviews (Smart, 2016). This, Glassdoor claims, encourages individuals with more moderate opinions to contribute rather than restricting the sample to individuals with strong (either positive or negative) opinions.

One advantage to having multiple metrics is that having multiple correlated metrics has been shown to be much more reliable for inference than single metrics

[2]Not all reviews have scores for every metric.

[3]For instance, in the job satisfaction data used by Clark and Oswald (1996), over 50% of individuals rated their job a 6 or 7 on a 7 point scale while less than 5% rated their job a 1 or 2.

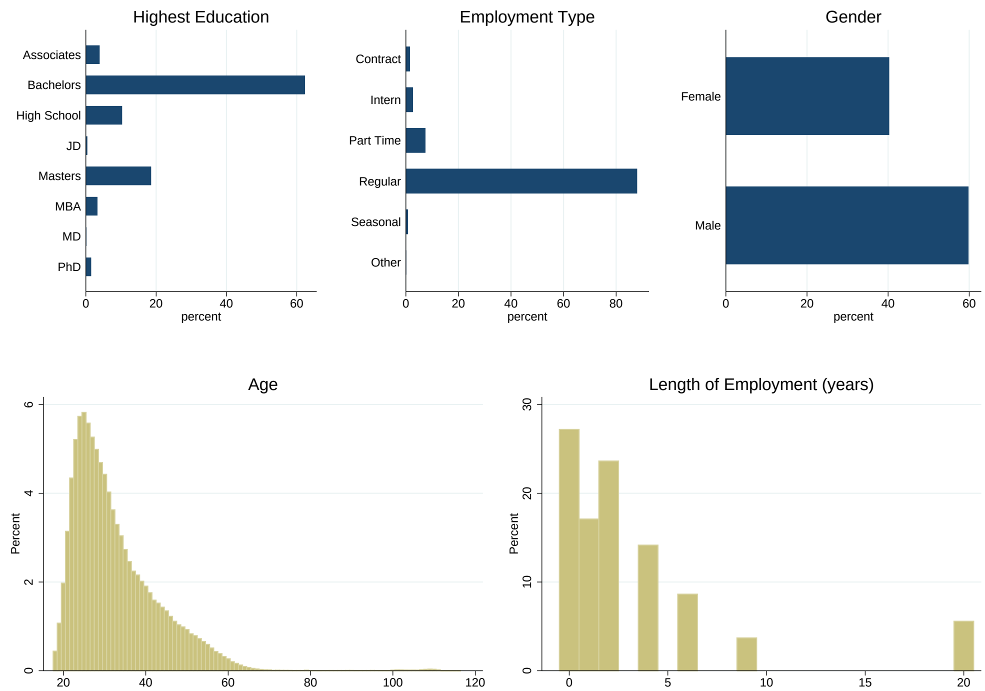

Figure 2.2. Demographics

Note: Demographic statistics reported where available. Gender statistics omitted for indivdiuals who responded “Prefer Not to State” or “Unknown”.

(Wanous et al., 1997). Krueger and Schkade (2008) find that utilizing multiple correlated metrics of satisfaction results in higher test - retest correlations. Additionally, it is clear from the distributions of responses that these metrics are all capturing slightly different aspects of worker satisfaction. For instance, while overall rating of a job is unimodal with a mean of 3.2, ratings of senior leadership is bimodal with a mean of 2.85. Thus, these metrics offer additional variation from which I can capture determinants of job satisfaction.

2.2.2 Selection

As with any dataset comprised of volunteered survey data, concerns of selection arise. With the Glassdoor data there are two selection factors I will address: the first is

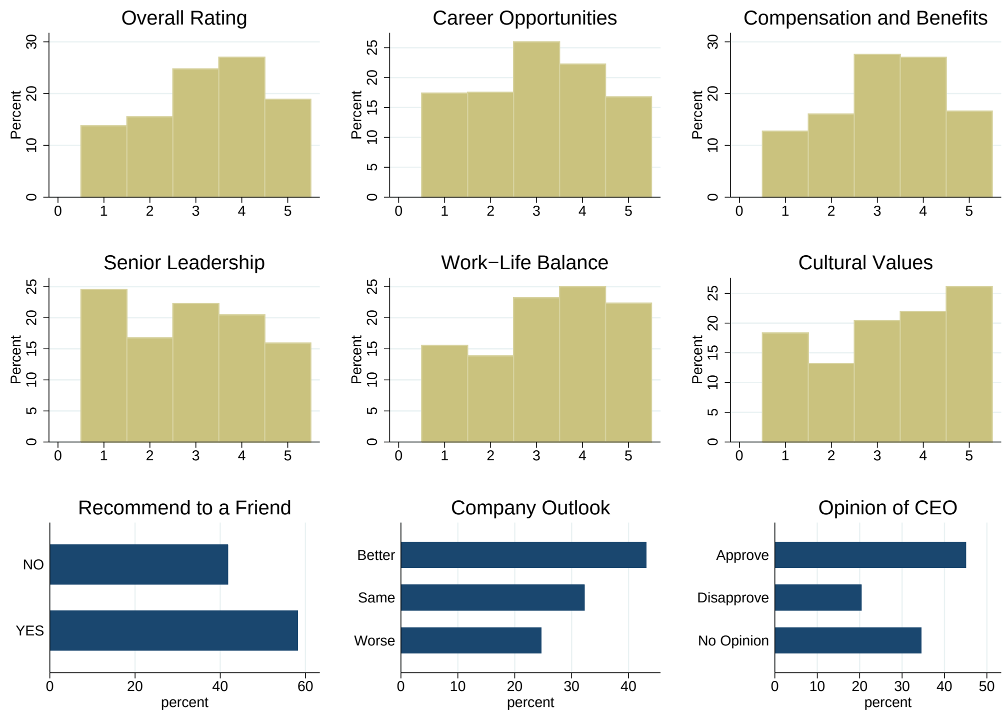

Figure 2.3. Distribution of reported satisfaction metrics

Note: While most ratings are on a 1-5 scale, the rating was not required to be an integer and, as a result, some intermediary ratings exist in the data.

selection into participation on Glassdoor. The second is selection into leaving a company review conditional on Glassdoor participation. I have little recourse regarding the former and thus my conclusions must be conditioned on the type of individual who would participate in Glassdoor job ratings. For the latter, I compare demographics and earnings for those who submitted both a salary and job review with the demographics and earnings of those who only submitted a salary (but not a job review).

Figure 2.2 provides an overview of the users who participate on Glassdoor. From these demographics, it is clear that Glassdoor users are likely not representative of the overall labor force. Glassdoor users are predominantly college educated, full time workers in the primes of their careers. Men also appear to be overrepresented among users who supplied demographic information. Due to the uniqueness of the Glassdoor data, there is no way of accounting for this type of selection in our results. With this in mind, the Glassdoor data does have an extensive range of professional jobs across a large variety of firms and so the results I present, while not generalizable across the entire labor force, are representative of career-professionals.

Figure 2.4. Demographic differences between those who left a job rating and those who did not

Note: Gender staistics omitted for indivdiuals who responded “Prefer Not to State” or “Unknown”. CDFs for age and compensation presented for comparison.

One aspect of selection in the dataset that I can address is that of leaving a review of a company. Figure 2.4 compares the demographics depicted by Figure 2.2 between those individuals who left a job rating in addition to their salary review and those individuals who only left a salary review. Among most of the observable demographics, those who left a review in addition to a salary, as opposed to just a salary, look very similar to those who only submitted a salary. One area in which the two groups are different is earnings. The median income for those who left a review is approximately $20,000 less than that of those who did not leave a review. This indicates that my analysis will overweight those who earn less than $100,000 per year in total compensation. Since this range accounts for the vast majority of workers, this is likely advantageous to my analysis.

2.3 Analysis

My analysis of worker satisfaction is divided into three parts: In the first part, I look at the relationship between absolute income and worker satisfaction utilizing each satisfaction metric separately. Second, I examine the effect relative income has on workers' "overall rating" for satisfaction using a variety of reference groups to examine which groups are most relevant and test for differential effects of comparing one's income to various groups. Lastly, I explore other determinants of satisfaction such as corporate culture and work-life balance and their importance for determining happiness relative to the importance of income.

2.3.1 Effects of Income on Job Satisfaction

Figure 2.5 shows income distributions conditional on various satisfaction ratings for each metric in my data. These are income and ratings averaged over a wide variety of occupations and demographics and thus little can be inferred from eyeballing trends in these graphs. Nevertheless, it does appear to be the case that workers with higher income tend to report higher job satisfaction across most of the metrics. Perhaps not surprisingly, income appears to be more highly correlated with some metrics than with others.

As expected, individuals with higher incomes appear to rate their compensation and benefits more highly. Income does not appear to play a very big role for many

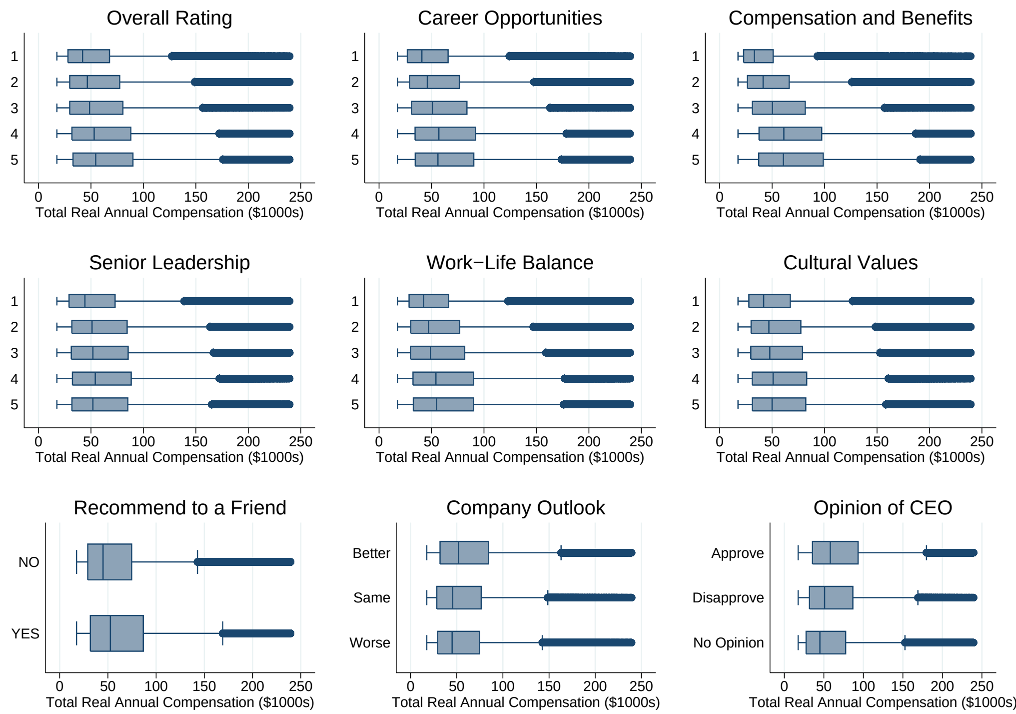

Figure 2.5. Distribution of income for each satisfaction rating

other satisfaction metrics. While the lowest rating for each satisfaction metric appears to be more tightly grouped among the lower earners, it is difficult to discern much of a noticeable difference in the distribution of income for mid-ranged (2-4) satisfaction responses. Of course, these are distributions aggregated over a wide range of occupations, demographics, and geographic areas. Fortunately, the Glassdoor data allows for me to control for many of these confounds to more clearly estimate a relationship between income and satisfaction.

A natural question that might arise for a firm concerned about worker satisfaction is: how much would a salary raise or bonus affect worker satisfaction. To begin to answer this question, I treat satisfaction as continuous and estimate an income elasticity for each of the six 5 point scale satisfaction metrics in my data. Figure 1.13 shows the income elasticity for each of these metrics and their relative magnitudes. Equation 2.1 will be my estimating equation for this baseline analysis.

$$log(\text{overall satisfaction}_i) = \beta log(\text{TRAC}_i) + \gamma X_i + \varepsilon_i \quad (2.1)$$

In Equation 2.1, β is my primary parameter of interest and can be interpreted as the elasticity of overall satisfaction with respect to total real annual income (TRAC). X_i is a vector of controls for the ith individual which include demographics such as age, education, and gender which might affect how an individual perceives their job and thus their reported job satisfaction. I also control for geography for the same reason. Lastly, these controls include indicators for the individual's specific industry, sector, occupation, employer, and if the job is the individual's current job. These controls will mitigate the possibility that employees in certain sectors or firms are systematically more or less happy than others.

Figure 1.13 shows the relative magnitudes of β with different satisfaction metrics used as the dependent variable in equation 2.1. While the income elasticity for every metric is statistically significant for each metric, in most cases it is quite small, suggesting that compensation only plays a limited role in an worker's happiness with their job. As expected, the compensation and benefits metric is most responsive to increases in earnings with an elasticity approximately double that of the other metrics. However, income elasticities for most other satisfaction metrics are more uniform. Overall rating and career opportunities are slightly more sensitive to higher pay than evaluations of senior leadership, work-life balance, and cultural values, but while the point estimates of the elasticities are statistically different, the differences in scale are negligible.

Of course, satisfaction is not measured in these data as a continuum but rather on an ordinal integer scale. Therefore, to stay true to the form of these data, I use an ordered logit to analyze the responses. In this analysis, I estimate the change in probability of

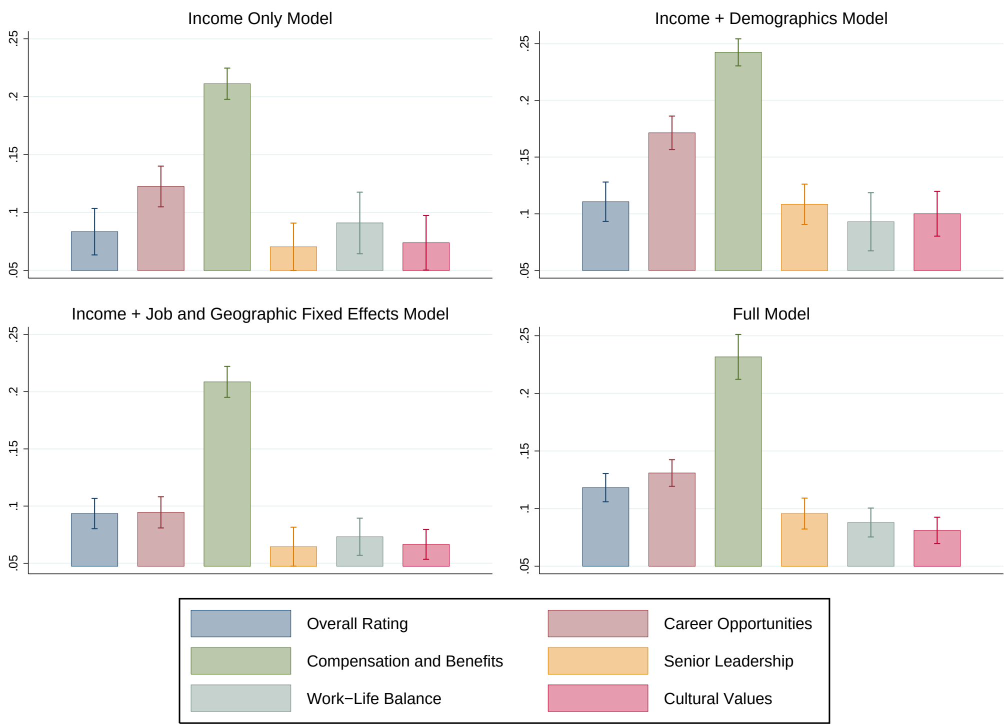

Figure 2.6. Elasticity estimates for each satisfaction rating with respect to income

Note: 95% confidence intervals calculated using standard errors clustered by major occupation group shown. Demographic controls include years of relevant experience, whether or not the job being reviewed is the respondent's current job, education level, age, age squared, and gender. Job and georephic fixed effects include indicators for the industry, sector, state, metro, occupation group, and employer.

reporting a given satisfaction level as income changes conditioning on a similar set of control variables as before.[4]

Figure 2.7 shows the estimated conditional probabilities of a given rating for overall job satisfaction as total real annual compensation changes. The ordered logit paints a similar, though more complete, picture as the raw conditioned income distributions of Figure 2.5. The probability of rating a job poorly (1 or 2) is monotonically decreasing in total compensation though not particularly quickly. As compensation crosses the

[4]The ordered logit is a computationally much more complicated model to estimate with a large number of indicator variables. As such, I use major occupation group instead of occupation, state instead of metro, and sector instead of industry and sector.

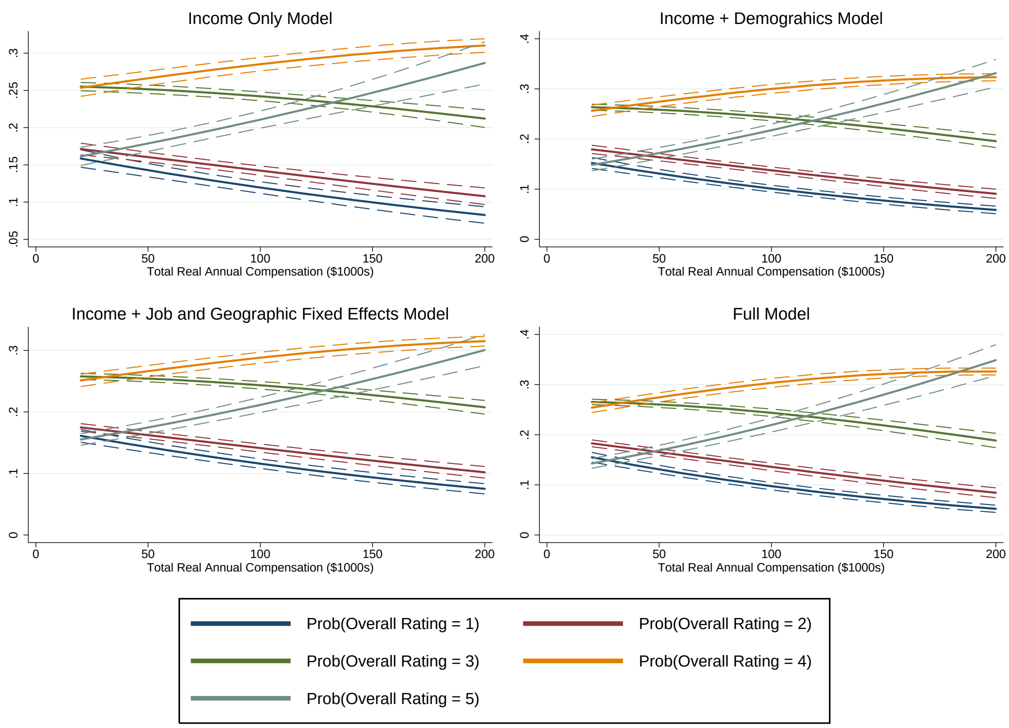

Figure 2.7. Ordered logit probabilities of overall satisfaction rating conditioning on income

Note: 95% confidence intervals calculated using standard errors clustered by major occupation group shown. Demographic controls include years of relevant experience, whether or not the job being reviewed is the respondent's current job, education level, age, age squared, and gender. Job and georephic fixed effects include indicators for sector, state, and major occupation group. Complete estimation results can be found in the Appendix.

$100,000 threshold, probability of rating the job very highly (5) grows at a faster rate while the probability of a middle rating (3 or 4) decreases or remains constant.

One important caveat to these results is that, while there is a large difference in the distribution of satisfaction between the two extremes of the compensation spectrum, there is relatively little effect of modest raises or bonuses on overall satisfaction. For instance, consider a worker earning a total annual compensation of $75,000 who rates his job as a 3 out of 5 overall. A $25,000 bonus (equivalent to a third of his current compensation) which raises his total compensation to $100,000 would only increase his probability of rating his job either a 4 or 5 by about 7 or 8 percent. This is consistent with the earlier elasticity estimations, which predict that the worker's job satisfaction in this example would increase by about 4%.[5]

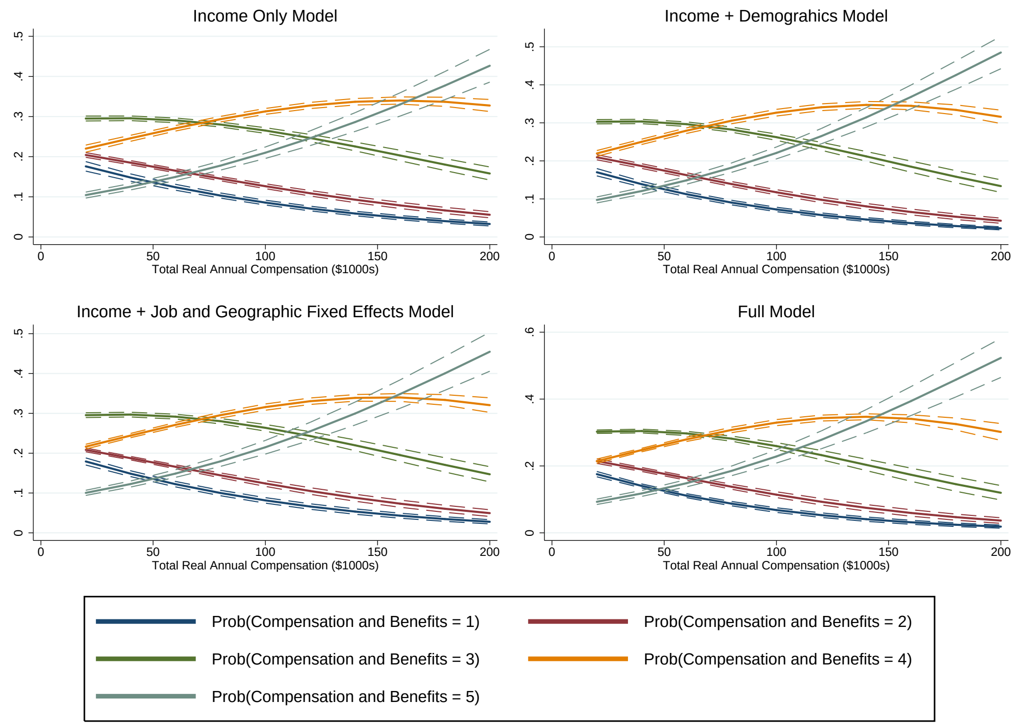

Figure 2.8. Ordered logit probabilities of compensation and benefits rating conditioning on income

Note: 95% confidence intervals calculated using standard errors clustered by major occupation group shown. Demographic controls include years of relevant experience, whether or not the job being reviewed is the respondent's current job, education level, age, age squared, and gender. Job and georephic fixed effects include indicators for sector, state, and major occupation group. Complete estimation results can be found in the Appendix.

Figure 2.8 shows probability estimates from an ordered logit on the compensation and benefit rating. For the most part, the trends in probability of a given rating are quite similar to those for overall rating. One distinction, however, is that there is a much

[5]Of course, with measures of subjective happiness, it is not entirely clear what an increase in satisfaction of 4% or a change in overall rating of a job from 3 to 4 tangibly means. Nevertheless, the gains do appear to be quite small given the substantial bonus awarded in the example.

sharper trade-off between the odds of rating compensation either 3 or 4 versus a 5 rating. This suggests that individuals with the highest incomes in this sample tend to have the strongest positive opinions about both their job and their compensation. However, the difference in slopes between probability changes in overall rating and in compensation and benefits rating suggests that higher compensation will make a worker happier with his compensation but that the effect on his overall job satisfaction are muted.

2.3.2 Effects of Relative Income on Job Satisfaction

In the previous analysis, compensation was taken without regard for potential effects of relative standing to a reference group. However, as previously described, determining which reference group is relevant is not theoretically clear. Moreover, it is not theoretically clear whether having relatively more or less income within a reference group should make an individual happier. On one hand, having relatively less income within a group might kindle jealousy of other members of the group who make more or might suggest unfair treatment. However, on the other hand, as evidenced by Clark et al. (2009), having less income than peers might signal room for growth or future earnings in a position, resulting in an individual being observed to be more satisfied while having relatively less income.

In order to disentangle the effects of relative income across various groups, I first obtain two measures of residual income. For the first measure, I obtain the residuals from regressing total compensation on occupation dummies interacted with metro dummies using all of the Glassdoor and H1-B salary data. Equation 2.2 is the estimating equation for this part of the first stage. This gives me an estimate of an individual's deviation from his expected compensation in his occupation in his locality. Next, I regress compensation on interactions of occupation, metro, and employer to give me an estimate of an individual's deviation from his expected compensation for his job and location within his firm. Equation 2.3 is the estimating equation for the later part of the first stage.

$$\text{TRAC}_i = \delta_1(\text{occupation} \times \text{metro}) + \varepsilon_{1,i} \tag{2.2}$$

$$\text{TRAC}_i = \delta_2(\text{occupation} \times \text{metro} \times \text{employer}) + \varepsilon_{2,i} \tag{2.3}$$

I begin by using a linear probability model to estimate how income, and the aforementioned residuals impact the probability that an employee is happy at his job. I define happiness as an employee having an overall satisfaction rating of 4 or 5.[6] The estimating equation for this linear probability model is given by equation 2.4.[7] Table 2.2 tabulates the estimates from this model.

$$\text{Prob}(\text{happy}_i) = \beta_0 \text{TRAC}_i + \beta_1 \varepsilon_{1,i} + \beta_2 \varepsilon_{2,i} + \gamma X_i + \tau_i \tag{2.4}$$

Just as in the direct analysis of the effects of income on employee satisfaction, the effects of income are extremely small. However, the decomposition between income, and residual income does add some nuance to these results. Without adding controls for an employee's perception of his career opportunities at his job, absolute income and a worker earning more relative to their occupation's average in their city both increase the probability of giving an overall rating of 4 or 5. However, compensation relative to the expected compensation conditional on an occupation within an employer (and metro) is negative, suggesting that having lower than expected compensation at a firm increases happiness, perhaps by signaling room for advancement. The results from

[6]I give results for alternatively defining happiness as a 3, 4 or 5 overall rating in the appendix. The qualitative results are unchanged.

[7]Note that equations 2.2 and 2.3 are estimated using a much larger dataset than equation 2.4 since the former are estimated from the universe of available salary data while the later is estimated only from observations for which I have more complete information.

Table 2.2. Linear Probability of Overall Rating 4 or 5

VARIABLES	(1)	(2)	(3)	(4)
Total Real Annual Compensation ($1000s)	0.001**	0.001***	0.001***	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Residual Compensation ($1000s)	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Residual Compensation within Employer ($1000s)	-0.000**	-0.000**	-0.001***	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.413***			
	(0.027)			
Observations	227,009	224,940	90,880	88,225
R-squared	0.008	0.039	0.087	0.398
Additional Controls	No	Yes	Yes	Yes
Demographic Controls	No	No	Yes	Yes
Career Opportunities	No	No	No	Yes

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

Note: Robust standard errors clustered by major occupation group in parentheses. I exclude residuals in the top and bottom 1% of their respective distributions. Additional controls include sector, state, and major occupation group. Demographic controls include years of relevant experience, an indicator for if the job is the respondant's current job, level of education, age, age-squared, and reported gender. Career opportunities denotes inclusion of dummy variables for individual rating of "Career Opportunities" on a 1-5 scale.

controlling for perceptions of career opportunities support this hypothesis. Once a worker's rating of their career opportunities is controlled for, both absolute income and residual income within an employer lose statistical significance. Residual compensation within an occupation, however, remains, indicating that relative income is the primary driver behind increased happiness from earning more.

To use the data in a way more appropriate to the structure of the data, I also estimate the above model as an ordered logit to better capture the nature of the ordinal ranking provided by the overall rating metric. Table 2.3 contains the odds ratios from this model. The results of using an ordered logit do not qualitatively change the findings from the linear probability model. As before, the effects of income on overall rating are quite small both in economic terms and relative to what other papers have found (Clark et al., 2009), though the ordered logit indicates a greater effect magnitude than the linear probability model.

Table 2.3. Ordered Logit Odds Ratios for Overall Rating

VARIABLES	(1)	(2)	(3)	(4)
Total Real Annual Compensation ($1000s)	1.003**	1.003***	1.006***	0.999
	(0.001)	(0.001)	(0.001)	(0.001)
Residual Compensation ($1000s)	1.005***	1.004***	1.005***	1.006***
	(0.001)	(0.001)	(0.001)	(0.001)
Residual Compensation within Employer ($1000s)	0.998*	0.998***	0.997***	0.999
	(0.001)	(0.001)	(0.001)	(0.001)
Constant Cut 1	0.168***	0.426***	0.167***	1.993***
	(0.017)	(0.065)	(0.030)	(0.303)
Constant Cut 2	0.467***	1.207	0.523***	11.165***
	(0.046)	(0.180)	(0.093)	(1.755)
Constant Cut 3	1.411***	3.764***	1.789***	75.498***
	(0.146)	(0.573)	(0.320)	(12.387)
Constant Cut 4	5.699***	15.788***	8.364***	738.813***
	(0.584)	(2.452)	(1.487)	(125.548)
Observations	227,009	224,942	90,880	88,225
Additional Controls	No	Yes	Yes	Yes
Demographic Controls	No	No	Yes	Yes
Career Opportunities	No	No	No	Yes

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

Note: Robust standard errors clustered by major occupation group in parentheses. I exclude residuals in the top and bottom 1% of their respective distributions. Additional controls include sector, state, and major occupation group. Demographic controls include years of relevant experience, an indicator for if the job is the respondant's current job, level of education, age, age-squared, and reported gender. Career opportunities denotes inclusion of dummy variables for individual rating of "Career Opportunities" on a 1-5 scale.

Consistent with the hypothesis that earning less relative to co-workers signals future opportunities, the odds ratio for residual compensation within an employer is less than one, meaning that higher earners within a firm are, on average, less satisfied with their jobs. In order to test if signaling growth within the firm is indeed the mechanism at work here, I include an ordered logit specification controlling for a worker's rating of his career opportunities. Controlling for perceived career opportunities shrinks the odds ratio of within employer residual compensation significantly and the odds ratio for both actual compensation and relative compensation within employer become very close to one and statistically insignificant indicating negligible effects on worker satisfaction.

2.3.3 Alternate Determinants of Satisfaction

One consistent theme throughout my analysis thus-far has been that the potential effects on worker satisfaction from modest increases in annual compensation appear to be quite small. While workers who work on the right tail of the compensation distribution are much more likely to rate their jobs highly compared to their counterparts on the left tail, it is neither practical nor probably a good idea for an employer to increase compensation on the scale required to achieve substantial gains to worker satisfaction. Moreover, my analysis of residual compensation would seem to indicate that income relative to others within an occupation is the only meaningful mechanism by which there are gains to satisfaction via higher earnings. Thus, in equilibrium, if all employers raised their wages with the aim of increasing worker satisfaction, there would be little net effect on happiness and employers would encounter a hedonic treadmill as discussed by Brickman and Campbell (1971).

This brings forth perhaps the most relevant question of this line of research from a business perspective: if having happier employees is good for a company, what is the best way to make employees happier? One advantage of the Glassdoor data is that it captures a multitude of satisfaction metrics targeted towards various aspects of a job. To this point, I have primarily focused on workers' "Overall Rating" of their job as the primary measure of satisfaction. However, the other metrics may capture elements of a job that workers could find more appealing than moderate compensation increases and which might be more cost effective for an employer to implement. Of course, one would expect that these metrics are all highly correlated with each other. Table 2.4 examines the degree to which the six 1-5 scale satisfaction metrics from the Glassdoor data are indeed correlated with one another.

Table 2.4. Correlation matrix of job satisfaction metrics

	Overall Rating	Career Opportunities	Compensation and Benefits	Senior Leadership	Work-Life Balance	Cultural Values
Overall Rating	1.000					
Career Opportunities	0.740	1.000				
Compensation and Benefits	0.618	0.583	1.000			
Senior Leadership	0.786	0.669	0.520	1.000		
Work-Life Balance	0.630	0.471	0.445	0.589	1.000	
Cultural Values	0.795	0.659	0.529	0.765	0.601	1.000

As expected, there is a high degree of correlation between the overall rating of the job and more targeted metrics. Nonetheless, two key takeaways emerge from looking at how the metrics are correlated with each other: first, the more specific metrics are not as strongly correlated with each other as they are with the overall rating, suggesting that they are indeed capturing worker sentiment about different aspects of his job satisfaction and that the overall rating is serving to aggregate these sentiments. Second, certain metrics are less correlated with the overall rating than others, suggesting that there are some components of employment which workers value more highly on average. In particular, compensation and benefits and work-life balance are notably less correlated with overall rating of a job than career opportunities, senior leadership, or cultural values.

In order to understand which factors are more important predictors of overall job satisfaction, I use a random forest to estimate relative importances of each satisfaction metric and various demographics. The random forest classifier attempts to predict the overall level of job satisfaction by randomly sampling the data with replacement (bootstrapping) and then estimating a decision tree regression model with the bootstrapped data. This process is repeated 10,000 times, each time using a random subset of the complete set of independent variables. For each iteration, the decision tree estimated from the bootstrapped data is used to fit the full dataset, thus measuring prediction error for that iteration. The ranked importances are computed using the normalized differences in average error with and without the inclusion of an independent variable. A higher relative importance indicates that inclusion of a particular independent variable in a model creates a better prediction of the overall satisfaction ranking.

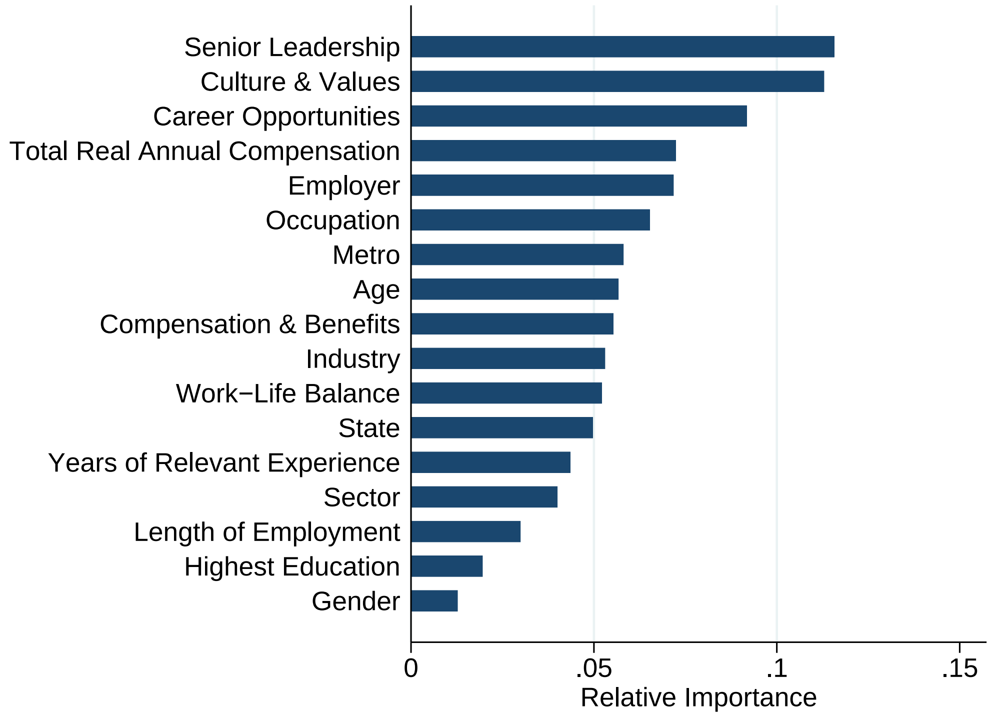

Figure 2.9. Ranked relative importance ranking from random forest classification

Note: Importances calculated using 10,000 trees.

Figure 2.9 shows the relative importances from the random forest with importances normalized to sum to one. One important caveat to note about this ranking is that random forest importances are biased towards continuous variables and factor variables of finer partition. Thus, independent variables such as total real annual compensation are likely to have overstated importance while the satisfaction metrics may have understated importance. Despite this, senior leadership and company culture stand out as the relatively most important predictors of overall satisfaction. This is consistent with the notion that other elements of the work environment play a greater role in determining overall satisfaction than pay.

Interestingly, demographics, while correlated with job satisfaction in estimations from the previous section, do not make for good predictors of overall satisfaction. This is particularly important to note for a firm since factors such as senior leadership and corporate culture can be more easily altered than the demographics of employees which thus suggests that employee satisfaction can be directly and meaningfully affected by a firm's internal policy.

2.4 Discussion

Traditional economic thought suggests that paying workers more should increase their surplus in the labor market and thus their utility. In the past few decades, more attention has been paid to the effects of wages taken in context, relative wages, and their effect on worker's utility as measured by their job satisfaction. The findings I present in this paper build on this literature by using a large dataset of employee satisfaction surveys and an even larger dataset of worker salaries provided by Glassdoor.com. Using these datasets I explore the effects of relative earnings both within occupation and within employer.

Relative compensation within an occupation appears to be a more significant driver of job satisfaction than absolute compensation. I find evidence consistent with the hypotheses of previous research suggesting that higher relative compensation within an employer negatively affects job satisfaction through a signaling mechanism by controlling for employee's reported rating of their career opportunities within their current firm. Controlling for career opportunities effectively eliminates the effects of absolute and within firm relative compensation on job satisfaction in comparison to within occupation relative compensation.

While my results are consistent with hypotheses made from previous findings, they deviate in terms of the magnitude of the effects of compensation on job satisfaction. Compensation, even relative compensation, appears to play only a minor role in a worker's happiness at his job. Other elements of the job such as the opportunities offered within a firm, senior leadership, and corporate culture appear to play a significantly larger role in determining overall satisfaction with a job. This provides a direction for future research on worker satisfaction as the exact factors which affect these aspects of a job remain an open question.[8]

Acknowledgements

Chapter 2, in part is currently being prepared for submission for publication of the material. Leah-Martin, Vincent. The dissertation author was the sole investigator and author of this material.

[8]Cappelli and Sherer (1988) address alternative factors which affect worker satisfaction within airlines, finding that workers are more satisfied with less supervision and more control over their work.

Chapter 3

Putting Risk Preferences to Work

3.1 Introduction

Standard labor supply theory assumes workers operate under a fixed wage schedule and have at least partial flexibility to set their hours. However, many occupations such as farmers, fishermen, independent contractors, taxi drivers, waiters, and the self-employed face uncertainty in their hourly wages. Such uncertainty could potentially lead to substantial variation in realized income over a given time-frame. Thus two questions arise: 1) does this variation in realized income affect an individual's labor supply decision and 2) if so, are individual's risk preferences a driving factor of labor supply with wage uncertainty?

Examining this kind of uncertainty in the field is extremely difficult, if not impossible. In most cases, the true distribution of the uncertain wage is difficult to observe and could only be estimated via possibly endogenous realizations from that distribution. Moreover, it would be impossible to measure exogenous changes to the wage distribution to obtain any kind of comparative statics for a distributional change on effort provision. Lastly, it may be impractical to obtain estimates for the risk preferences of workers in the field and even if such estimates were obtained, there is likely selection into certain industries resulting in sorting more risk-seeking individuals into particular jobs and more risk-averse individuals away from those jobs. For these reasons, this topic of research is best suited for a controlled laboratory environment in which the wage distribution may be exogenously controlled.

The objective of this paper is to empirically determine the sign of the effect of a change in wage distribution on labor supply (effort). Due to the potential complexities which could arise in the study of such comparative statics, I focus on the effect of a mean-preserving spread the wage distribution. In order to account for individual preferences, I separately elicit risk preferences and estimate an effect of the mean-preserving spread conditional for potential heterogeneous effects depending on whether or not the subject is estimated to be risk-averse or risk-loving and the strength of those preferences.

One obstacle to gaining a clearer understanding of the comparative statics of risk in a labor supply setting is a theoretical ambiguity about what the effect should be. For instance, Hicks (1963) argues that uncertainty in wages would reduce labor supply while Knight (1964) contends the opposite. Block and Heineke (1973) examine wage uncertainty in an expected utility framework and finds that changes in labor supply are ambiguous with respect to increased uncertainty in the wage.

The ambiguity can be simply explained by comparing two competing intuitions for how a risk averse worker might response to an increase in wage risk: On one hand, the risk averse worker sees an increase in earnings risk with each additional hour he works and thus is disinclined to work more as wage risk increases. On the other hand, the worker might want to insure against a negative outcome as wage risk rises and his expected earnings are increasing in hours worked which incentivizes the worker to work more under increased wage risk. Which intuition wins out to generate the overall effect of increased wage risk will depend on individual preferences and thus is reflected in theoretical ambiguity.

Chiu and Eeckhoudt (2010) formalizes the intuition described above, characterizing something akin to a Hicks-Slutsky decomposition but for changes in wage risk rather than changes in a fixed wage. This formalization allows for clearer predictions for how an individual, with particular risk preferences, will respond to changes in the wage distribution. Ultimately, whether an individual supplies more or less labor under an Nth order risk increase depends on the magnitude of the Nth order curvature of the individual's utility.[1] Jouini et al. (2013) further generalizes this framework and shows that primitive assumptions about risk preferences may provide clear predictions of economic behavior in numerous environments. I discuss these results in more detail in the next section.

Despite the difficulty in studying this topic, several attempts have been made to relate risk preferences, stochastic wages, and labor supply. Existing empirical work suggests a positive effect of wage uncertainty on labor supply. Grossberg (1989) finds that increased variance in the aggregate wage rate has a positive effect on aggregate labor supply using US time series data dating from 1950 to 1985. More targeted studies such as Low (2005) and Parker et al. (2005) suggest that individuals who face more wage uncertainty, such as younger workers and self-employed respectively, work longer hours with lower average wages, also indicating a positive labor supply response to wage uncertainty.

However, the evidence that risk preferences play a significant role in labor supply within particular settings is less clear. Chetty (2006), for instance, estimates a tight upper bound on the curvature of utility of consumption implying that risk aversion is bounded in a labor supply setting. Studying the potential for such risk preferences in settings where a wealth of individual choice data are available, in particular, with taxi drivers, has also produced mixed results as to the importance of risk preferences in settings with

[1]Ekern (1980) provides a formal definition for an Nth order risk increase which can be thought of as an increase in the Nth moment of a distribution. I define the Nth order curvature of a function as: $\frac{f^{N+1}(x)}{f^{N}(x)}$ where $f^{N}(x)$ is the Nth derivative of $f(x)$.

stochastic wages (Camerer et al., 1997; Chou, 2002; Farber, 2005; Fehr and Goette, 2007; Farber, 2008; Crawford and Meng, 2011; Farber, 2015; Leah-Martin, 2016).

With the question of the effect of interacting risk preferences and uncertainty in wages open from both a theoretical and empirical perspective, I use an experimental setting to attempt to disentangle potentially competing effects of risk preferences on changes in the wage distribution. The remainder of the paper is organized as follows: Section 2 outlines the theoretical framework for risk preferences in a labor supply setting with wage uncertainty. Section 3 describes my experimental design. Section 4 presents the results from this experiment. Lastly, section 5 discusses the results in the context of prior theoretical and empirical work.

3.2 Theoretical Framework

Understanding how an individual responds to wage uncertainty is substantially more complex than ordinary decisions under uncertainty. A decision-maker has preferences over income, I, which can be used for general consumption and is described as such in a standard labor supply model, and hours worked H. I assume his preferences can be rationalized by a differentiable utility function $U(I,H)$ where $\frac{\partial U}{\partial I} > 0$ and $\frac{\partial U}{\partial H} < 0$. For simplicity, I will also assume that I and H are additively separable in the utility function and so the individual's preferences may be written as:[2]

$$U(I,H) = u_I(I) + u_H(H)$$

In a standard labor supply mode, I is the sum of non-wage income and wage income and thus $I = I_0 + wH$ for some fixed wage w. In the setting of this paper, w is a random variable, which I will call $\tilde{w}$ whose distribution is characterized by a CDF

[2] In the experiment, the wage is piece-rate rather than an hourly wage. However, all of the intuition described in this section holds when changing hours to effort for completing a task.

$F(w)$. Without loss of generality, I will assume $F(w)$ is a discrete distribution and thus the random variable $\tilde{w}$ can be described by the pairings of potential wage outcomes and the probability of that outcome:

$$\tilde{w} = (w_1, p_1; w_2, p_2; ...w_n, p_n)$$

Where $0 \leq w_1 < w_2 < ... < w_n$. In this setting, the realization of the stochastic wages w_i is not known ex-ante and applies to all hours worked. What this means is that a decision-maker is choosing over lotteries which are a function of hours work. For simplicity, let $I_0 = 0$, then the lottery of earnings for working a given number of hours H is:

$$\tilde{w}(H) = (Hw_1, p_1; Hw_2, p_2; ...Hw_n, p_n)$$

I can now succinctly write the expected utility maximization problem of the decision maker as:

$$\max_H EU(H) = \sum_{i=1}^{n} p_i u_I(Hw_i) + u_H(H) \tag{3.1}$$

In this problem, the curvature of u_I is the determinant of a decision-maker's risk preferences. The problem, as framed above follows the setup of Chiu and Eeckhoudt (2010). In choosing the amount of time to work, the decision-maker is implicitly choosing over lotteries of outcome earnings. To better understand the trade-off a risk averse individual is making when choosing whether or not to work an extra hour, it is useful to note two properties of $\tilde{w}(H)$ in H described by the inequality in 3.2.

$$\frac{\partial E[\tilde{w}(H)^N]}{\partial H} > \frac{\partial E[\tilde{w}(H)]}{\partial H} > 0 \tag{3.2}$$

The above inequality describes the effect of working longer under stochastic wages. On one hand, working longer linearly increases the decision-maker's expected earnings. However, working longer also has the effect of increasing other moments of the earnings distribution at a faster rate than the expected value of earnings. For example, working an extra hour will increase expected earnings, which a risk averse decision-maker likes. However, working longer would also increase the variance of earnings which a risk averse decision-maker dislikes.

The above intuition is captured by the decomposition of changes in the underlying wage distribution done by Chiu and Eeckhoudt (2010). In their characterization, the overall labor supply effects for a mean-preserving spread of the wage distribution are broken down to the following components:[3]

1. Assuming $u_I'' < 0$ (the decision-maker is risk averse), labor supply decreases as the wage variance increases due to increased wage risk.

2. Assuming $u_I''' \leq 0$, labor supply increases as wage variance increases since marginal utility of additional income is increased by the increased risk.

Ultimately, the theoretical characterization of this problem leaves a somewhat ambiguous result. The result becomes unambiguous only by imposing a significant degree of structure on the curvature properties of an individual's utility of earnings. Ideally, for an applicable result, I would like to be able to say whether or not there is an effect of being risk averse or risk loving on labor supply under wage uncertainty and whether or not labor supply changes in the degree to which an individual is one or the other.

[3]The generalization of the framework described here where consumption and leisure are not additively separable includes a third effect under which the increase in risk reduces the marginal utility of leisure and thus increased labor supply is desirable.

3.3 Design

The experiment I conduct consists of two stages. The first stage is designed to measure the labor supply choice of subjects. In the second stage, I separately elicit the risk preferences of subjects. Both stages occur during the same session with one immediately following the completion of the other.

The experiment was run at the University of California, San Diego using the subject pool from the UCSD EconLab. All subjects were volunteers and were paid a $5 show-up fee. In order to mitigate possible peer effects on individual subjects' labor supply choices, each subject participated in the experiment alone in an office room and start times were staggered. At no point did subjects in the experiment see each other for the duration of the experiment.

3.3.1 Eliciting Effort

In order to simulate a labor supply environment in the first stage, I use the real effort task of Abeler et al. (2011). In this task, a computer generates tables of 150 zeros and ones. The subject's task is to count the number of zeros in the table and submit that number to the computer. If the subject counted the number of zeros correctly within three attempts, then the computer generates a new table and the subject may choose to either submit a count for the new table or stop working. Once the subject decides to stop working, the first stage of the experiment ends. Subjects were allowed to work on as many tasks as they wished for up to one hour.

Subjects earned a piece-rate wage per completed task. In order to simulate wage uncertainty, subjects were only paid for this stage based on the number of tasks completed and uncertainty was only introduced to the possible wage rate applied to all tasks completed. In a control group, subjects received a fixed $0.10 per completed task.

To test the predictions derived by Chiu and Eeckhoudt (2010), the treatment groups involved subjects being paid based on a mean-preserving spread around \$0.10. In a low variance treatment, subjects were paid either \$0.05 per task or \$0.15 per task with equal probability. In a high variance treatment, subjects were paid either \$0.00 per task or \$0.20 per task with equal probability. Subjects in the treatment groups knew the distribution of wages and that either wage was equally likely to be realized. While completing tasks, subjects were also told what their total earnings would be for both possible wages if they decided to stop at that time.

During piloting, I observed a large number of subjects were working on the counting task for the full 60 minutes. This behavior was not observed by previous experiments which used this effort elicitation device (Abeler et al., 2011; Gneezy et al., 2016). Upon reviewing the interface used by previous papers, I realized that the interface used for the counting task included a countdown timer in my version which let subjects know how much of their 60 minutes were remaining. In a second pilot, without a timer, subjects were significantly less prone to work for the entire 60 minutes. In order to test for the robustness of any potential risk preference effects, I modified the experiment to randomly treat subjects by showing them a timer or removing it. Table 3.1 shows the composition of subjects used for each treatment in the resulting 2×3 design.

Table 3.1. Number of subjects per treatment group

	No Timer	Timer	Total
Fixed Wage	24	33	57
Low Variance Wage	18	37	55
High Variance Wage	21	32	53
Total	63	102	165

3.3.2 Eliciting Risk Preferences

In the second stage of the experiment, I measure subjects' risk preferences using convex risk budgets based on those used by Andreoni and Harbaugh (2015). With this risk preference elicitation method, subjects make choices on a linear budget trading off probability of winning a lottery and the amount that the lottery pays off. Assuming that preferences are convex in this space, this elicitation allows for very precise estimation of preferences, giving subjects a very fine choice space without overwhelming the subject with options.[4]

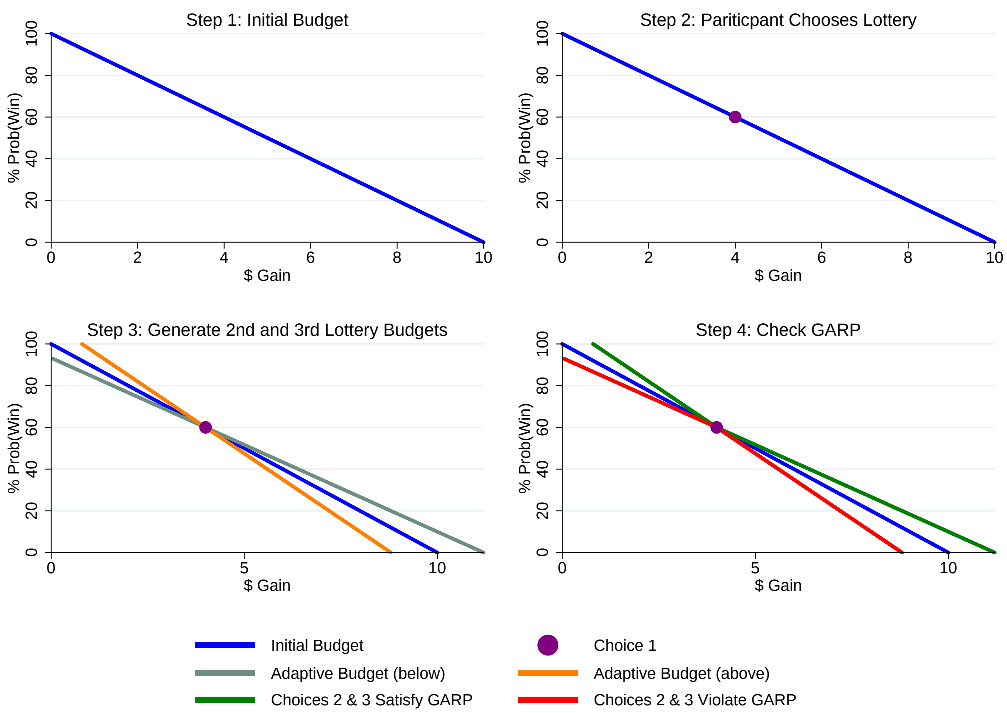

Figure 3.1. Example of budgets procedurally generated from choice under initial budget

In an effort to elicit more precise estimates of subjects' risk preferences, and to determine if a given subject's risk preferences are rationalizable by a utility function, I

[4]Subjects could change their probability of winning in 1% increments. Thus, subjects were, essentially, presented with 100 lottery options to pick between. Through the use of a slider to make this choice, subjects could easily and quickly survey their options and converge to picking the lottery they liked best.

use a technique to dynamically generate budgets based on an initial choice described by Leah-Martin and Andreoni (2016). In addition to their initial budget choice, subjects also are presented with two additional budgets. These additional budgets are determined in a way that provides a powerful test of GARP and allows for more precise risk preference estimation using fewer choices.[5]

Figure 3.1 describes the process by with the 2nd and 3rd budgets are generated so as to provide a powerful test of GARP. After the subject makes his first choice, the 2nd and 3rd budgets are created in such a way that the "price" of probability is either increased or decreased by 20% relative to the initial budget. The budgets are then positioned in such a way that they intersect the choice made for the initial budget. The budget with increased price thus intersects the original price from below and the decreased price budget intersects from above.

This elicitation method results in a non-parametric measure of risk preferences which can be easily extended for parametric estimation. For any given budget, choosing a higher probability of winning is more desirable with higher degrees of risk aversion. Given a utility function, I can also directly estimate a parameter of risk aversion which would rationalize a given choice on a budget. For simplicity, I estimate a risk parameter from a CRRA utility function, $u(x) = x^{\alpha}$. Here, more risk averse behavior is reflected in a smaller α between 0 and 1. Risk loving behavior is reflected in a larger α between 1 and ∞. For each budget, I calculate an α implied by the subject's choice given that budget. I then calculate $\bar{\alpha}$ as the average α for all three budgets if no choice violates GARP.

In order to provide a measure of risk aversion, I rescale the estimated $\bar{\alpha}$ into a coefficient of risk aversion described in equation 3.3. Figure 3.2 shows the distribution

[5] Approximately 37% of subjects' preferences violate GARP. Andreoni and Harbaugh (2015) observe approximately 35% of subjects violating.

of risk aversion for subjects with rationalizable risk preferences.

$$\rho = \frac{1 - \bar{\alpha}}{\bar{\alpha}} \tag{3.3}$$

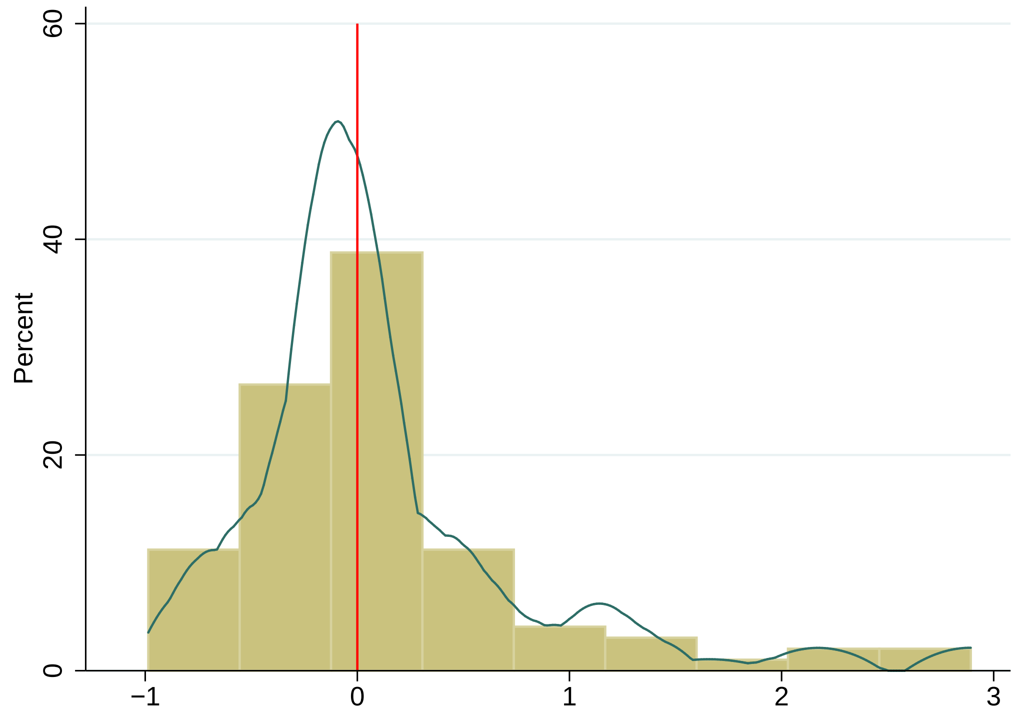

Figure 3.2. Estimated coefficient of risk aversion (ρ)

Note: $\rho = 0$ indicates risk neutrality and is indicated for illustrative purposes.

Of course, requiring that subjects' risk preferences be consistent with GARP in this setting places a large burden on an already small dataset as it results in dropping just over one-third of all observations. As a robustness check of results, I also conduct my analysis using an alternative risk preference estimate. In this alternative, if removing only one budget would allow a subject's preferences to be rationalizable, I remove that budget and estimate the subject's risk preferences from the remaining two budgets. This forces the subject's choices to be consistent with GARP. Doing so allows me to estimate the subject's coefficient of risk aversion for almost all of the subjects who violated GARP when all three budgets were considered. Figure 3.3 shows the distribution of risk preferences when forcing GARP in this manner. Estimates for all analyses done in the next section using the forced ρ may be found in the Appendix.

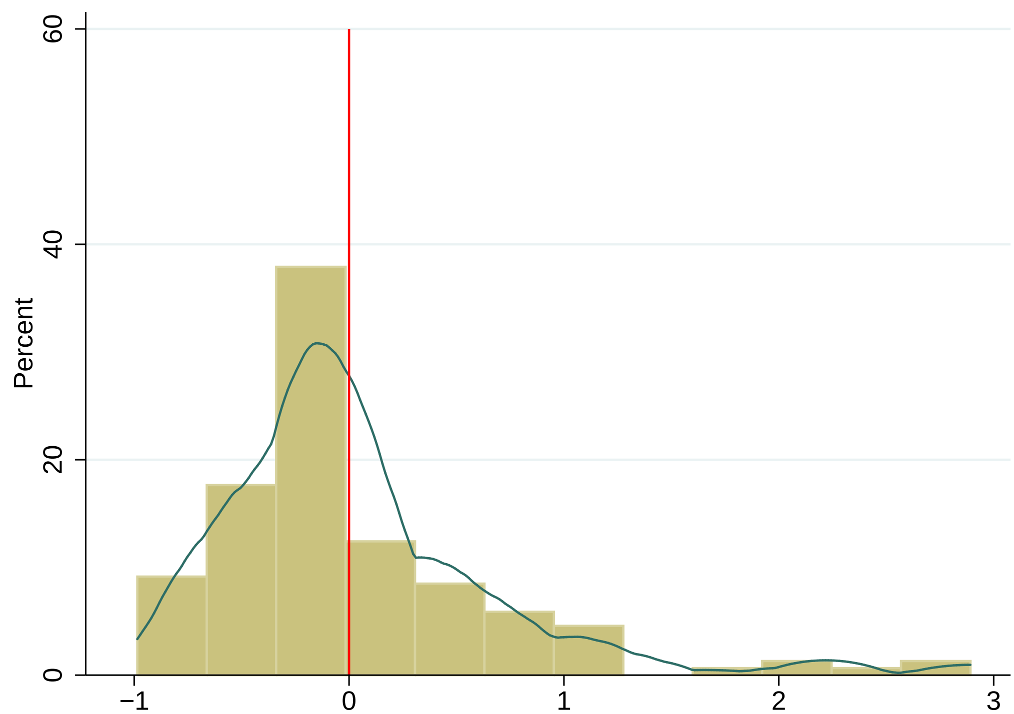

Figure 3.3. Forced estimated coefficient of risk aversion (ρ_F)

Note: $\rho = 0$ indicates risk neutrality and is indicated for illustrative purposes.

While forcing a subject's risk preferences to be consistent with GARP may not be theoretically desirable, doing so has several advantages. First, it allows for a more complete analysis of my experimental data since a much smaller proportion of observations need be dropped. Second, using these forced coefficients of risk aversion provides a robustness check of treatment effects since one may not believe that all individuals' risk preferences would be consistent with GARP. Finally, cross-validation tests show that forcing an estimation of risk preferences actually improves model fit over using only GARP-consistent subjects.

3.4 Results

In this section, I detail the experimental results of this study. One difficulty to applying theory in this setting is that the key component of the cost function in ambiguous. Effort in a real-effort task may take a variety of forms: it may be the amount of time spent working on the task which has the advantage of being able to interpret the cost function as an opportunity cost. Effort may also be a combination of unobservable elements which result in a higher output. Lastly, effort may be productivity which manifests in subjects completing individual tasks faster. I separate my analysis of the experimental results into each of these effort measures in order to provide a clearer picture of the mechanisms behind any treatment effects on effort.

For each effort measure, Equation 3.4 will be my baseline model to estimate both the correlation between risk preferences and the effort measure and the effect of introducing and increasing wage risk. In this equation, β_1, β_2 β_3, β_4, and γ are vectors of varying length of coefficients to be estimated. X_i is a set of demographic and day of the week controls for each subject. The treatment effect of increasing wage risk without a timer can be observed in the difference between coefficients within the β_3 vector. The β_4 vector will give how the timer affects risk and risk preferences in determining effort.

$$\begin{aligned}\text{Outcome}_i = {} & \beta_1 \mathbb{1}\{\text{Wage Treatment}_i\} + \beta_2 \mathbb{1}\{\text{Timer Treatment}_i\} \\ & + \beta_3 \mathbb{1}\{\text{Wage Treatment}_i\} \times \rho \\ & + \beta_4 \mathbb{1}\{\text{Wage Treatment}_i\} \times \mathbb{1}\{\text{Timer Treatment}_i\} \times \rho \\ & + \gamma X_i + \varepsilon_i\end{aligned} \tag{3.4}$$

3.4.1 Effort Measure 1: Time Spent

The first measure of effort I examine is the total amount of time spent counting 0s. As previously mentioned, this measure adds some interpretability to the cost function of an individual since time spent working is time that could have been spent doing other activities which are presumably more desirable than counting 0s for an experiment.

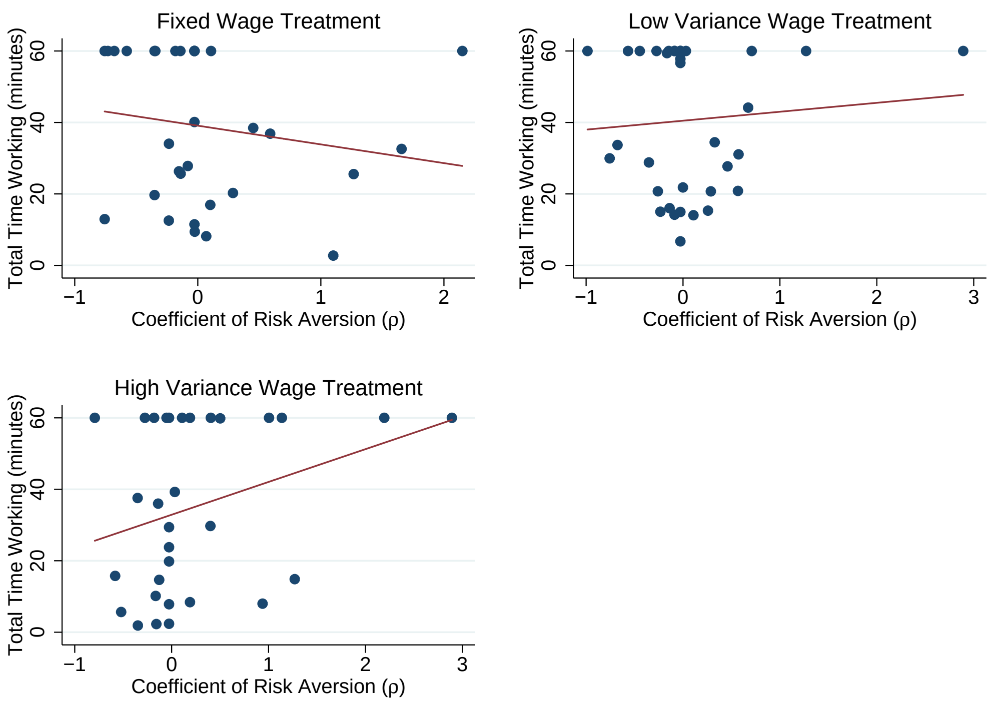

Figure 3.4. Linear relationship between estimated coefficient of risk aversion and total time spent working

Figure 3.4 depicts a best fit of total time spent working and the estimated coefficient of risk aversion broken down by wage treatment but not by timer treatment. The "Fixed Wage Treatment" serves as a control since subjects know with certainty how much they will receive from the first part of the experiment. In this treatment, total time spent working is negatively correlated with the coefficient of risk aversion. This result, intuitively, is not surprising. Risk aversion is a feature of utility functions which exhibit diminishing marginal utility of wealth. Holding the cost function constant, an individual with increasing marginal utility of wealth would work more than an individual with diminishing marginal utility of wealth under a fixed wage.

The treatment effect of increasing the wage variance can be seen in the increasing slope of the relationship between ρ and total time spent working. As the wage variance increases, more risk averse individuals appear to work relatively longer indicating that the effect of increasing earnings risk is mitigated by higher marginal utility of income from the increasing spread of earnings.

Table 3.2 provides a more detailed analysis of the treatment effect including additional control variables. As before, at baseline the effect of ρ is negative on time spent working. However, as wage variance increases, so too does the effect of increased risk aversion on time spent working with the slop increasing in the amount of wage variance. This is consistent with the summary result that risk aversion increases the amount of time spent working when there is greater variance in the wage distribution.

In order to test for the effects of including a timer, I include the same treatment effect variables interacted with the presence of a timer. I then perform an F-test to test the joint null hypothesis that risk preferences and uncorrelated with time spent working for each separate treatment group. The F-test fails to reject this hypothesis, indicating that simply including a countdown timer indicating how much time is remaining to work on the task eliminates the effect of risk preferences on time spent working. This result suggests that risk preferences, while having an effect on labor supply, may easily be overridden by other institutional considerations such as expectations and norms.

One potential issue with using OLS to estimate the treatment effect on time spent is that subjects are constrained to working no more than 60 minutes resulting in censoring above. This censoring can clearly be seen in the plots of Figure 3.4 and results in biasing the slope estimates in OLS toward 0. To correct for this, I also estimate a

Table 3.2. OLS estimates for time spent working on tasks

VARIABLES	(1)	(2)	(3)
Low Variance Wage	2.84	5.71	4.81
	(5.08)	(5.59)	(5.55)
High Variance Wage	-3.82	-2.26	-3.07
	(5.52)	(5.54)	(5.90)
Timer	10.21**	11.23**	11.32**
	(4.68)	(4.80)	(4.99)
ρ	-36.94***	-43.84***	-43.71***
	(6.51)	(7.13)	(6.64)
Low Variance Wage × ρ	17.00	28.20*	32.25*
	(16.26)	(16.31)	(16.41)
High Variance Wage × ρ	40.97**	50.33***	46.43***
	(16.31)	(15.95)	(15.06)
Timer × ρ	38.33***	44.78***	46.30***
	(8.62)	(9.19)	(8.74)
Low Variance Wage × Timer × ρ	-14.46	-24.65	-30.71*
	(16.86)	(16.79)	(16.79)
High Variance Wage × Timer × ρ	-34.30*	-42.27**	-40.97**
	(17.80)	(17.87)	(16.95)
Observations	98	98	98
R-squared	0.193	0.304	0.347
Demographic Controls	No	Yes	Yes
Day of Week Controls	No	No	Yes
Timer + Risk Effect = 0 (p-value)	0.119	0.110	0.191

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

Note: Robust standard errors in parentheses. Demographic controls include gender, age, major, and an indicator for if the subject went to high school in the United States. Timer + Risk Effect = 0 displays a p-value for an F-test of the null hypothesis that ρ is uncorrelated with the outcome for each treatment with the inclusion of a timer.

tobit specification which corrects slope and level estimates to account for the censoring. Table 3.3 shows the coefficient estimates corrected for the censoring at 60 minutes. The qualitative results remain unchanged from OLS.

Table 3.3. Tobit estimates for time spent working on tasks

VARIABLES	(1)	(2)	(3)
Low Variance Wage	2.07	5.85	5.03
	(8.18)	(8.36)	(8.07)
High Variance Wage	-6.75	-4.96	-5.71
	(8.36)	(7.76)	(8.12)
Timer	13.23*	14.51**	13.57**
	(6.97)	(6.59)	(6.76)
ρ	-59.38***	-69.82***	-71.08***
	(18.86)	(15.99)	(14.00)
Low Variance Wage $\times$ ρ	26.21	46.97*	54.50**
	(27.65)	(25.86)	(23.35)
High Variance Wage $\times$ ρ	64.28**	81.75***	77.96***
	(26.00)	(23.15)	(20.43)
Timer $\times$ ρ	59.42***	69.06***	74.61***
	(21.25)	(18.48)	(16.79)
Low Variance Wage $\times$ Timer $\times$ ρ	-16.55	-38.96	-51.20**
	(29.89)	(27.72)	(25.39)
High Variance Wage $\times$ Timer $\times$ ρ	-45.88	-60.45**	-63.54***
	(29.86)	(27.10)	(23.97)
Observations	98	98	98
Demographic Controls	No	Yes	Yes
Day of Week Controls	No	No	Yes
Timer + Risk Effect = 0 (p-value)	0.231	0.099*	0.140

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

Note: Robust standard errors in parentheses. Demographic controls include gender, age, major, and an indicator for if the subject went to high school in the United States. Timer + Risk Effect = 0 displays a p-value for an F-test of the null hypothesis that ρ is uncorrelated with the outcome for each treatment with the inclusion of a timer.

Lastly, I examine a much coarser measure of time spent working, the probability that a subject will work for the entire 60 minutes. Table 3.4 show the coefficients from this estimation. These results are again very much consistent with those of the OLS and tobit treating time as a continuous variable. With this metric, however, the F-test that the

Table 3.4. Logit estimates for probability of working on task for 60 minutes

VARIABLES	(1)	(2)	(3)
Low Variance Wage	0.15	0.30	0.34
	(0.56)	(0.65)	(0.72)
High Variance Wage	0.07	0.20	0.21
	(0.62)	(0.73)	(0.87)
Timer	1.28**	1.62**	1.39*
	(0.60)	(0.67)	(0.77)
ρ	-10.10***	-11.05***	-10.84***
	(3.26)	(3.04)	(2.98)
Low Variance Wage × ρ	3.86	2.99	1.78
	(6.25)	(8.10)	(7.96)
High Variance Wage × ρ	10.50***	12.25***	11.93***
	(3.39)	(3.11)	(2.99)
Timer × ρ	9.91***	10.77***	10.80***
	(3.35)	(3.20)	(3.13)
Low Variance Wage × Timer × ρ	-3.16	-2.55	-1.72
	(6.33)	(8.21)	(8.10)
High Variance Wage × Timer × ρ	-9.48***	-11.20***	-11.32***
	(3.56)	(3.38)	(3.23)
Observations	98	92	91
Demographic Controls	No	Yes	Yes
Day of Week Controls	No	No	Yes
Timer + Risk Effect = 0 (p-value)	0.344	0.602	0.776

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

Note: Robust standard errors in parentheses. Demographic controls include gender, age, major, and an indicator for if the subject went to high school in the United States.

timer eliminates the effect of risk preferences is much less ambiguous as the F-statistic is much smaller than with the previous models.

3.4.2 Effort Measure 2: Tasks Completed

While time spent working is a convenient measure due to clean interpretability, it may not be the most relevant metric from a practical point of view. Output, or in this setting, the total number of tasks completed, may be the desired metric from a principal-agent perspective. Additionally, output does not carry the complication that it is censored from above as time working was.

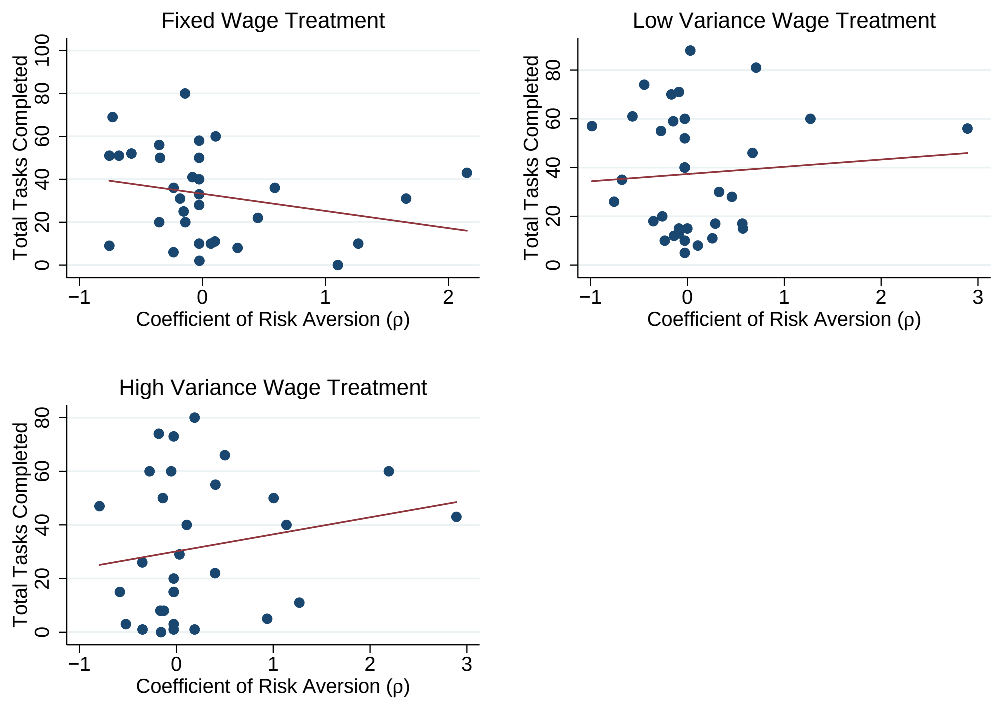

Figure 3.5. Linear relationship between estimated coefficient of risk aversion and number of tasks completed

Figure 3.5 shows the relationship between ρ and total number of tasks completed. These scatterplots, and the best fit lines for each, are somewhat easier to interpret due to censoring not being a factor. Just as with time working though, a similar trend emerges. The total tasks completed is decreasing in risk aversion under a fixed wage. However, with wage uncertainty, increased risk aversion is correlated with an increase in the number of tasks completed. The effect appears to grow larger (more positive) as the wage variance increases.

Table 3.5. OLS estimates for total tasks completed

VARIABLES	(1)	(2)	(3)
Low Variance Wage	5.41	8.07	7.89
	(5.61)	(6.08)	(5.80)
High Variance Wage	-0.66	1.28	0.77
	(6.08)	(6.17)	(6.52)
Timer	10.50**	11.44**	11.95**
	(5.22)	(5.29)	(5.79)
ρ	-39.68***	-49.40***	-51.00***
	(9.84)	(9.20)	(8.63)
Low Variance Wage $\times$ ρ	23.84	37.49**	44.69**
	(18.74)	(17.99)	(20.12)
High Variance Wage $\times$ ρ	42.42**	54.69***	50.10***
	(18.27)	(18.27)	(17.00)
Timer $\times$ ρ	38.24***	47.36***	50.50***
	(11.11)	(10.51)	(9.87)
Low Variance Wage $\times$ Timer $\times$ ρ	-18.17	-31.12*	-39.51*
	(19.32)	(18.51)	(20.32)
High Variance Wage $\times$ Timer $\times$ ρ	-35.95*	-48.02**	-45.87**
	(19.64)	(20.23)	(19.15)
Observations	98	98	98
R-squared	0.15	0.24	0.28
Demographic Controls	No	Yes	Yes
Day of Week Controls	No	No	Yes
Timer + Risk Effect = 0 (p-value)	0.419	0.524	0.604

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

Note: Robust standard errors in parentheses. Demographic controls include gender, age, major, and an indicator for if the subject went to high school in the United States. Timer + Risk Effect = 0 displays a p-value for an F-test of the null hypothesis that ρ is uncorrelated with the outcome for each treatment with the inclusion of a timer.

Table 3.5 shows more detailed OLS estimates of the effect of risk preferences in each treatment. As before, risk aversion negatively correlates with tasks completed until wage uncertainty is introduced. This is not entirely unexpected given that wage uncertainty causes risk averse individuals to work longer and time spent working is highly correlated with the number of tasks completed.

Also consistent with the previous results, I fail to reject that inclusion of a timer completely mitigates the effect of both the wage variance treatment and risk preferences. This is particularly noteworthy for external applications of these results since it would be difficult to find a labor situation without some non-risk related constraints of varying salience.

3.4.3 Effort Measure 3: Productivity

To this point, these results indicate that increased risk aversion increases time spent working and output under wage uncertainty. Of course, these two metrics are mechanically correlated and it is unclear if risk averse individuals with wage uncertainty are producing more output because they are working longer or because they are working harder. To answer this question, I leveraged detailed data from the experiment on how long it took each subject to complete each individual task.

Figure 3.6 shows the 25th, 50th, and 75th percentile of time taken to complete the nth task measured from the instant that particular table of 0s and 1s was generated up until the number of 0s was successfully counted. The distribution for each task is conditional on a particular subject reaching that task and thus the number of observations for each distribution is decreasing in the task number.

Ex-ante, one could envision the median time to complete a task to be either increasing or decreasing in the task number. If subjects are getting tired but continuing to work, then one would imagine it would take longer and longer to do each subsequent task. However, since this distribution is conditional on subjects completing the nth task,

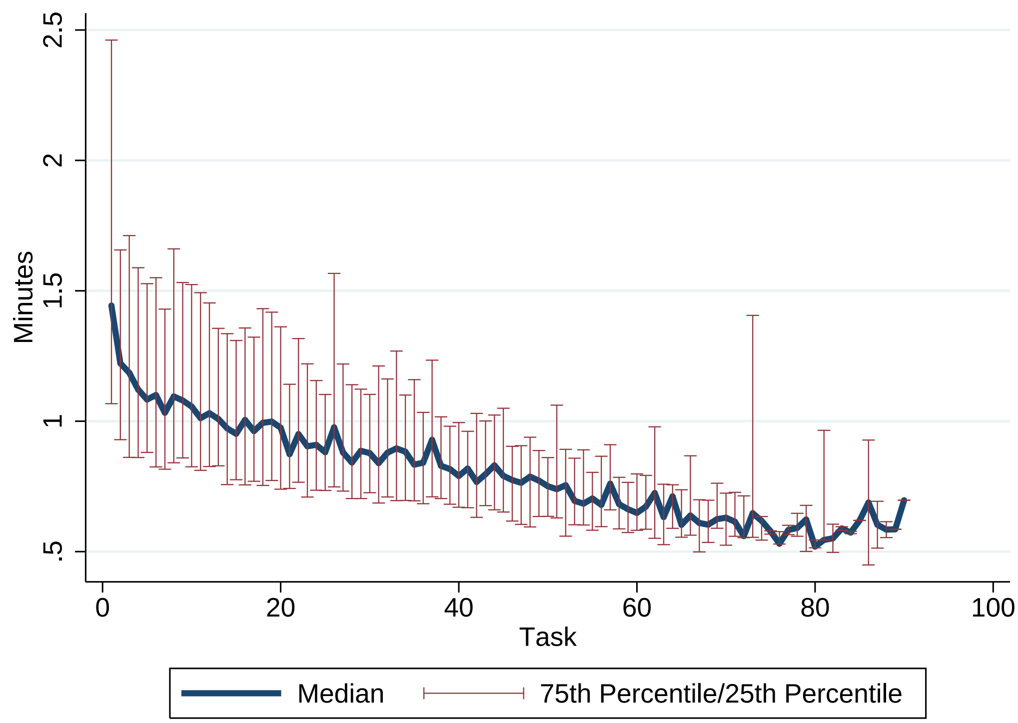

Figure 3.6. Distribution of time to complete task n

it may also be the case that more productive subjects are the ones who continue to work while less productive subjects stop earlier. Figure 3.7 indicates that the latter is the case, as subjects with a lower median task completion time, tend to work for longer.

To get a more exhaustive picture of how either risk preferences or changing the wage variance affect productivity, I examine various components of the distribution of the amount of time taken on individual tasks. The estimates from Table 3.6 show these distributional components regressed on the same model used in the previous analyses. From columns 1 and 2, risk preferences do appear to be correlated with productivity with more risk averse individuals working slower on average under the baseline fixed wage. However, with wage uncertainty, the risk averse individuals do appear to be more productive though the effect is not monotonic in the degree of wage uncertainty.

From columns 4, 6, and 7, it is clear that the effect on the average and median

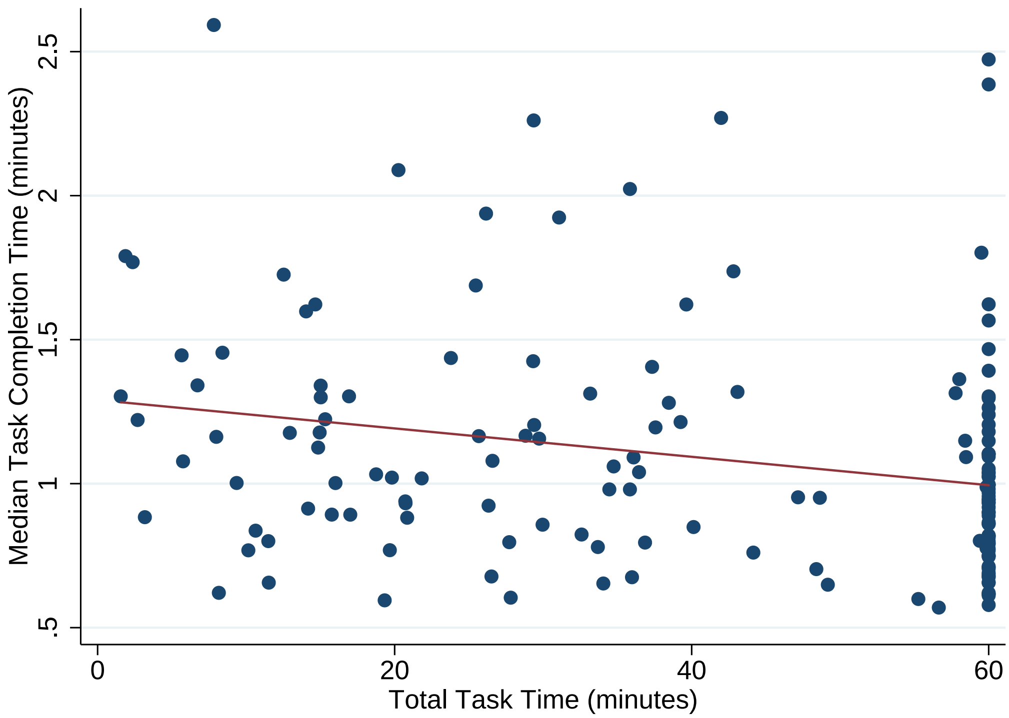

Figure 3.7. Scatter plot of total time spent working on tasks and median task completion time

task time differences are driven from shortening the bottom of the distribution of time to complete tasks. The fastest task completed is done slower by more risk averse individuals with fixed wages and is done faster with more wage uncertainty though, again, the effect of increased uncertainty is non-monotonic. Despite this shift from the bottom, there is no effect on the slowest task completed (column 3) or the spread of the time to complete tasks (column 5).

As before, the effects of both risk preferences and the addition of uncertainty are completely muted by the inclusion of a timer. This holds true for all aspects of the distribution which are statistically significantly changed by either risk preferences or the risk treatments.

Table 3.6. OLS estimates for various measures of the distribution of task completion time

VARIABLES	(1) Avg. Task Time	(2) Median Task Time	(3) Max Task Time	(4) Min Task Time	(5) Task Time Std. Dev.	(6) Task Time 25th Pctle	(7) Task Time 75th Pctle
Low Variance Wage	-0.18	-0.21	0.28	-0.07	-0.02	-0.10	-0.31
	(0.16)	(0.16)	(0.53)	(0.12)	(0.13)	(0.11)	(0.20)
High Variance Wage	-0.18	-0.14	-0.19	-0.07	-0.05	-0.10	-0.29
	(0.15)	(0.17)	(0.46)	(0.12)	(0.13)	(0.11)	(0.21)
Timer	-0.34	-0.32	0.29	-0.33**	-0.08	-0.33**	-0.46
	(0.24)	(0.24)	(0.69)	(0.15)	(0.21)	(0.15)	(0.31)
ρ	1.28***	1.31***	-0.42	1.00***	0.29	1.11***	1.63***
	(0.44)	(0.49)	(1.05)	(0.31)	(0.35)	(0.38)	(0.59)
Low Variance Wage $\times$ ρ	-2.15*	-2.09**	-1.08	-1.72**	-0.61	-1.89**	-2.52**
	(1.08)	(0.93)	(4.26)	(0.67)	(1.00)	(0.78)	(1.12)
High Variance Wage $\times$ ρ	-1.10**	-1.28**	0.36	-1.02***	-0.03	-1.23***	-0.98
	(0.51)	(0.51)	(1.67)	(0.36)	(0.45)	(0.44)	(0.67)
Timer $\times$ ρ	-1.07**	-1.04**	0.65	-0.94***	-0.18	-1.06***	-1.30**
	(0.44)	(0.50)	(1.07)	(0.31)	(0.34)	(0.38)	(0.60)
Low Variance Wage $\times$ Timer $\times$ ρ	1.89*	1.77*	0.95	1.59**	0.49	1.77**	2.15*
	(1.03)	(0.92)	(4.07)	(0.63)	(0.95)	(0.75)	(1.10)
High Variance Wage $\times$ Timer $\times$ ρ	0.91*	0.88	0.51	0.84**	0.14	1.07**	0.66
	(0.52)	(0.56)	(1.58)	(0.37)	(0.43)	(0.45)	(0.73)
Observations	96	96	96	96	93	96	96
R-squared	0.20	0.23	0.15	0.24	0.13	0.28	0.24
Demographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day of Week Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Timer + Risk Effect = 0 (p-value)	0.673	0.368	0.010***	0.236	0.052*	0.276	0.600

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

Note: Robust standard errors in parentheses.

3.5 Discussion

The study of risk preferences with respect to labor supply is inherently difficult, primarily due to the challenge of observing exogenous changes to the wage distribution. This difficulty is compounded by theoretically ambiguous predictions for how a risk averse (loving) individual should behave under wage risk. By using a controlled experimental setting, I am able to estimate a direction of effect for both risk preferences at baseline and changes in wage uncertainty on various labor supply metrics. The results suggest that: 1) risk preferences, even at baseline, can affect labor supply decisions. 2) Under wage uncertainty, risk averse individuals are relatively more productive and work longer than without wage uncertainty.

These results, however, come with a very strong caveat. The effect of both risk preferences and the wage distribution is entirely mitigated by the simple inclusion of a countdown timer, the purpose of which is to remind subjects how much time they have remaining to work on their effort task. Importantly, the time constraint was the same with and without the timer so the effect of the reminder drives the elimination of the risk and risk preference effects. In many, if not most, settings, one would expect at least some degree of salience of institutional constraints and social norms. For example, while taxi drivers may work 10-12 hour shifts, depending on their city, there exists a social norm of 8 hours being a full work day. These results indicate that such norms have an extremely strong effect and may override other preferences and certain aspects of an individual's economic environment.

Acknowledgements

Chapter 3, in part is currently being prepared for submission for publication of the material. Leah-Martin, Vincent. The dissertation author was the sole investigator and

author of this material.

Appendix A

A.1 Chapter 1 Estimation Results

Table A.1. Hazard of Stopping—Income Linear Spline Regression (SF)

VARIABLES	(1)	(2)	(3)
Shift Income \$0-\$50	-0.047***	-0.053***	-0.068***
	(0.006)	(0.005)	(0.006)
Shift Income \$50-\$100	-0.050***	-0.050***	-0.060***
	(0.004)	(0.004)	(0.004)
Shift Income \$100-\$150	-0.056***	-0.053***	-0.061***
	(0.004)	(0.003)	(0.004)
Shift Income \$150-\$200	-0.037***	-0.034***	-0.041***
	(0.004)	(0.004)	(0.004)
Shift Income \$200-\$250	-0.008*	-0.004	-0.012***
	(0.004)	(0.004)	(0.004)
Shift Income \$250-\$300	0.016***	0.020***	0.013***
	(0.005)	(0.005)	(0.005)
Shift Income \$300-\$350	0.026***	0.033***	0.025***
	(0.006)	(0.006)	(0.006)
Shift Income \$350-\$400	0.010	0.016**	0.008
	(0.008)	(0.008)	(0.007)
Shift Income \$400-\$450	-0.013	-0.006	-0.012
	(0.011)	(0.010)	(0.010)
Shift Income \$450+	0.063***	0.065***	0.055***
	(0.008)	(0.008)	(0.008)
Shift Hours	0.029***	0.031***	0.033***
	(0.001)	(0.001)	(0.001)
Warm	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Cold	0.001	0.000	0.000
	(0.001)	(0.001)	(0.001)
Rain	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)
Observations	1,549,629	1,549,629	1,505,446
R-squared	0.122	0.138	0.147
Hour of Day Effects	Yes	Yes	Yes
Day of Week Effects	Yes	Yes	Yes
Month & Year FEs	Yes	Yes	Yes
Driver Fixed Effects	No	Yes	Yes
Destination Effects	No	No	Yes

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

Note: Robust standard errors clustered by driver in parentheses. Coefficients represent linear slopes within the given income range.

Table A.2. Hazard of Stopping—Income Linear Spline Regression (NYC)

VARIABLES	(1)	(2)	(3)
Shift Income \$0-\$50	0.002**	-0.007***	-0.027***
	(0.001)	(0.001)	(0.001)
Shift Income \$50-\$100	-0.025***	-0.027***	-0.035***
	(0.000)	(0.000)	(0.000)
Shift Income \$100-\$150	-0.029***	-0.030***	-0.037***
	(0.000)	(0.000)	(0.000)
Shift Income \$150-\$200	-0.015***	-0.014***	-0.021***
	(0.000)	(0.000)	(0.000)
Shift Income \$200-\$250	0.021***	0.025***	0.018***
	(0.001)	(0.001)	(0.001)
Shift Income \$250-\$300	0.039***	0.044***	0.036***
	(0.001)	(0.001)	(0.001)
Shift Income \$300-\$350	0.013***	0.019***	0.011***
	(0.001)	(0.001)	(0.001)
Shift Income \$350-\$400	0.010***	0.015***	0.007***
	(0.001)	(0.001)	(0.001)
Shift Income \$400-\$450	-0.020***	-0.013***	-0.021***
	(0.002)	(0.002)	(0.002)
Shift Income \$450+	0.131***	0.141***	0.127***
	(0.002)	(0.002)	(0.002)
Shift Hours	0.017***	0.020***	0.021***
	(0.000)	(0.000)	(0.000)
Warm	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)
Cold	0.000***	0.000***	0.001***
	(0.000)	(0.000)	(0.000)
Rain	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
Observations	153,562,159	153,562,159	145,533,591
R-squared	0.123	0.140	0.157
Hour of Day Effects	Yes	Yes	Yes
Day of Week Effects	Yes	Yes	Yes
Month & Year FEs	Yes	Yes	Yes
Driver Fixed Effects	No	Yes	Yes
Destination Effects	No	No	Yes

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

Note: Robust standard errors clustered by driver in parentheses. Coefficients represent linear slopes within the given income range.

Table A.3. Hazard of Stopping—Income & Hours Linear Spline Regression (SF)

VARIABLES	(1)	(2)	(3)
Shift Income $0-$50	-0.036***	-0.048***	-0.104***
	(0.007)	(0.007)	(0.007)
Shift Income $50-$100	-0.020***	-0.009**	-0.030***
	(0.004)	(0.004)	(0.004)
Shift Income $100-$150	-0.013***	-0.004	-0.014***
	(0.003)	(0.003)	(0.003)
Shift Income $150-$200	-0.002	0.005*	-0.002
	(0.003)	(0.003)	(0.003)
Shift Income $200-$250	0.001	0.007**	0.000
	(0.004)	(0.003)	(0.003)
Shift Income $250-$300	-0.000	0.005	-0.001
	(0.004)	(0.004)	(0.004)
Shift Income $300-$350	-0.007	-0.001	-0.008
	(0.006)	(0.006)	(0.006)
Shift Income $350-$400	-0.025***	-0.022***	-0.028***
	(0.008)	(0.008)	(0.008)
Shift Income $400-$450	-0.037***	-0.034***	-0.038***
	(0.010)	(0.010)	(0.010)
Shift Income $450+	0.071***	0.068***	0.060***
	(0.009)	(0.008)	(0.008)

Table continues on next page.

Hazard of Stopping—Income & Hours Linear Spline Regression (SF) (continued)

VARIABLES	(1)	(2)	(3)
Shift Hours 0-1	0.028***	0.036***	0.060***
	(0.003)	(0.004)	(0.003)
Shift Hours 1-2	0.021***	0.018***	0.024***
	(0.002)	(0.002)	(0.002)
Shift Hours 2-3	0.019***	0.017***	0.022***
	(0.002)	(0.002)	(0.002)
Shift Hours 3-4	0.008***	0.007***	0.010***
	(0.001)	(0.001)	(0.001)
Shift Hours 4-5	0.008***	0.007***	0.009***
	(0.002)	(0.001)	(0.001)
Shift Hours 5-6	0.011***	0.011***	0.013***
	(0.002)	(0.002)	(0.002)
Shift Hours 6-7	0.018***	0.019***	0.020***
	(0.002)	(0.002)	(0.002)
Shift Hours 7-8	0.028***	0.028***	0.030***
	(0.004)	(0.004)	(0.004)
Shift Hours 8-9	0.148***	0.151***	0.150***
	(0.007)	(0.007)	(0.007)
Shift Hours 9+	0.014***	0.020***	0.020***
	(0.004)	(0.004)	(0.004)
Warm	0.001	0.000	0.000
	(0.000)	(0.000)	(0.000)
Cold	0.001**	0.001*	0.001
	(0.001)	(0.001)	(0.001)
Rain	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)
Observations	1,549,629	1,549,629	1,505,446
R-squared	0.135	0.152	0.160
Hour of Day Effects	Yes	Yes	Yes
Day of Week Effects	Yes	Yes	Yes
Month & Year FEs	Yes	Yes	Yes
Driver Fixed Effects	No	Yes	Yes
Destination Effects	No	No	Yes

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

Note: Robust standard errors clustered by driver in parentheses. Coefficients represent linear slopes within the given income or hours range.

Table A.4. Hazard of Stopping—Income & Hours Linear Spline Regression (NYC)

VARIABLES	(1)	(2)	(3)
Shift Income $0-$50	-0.049***	-0.067***	-0.119***
	(0.001)	(0.001)	(0.001)
Shift Income $50-$100	-0.011***	-0.014***	-0.028***
	(0.000)	(0.000)	(0.000)
Shift Income $100-$150	-0.004***	-0.002***	-0.010***
	(0.000)	(0.000)	(0.000)
Shift Income $150-$200	0.001**	0.005***	-0.001***
	(0.000)	(0.000)	(0.000)
Shift Income $200-$250	0.010***	0.016***	0.009***
	(0.001)	(0.001)	(0.001)
Shift Income $250-$300	0.009***	0.014***	0.007***
	(0.001)	(0.001)	(0.001)
Shift Income $300-$350	-0.007***	-0.004***	-0.011***
	(0.001)	(0.001)	(0.001)
Shift Income $350-$400	0.018***	0.018***	0.012***
	(0.002)	(0.001)	(0.001)
Shift Income $400-$450	0.006***	0.008***	0.002
	(0.002)	(0.002)	(0.002)
Shift Income $450+	0.165***	0.168***	0.156***
	(0.002)	(0.002)	(0.002)

Table continues on next page.

Hazard of Stopping—Income & Hours Linear Spline Regression (NYC) (continued)

VARIABLES	(1)	(2)	(3)
Shift Hours 0-1	0.043***	0.049***	0.066***
	(0.000)	(0.000)	(0.000)
Shift Hours 1-2	0.017***	0.021***	0.027***
	(0.000)	(0.000)	(0.000)
Shift Hours 2-3	0.010***	0.013***	0.016***
	(0.000)	(0.000)	(0.000)
Shift Hours 3-4	0.008***	0.010***	0.011***
	(0.000)	(0.000)	(0.000)
Shift Hours 4-5	0.006***	0.007***	0.009***
	(0.000)	(0.000)	(0.000)
Shift Hours 5-6	0.006***	0.006***	0.008***
	(0.000)	(0.000)	(0.000)
Shift Hours 6-7	0.017***	0.018***	0.020***
	(0.000)	(0.000)	(0.000)
Shift Hours 7-8	0.019***	0.021***	0.022***
	(0.001)	(0.001)	(0.001)
Shift Hours 8-9	0.078***	0.082***	0.082***
	(0.001)	(0.001)	(0.001)
Shift Hours 9+	0.006***	0.011***	0.012***
	(0.000)	(0.000)	(0.000)
Warm	0.000***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)
Cold	0.000***	0.000***	0.001***
	(0.000)	(0.000)	(0.000)
Rain	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)
Observations	153,562,159	153,562,159	145,533,591
R-squared	0.126	0.143	0.161
Hour of Day Effects	Yes	Yes	Yes
Day of Week Effects	Yes	Yes	Yes
Month & Year FEs	Yes	Yes	Yes
Driver Fixed Effects	No	Yes	Yes
Destination Effects	No	No	Yes

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

Note: Robust standard errors clustered by driver in parentheses. Coefficients represent linear slopes within the given income or hours range.

Table A.4. Hazard of Stopping—Binned Income (SF)

VARIABLES	(1)	(2)	VARIABLES	(3)	(4)
Income $100-149	-0.031***	-0.036***	Income $50-100	-0.023***	-0.028***
	(0.002)	(0.002)		(0.002)	(0.002)
Income $150-199	-0.049***	-0.057***	Income $100-150	-0.047***	-0.056***
	(0.003)	(0.003)		(0.003)	(0.003)
Income $200-224	-0.056***	-0.065***	Income $150-200	-0.068***	-0.080***
	(0.004)	(0.004)		(0.004)	(0.004)
Income $225-249	-0.056***	-0.067***	Income $200-250	-0.076***	-0.091***
	(0.004)	(0.004)		(0.005)	(0.005)
Income $250-274	-0.052***	-0.064***	Income $250-300	-0.071***	-0.090***
	(0.004)	(0.004)		(0.005)	(0.006)
Income $275-299	-0.047***	-0.061***	Income $300-350	-0.057***	-0.078***
	(0.005)	(0.005)		(0.006)	(0.006)
Income $300-349	-0.034***	-0.050***	Income $350-400	-0.047***	-0.072***
	(0.005)	(0.005)		(0.006)	(0.006)
Income $350-399	-0.023***	-0.042***	Income $400-450	-0.037***	-0.065***
	(0.006)	(0.006)		(0.007)	(0.007)
Income $400+	-0.001	-0.025***	Income $450+	-0.015**	-0.049***
	(0.006)	(0.006)		(0.007)	(0.007)
Shift Hours	0.028***	0.030***	Shift Hours	0.030***	0.032***
	(0.001)	(0.001)		(0.001)	(0.001)
Warm	0.000	-0.000	Warm	0.000	0.000
	(0.000)	(0.000)		(0.000)	(0.000)
Cold	0.001	0.000	Cold	0.001	0.000
	(0.001)	(0.001)		(0.001)	(0.001)
Rain	0.001	0.001	Rain	0.001	0.001
	(0.001)	(0.001)		(0.001)	(0.001)
Observations	1,549,629	1,505,446		1,549,629	1,505,446
R-squared	0.137	0.145		0.137	0.146
Hour of Day Effects	Yes	Yes		Yes	Yes
Day of Week Effects	Yes	Yes		Yes	Yes
Month & Year FEs	Yes	Yes		Yes	Yes
Driver Fixed Effects	Yes	Yes		Yes	Yes
Destination Effects	No	Yes		No	Yes

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

Note: Robust standard errors clustered by driver in parentheses. Columns 1 and 2 correspond to income bins used by Farber (2015).

Table A.5. Hazard of Stopping—Binned Income (NYC)

VARIABLES	(1)	(2)	VARIABLES	(3)	(4)
Income $100-149	-0.015***	-0.019***	Income $50-100	-0.009***	-0.014***
	(0.000)	(0.000)		(0.000)	(0.000)
Income $150-199	-0.025***	-0.031***	Income $100-150	-0.022***	-0.031***
	(0.000)	(0.000)		(0.000)	(0.000)
Income $200-224	-0.025***	-0.033***	Income $150-200	-0.033***	-0.044***
	(0.000)	(0.000)		(0.000)	(0.000)
Income $225-249	-0.019***	-0.028***	Income $200-250	-0.031***	-0.045***
	(0.000)	(0.000)		(0.001)	(0.001)
Income $250-274	-0.009***	-0.019***	Income $250-300	-0.014***	-0.031***
	(0.001)	(0.001)		(0.001)	(0.001)
Income $275-299	0.002***	-0.010***	Income $300-350	0.002***	-0.018***
	(0.001)	(0.001)		(0.001)	(0.001)
Income $300-349	0.013***	-0.001	Income $350-400	0.008***	-0.015***
	(0.001)	(0.001)		(0.001)	(0.001)
Income $350-399	0.020***	0.003***	Income $400-450	0.017***	-0.010***
	(0.001)	(0.001)		(0.001)	(0.001)
Income $400+	0.043***	0.021***	Income $450+	0.053***	0.019***
	(0.001)	(0.001)		(0.001)	(0.001)
Shift Hours	0.019***	0.020***	Shift Hours	0.020***	0.021***
	(0.000)	(0.000)		(0.000)	(0.000)
Warm	0.001***	0.001***	Warm	0.001***	0.001***
	(0.000)	(0.000)		(0.000)	(0.000)
Cold	0.000***	0.001***	Cold	0.000***	0.001***
	(0.000)	(0.000)		(0.000)	(0.000)
Rain	0.000***	0.000***	Rain	0.000***	0.000***
	(0.000)	(0.000)		(0.000)	(0.000)
Observations	153,562,159	145,533,591		153,562,159	145,533,591
R-squared	0.139	0.156		0.139	0.156
Hour of Day Effects	Yes	Yes		Yes	Yes
Day of Week Effects	Yes	Yes		Yes	Yes
Month & Year FEs	Yes	Yes		Yes	Yes
Driver Fixed Effects	Yes	Yes		Yes	Yes
Destination Effects	No	Yes		No	Yes

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

Note: Robust standard errors clustered by driver in parentheses. Columns 1 and 2 correspond to income bins used by Farber (2015).

Table A.6. Hazard of Stopping—Binned Income & Hours (SF)

	(1)	(2)		(3)	(4)
VARIABLES			VARIABLES		
Income \$100-149	0.011***	0.006***	Income \$50-100	-0.003**	-0.011***
	(0.001)	(0.001)		(0.001)	(0.001)
Income \$150-199	0.013***	0.006***	Income \$100-150	-0.004*	-0.017***
	(0.002)	(0.002)		(0.002)	(0.002)
Income \$200-224	0.018***	0.010***	Income \$150-200	-0.003	-0.019***
	(0.002)	(0.002)		(0.003)	(0.003)
Income \$225-249	0.021***	0.011***	Income \$200-250	0.001	-0.018***
	(0.003)	(0.003)		(0.004)	(0.004)
Income \$250-274	0.024***	0.013***	Income \$250-300	0.005	-0.016***
	(0.003)	(0.003)		(0.004)	(0.004)
Income \$275-299	0.025***	0.012***	Income \$300-350	0.009*	-0.015***
	(0.003)	(0.003)		(0.005)	(0.004)
Income \$300-349	0.027***	0.013***	Income \$350-400	0.002	-0.025***
	(0.004)	(0.004)		(0.005)	(0.005)
Income \$350-399	0.020***	0.003	Income \$400-450	-0.002	-0.031***
	(0.005)	(0.004)		(0.006)	(0.006)
Income \$400+	0.019***	-0.000	Income \$450+	0.019***	-0.013*
	(0.006)	(0.006)		(0.007)	(0.007)

Table continues on next page.

Hazard of Stopping—Binned Income & Hours (SF) (continued)

VARIABLES	(1)	(2)	VARIABLES	(3)	(4)
Hour 3-5	0.022***	0.026***	Hour 2	0.018***	0.024***
	(0.002)	(0.002)		(0.001)	(0.001)
Hour 6	0.040***	0.047***	Hour 3	0.033***	0.041***
	(0.003)	(0.003)		(0.002)	(0.002)
Hour 7	0.067***	0.075***	Hour 4	0.044***	0.055***
	(0.003)	(0.003)		(0.003)	(0.003)
Hour 8	0.132***	0.140***	Hour 5	0.050***	0.063***
	(0.005)	(0.005)		(0.003)	(0.003)
Hour 9	0.236***	0.244***	Hour 6	0.058***	0.073***
	(0.008)	(0.008)		(0.004)	(0.003)
Hour 10	0.329***	0.335***	Hour 7	0.072***	0.088***
	(0.013)	(0.013)		(0.004)	(0.004)
Hour 11	0.287***	0.293***	Hour 8	0.098***	0.116***
	(0.016)	(0.016)		(0.005)	(0.004)
Hour 12	0.219***	0.228***	Hour 9	0.163***	0.181***
	(0.012)	(0.012)		(0.006)	(0.006)
Hour 13+	0.271***	0.279***	Hour 10+	0.291***	0.309***
	(0.012)	(0.011)		(0.008)	(0.008)
Warm	0.001	0.001	Warm	0.001	0.001
	(0.000)	(0.000)		(0.000)	(0.000)
Cold	0.001**	0.001*	Cold	0.001*	0.001
	(0.001)	(0.001)		(0.001)	(0.001)
Rain	0.001	0.001	Rain	0.001	0.001
	(0.001)	(0.001)		(0.001)	(0.001)
Observations	1,549,629	1,505,446		1,549,629	1,505,446
R-squared	0.152	0.159		0.151	0.158
Hour of Day Effects	Yes	Yes		Yes	Yes
Day of Week Effects	Yes	Yes		Yes	Yes
Month & Year FEs	Yes	Yes		Yes	Yes
Driver Fixed Effects	Yes	Yes		Yes	Yes
Destination Effects	No	Yes		No	Yes

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

Note: Robust standard errors clustered by driver in parentheses. Columns 1 and 2 correspond to income and hour bins used by Farber (2015).

Table A.7. Hazard of Stopping—Binned Income & Hours (NYC)

	(1)	(2)		(3)	(4)
VARIABLES			VARIABLES		
Income $100-149	0.008***	0.004***	Income $50-100	-0.004***	-0.009***
	(0.000)	(0.000)		(0.000)	(0.000)
Income $150-199	0.013***	0.007***	Income $100-150	-0.004***	-0.013***
	(0.000)	(0.000)		(0.000)	(0.000)
Income $200-224	0.018***	0.010***	Income $150-200	-0.003***	-0.014***
	(0.000)	(0.000)		(0.000)	(0.000)
Income $225-249	0.022***	0.013***	Income $200-250	0.003***	-0.011***
	(0.000)	(0.000)		(0.000)	(0.000)
Income $250-274	0.027***	0.016***	Income $250-300	0.012***	-0.005***
	(0.000)	(0.000)		(0.000)	(0.000)
Income $275-299	0.030***	0.018***	Income $300-350	0.016***	-0.004***
	(0.000)	(0.000)		(0.001)	(0.001)
Income $300-349	0.029***	0.014***	Income $350-400	0.019***	-0.003***
	(0.001)	(0.001)		(0.001)	(0.001)
Income $350-399	0.022***	0.005***	Income $400-450	0.041***	0.016***
	(0.001)	(0.001)		(0.001)	(0.001)
Income $400+	0.058***	0.036***	Income $450+	0.112***	0.083***
	(0.001)	(0.001)		(0.001)	(0.001)

Table continues on next page.

Hazard of Stopping—Binned Income & Hours (NYC) (continued)

VARIABLES	(1)	(2)	VARIABLES	(3)	(4)
Hour 3-5	0.018***	0.019***	Hour 2	0.020***	0.021***
	(0.000)	(0.000)		(0.000)	(0.000)
Hour 6	0.032***	0.034***	Hour 3	0.033***	0.035***
	(0.000)	(0.000)		(0.000)	(0.000)
Hour 7	0.054***	0.057***	Hour 4	0.042***	0.046***
	(0.000)	(0.000)		(0.000)	(0.000)
Hour 8	0.091***	0.094***	Hour 5	0.050***	0.054***
	(0.001)	(0.001)		(0.000)	(0.000)
Hour 9	0.139***	0.144***	Hour 6	0.056***	0.061***
	(0.001)	(0.001)		(0.000)	(0.000)
Hour 10	0.197***	0.202***	Hour 7	0.067***	0.073***
	(0.001)	(0.001)		(0.001)	(0.001)
Hour 11	0.238***	0.242***	Hour 8	0.089***	0.096***
	(0.002)	(0.002)		(0.001)	(0.001)
Hour 12	0.183***	0.190***	Hour 9	0.125***	0.133***
	(0.002)	(0.002)		(0.001)	(0.001)
Hour 13+	0.175***	0.184***	Hour 10+	0.198***	0.207***
	(0.001)	(0.001)		(0.001)	(0.001)
Warm	0.001***	0.001***	Warm	0.001***	0.001***
	(0.000)	(0.000)		(0.000)	(0.000)
Cold	0.000***	0.001***	Cold	0.000***	0.001***
	(0.000)	(0.000)		(0.000)	(0.000)
Rain	0.000***	0.001***	Rain	0.000***	0.001***
	(0.000)	(0.000)		(0.000)	(0.000)
Observations	153,562,159	145,533,591		153,562,159	145,533,591
R-squared	0.143	0.160		0.142	0.159
Hour of Day Effects	Yes	Yes		Yes	Yes
Day of Week Effects	Yes	Yes		Yes	Yes
Month & Year FEs	Yes	Yes		Yes	Yes
Driver Fixed Effects	Yes	Yes		Yes	Yes
Destination Effects	No	Yes		No	Yes

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

Note: Robust standard errors clustered by driver in parentheses. Columns 1 and 2 correspond to income and hour bins used by Farber (2015).

Table A.8. Expected Hourly Earnings (SF)

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES						
Rain	0.643***	0.582***	0.092	0.012	0.018	-0.005
	(0.109)	(0.107)	(0.097)	(0.084)	(0.084)	(0.081)
Cold	0.205*	0.181	-0.574***	-0.267***	-0.235**	-0.185*
	(0.124)	(0.123)	(0.115)	(0.100)	(0.099)	(0.097)
Warm	-1.056***	-0.936***	-0.343***	-0.117	-0.076	-0.091
	(0.148)	(0.144)	(0.130)	(0.111)	(0.110)	(0.107)
L1.Avg. Wage				0.521***	0.450***	0.399***
				(0.012)	(0.014)	(0.014)
Constant	29.061***	30.780***	27.127***	8.718***	6.769***	2.215***
	(0.357)	(0.372)	(0.395)	(0.567)	(0.699)	(0.845)
Observations	9,277	9,277	9,277	9,273	9,245	9,181
R-squared	0.417	0.445	0.558	0.678	0.685	0.705
Hour of Day Effects	Yes	Yes	Yes	Yes	Yes	Yes
Day of Week Effects	No	Yes	Yes	Yes	Yes	Yes
Month & Year FEs	No	No	Yes	Yes	Yes	Yes
Number of Lags	0	0	0	1	8	24
RMSE	4.738	4.626	4.129	3.522	3.488	3.388

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

Note: Robust standard errors in parentheses.

Table A.9. Expected Hourly Earnings (NYC)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Rain	-0.511***	-0.444***	-0.051	-0.015	0.004	-0.060
	(0.153)	(0.152)	(0.074)	(0.047)	(0.047)	(0.047)
Cold	2.110***	2.180***	-0.261**	-0.079	-0.087	-0.066
	(0.146)	(0.144)	(0.108)	(0.076)	(0.074)	(0.070)
Warm	2.256***	2.329***	0.111	0.085	0.086	0.091
	(0.136)	(0.134)	(0.102)	(0.071)	(0.069)	(0.066)
L1.Avg. Wage				0.692***	0.644***	0.608***
				(0.015)	(0.026)	(0.025)
Constant	21.594***	22.535***	24.731***	23.108***	24.354***	20.993***
	(0.374)	(0.393)	(0.289)	(0.301)	(0.782)	(1.035)
Observations	8,759	8,759	8,759	8,758	8,751	8,735
R-squared	0.265	0.278	0.789	0.894	0.900	0.908
Hour of Day Effects	Yes	Yes	Yes	Yes	Yes	Yes
Day of Week Effects	No	Yes	Yes	Yes	Yes	Yes
Month & Year FEs	No	No	Yes	Yes	Yes	Yes
Number of Lags	0	0	0	1	8	24
RMSE	5.459	5.414	2.928	2.074	2.018	1.940

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

Note: Robust standard errors in parentheses.

Table A.10. Coefficients for Estimated Expected Shift Wage

VARIABLES	(1)	(2)
PM Shift	1.192***	1.906***
	(0.081)	(0.008)
Monday	-1.814***	-0.902***
	(0.107)	(0.008)
Tuesday	-3.285***	-0.508***
	(0.108)	(0.008)
Wednesday	-2.542***	-0.208***
	(0.108)	(0.008)
Thursday	-1.301***	0.627***
	(0.108)	(0.008)
Friday	0.255**	0.868***
	(0.109)	(0.008)
Saturday	-0.581***	0.155***
	(0.107)	(0.008)
Constant	28.894***	30.213***
	(0.153)	(0.009)
Observations	84,365	7,107,781
R-squared	0.362	0.347
City	SF	NYC
Month & Year Fixed Effects	Yes	Yes
RMSE	7.819	5.250

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

Note: Standard errors in parentheses.

Table A.11. Estimated Labor Supply Parameters from Estimated Driver Elasticity: Observed Average Wage (SF)

VARIABLES	(1)	(2)	(3)
k	0.000**	-0.000**	-0.000*
	(0.000)	(0.000)	(0.000)
PM Shift			-0.843***
			(0.103)
Constant		8.207***	8.411***
		(0.046)	(0.099)
Observations	79,052	79,052	79,052
R-squared	0.004	0.004	0.215
RMSE	8.553	2.506	2.236
Additional Fixed Effects	No	No	Yes

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

Note: Robust standard errors clustered by driver in parentheses.

Table A.12. Estimated Labor Supply Parameters from Estimated Driver Elasticity: Observed Average Wage (NYC)

VARIABLES	(1)	(2)	(3)
k	0.071***	-0.006***	-0.015***
	(0.002)	(0.000)	(0.001)
PM Shift			-1.409***
			(0.018)
Constant		8.816***	8.951***
		(0.008)	(0.015)
Observations	6,568,159	6,568,159	6,568,159
R-squared	0.052	0.004	0.347
RMSE	8.892	2.542	2.063
Additional Fixed Effects	No	No	Yes

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

Note: Robust standard errors clustered by driver in parentheses.

A.2 Chapter 1 Data Cleaning Procedures

This appendix narrates the cleaning procedures used in Chapter 1.

San Francisco data were originally obtained from the SFMTA broken up by month. I first merged the data into a file referred to as "SF_data_original.dta". This file is a single file containing all of the raw data and the origin file for cleaning purposes.

I began by constructing shifts for New York City and San Francisco separately. Shifts are partitioned for both datasets by a 6 hour or more break between trips. Trips which are recorded as having earned a negative amount of money (3,629 trips in NYC; 68 trips in SF) are changed to earning $0 so as to not affect calculations of shift cumulatives.

In the data there are a number of trips within trips (based on the recorded start and end times). It is unclear whether or not these are recording errors or cabs serving multiple people (ie. splitting fares between individuals with different destinations). My computation of cumulative time worked in shift is unaffected by trips within trips.[1] Time worked is measured as the time between the beginning of the first trip of a shift and the time a shift ended. Additionally, while the first trip of shift is computed by sorting trips by start date and time, the final trip of the shift is computed by sorting trips by end date and time.

With all of the above changes made, I calculate a running total for hours worked and income for each shift. I then eliminate shifts which meet any of the following criteria:

1. Shifts with an average wage greater than the 99th percentile of all shifts or less than or equal to 0.
2. Shifts which last longer than the 99th percentile of all shifts or less than or equal to 0 hours.
3. Shifts which earn more revenue than the 99th percentile of all shifts or less than or equal to $0.

Additionally, for San Francisco, I remove any shift in which the cab was identified as a training vehicle.

[1]All results presented in the paper are robust to inclusion or exclusion of shifts with trips within trips. For shifts without trips within trips, using the cumulative trip and lag minutes within the shifts produces an identical calculated result to the methodology implemented.

A.3 Chapter 2 Data

Table A.13. Major occupation groups and frequencies in combined Glassdoor data

Group	Number	Percent
academic counseler	14,812	0
account executive	28,837	1
accounting	48,211	1
actuarial consultant	5,061	0
administrative	35,537	1
analyst	192,474	5
analytics	23,958	1
appointee	27,841	1
auditor	27,082	1
banker	5,520	0
beauty	8,193	0
bioinformatics scientist	5,527	0
branch manager	199,533	5
business analyst	82,843	2
business coordinator	5,675	0
business development	16,890	0
business operations	8,486	0
c suite	5,295	0
cad designer	16,599	0
caregiver	5,190	0

Continued on next page

Group	Number	Percent
chef	2,887	0
civil engineer	13,804	0
claims	10,268	0
client development manager	11,563	0
client services	8,218	0
clinical dietitian	1,191	0
clinical research	9,750	0
collections representative	5,827	0
communications associate	4,880	0
community manager	5,210	0
compliance consultant	3,834	0
computer programmer	73,375	2
computer sales	8,916	0
construction	5,150	0
contract manager	2,170	0
corporate account manager	38,396	1
corporate attorney	9,339	0
corporate controller	3,106	0
customer service	79,079	2
data specialist	3,761	0
database engineer	33,561	1
dealer	922	0
dean	576	0

Continued on next page

Group	Number	Percent
dentist	2,041	0
deputy manager	7,684	0
designer	26,812	1
driver	13,594	0
editor	22,034	1
engineer	121,571	3
environmental specialist	5,879	0
event coordinator	4,093	0
executive secretary	12,663	0
facility administrator	4,380	0
field sales manager	28,845	1
field services	9,881	0
finance specialist	68,426	2
food services	20,128	0
front desk	12,484	0
front end engineer	21,634	1
game artist	7,792	0
geologist	3,214	0
gis specialist	2,737	0
graphic designer	20,707	0
guest relations	4,056	0
hardware engineer	55,643	1
health educator	1,355	0

Continued on next page

Group	Number	Percent
hr specialist	27,624	1
information security specialist	6,812	0
insurance agent	2,572	0
internal medicine resident	10,411	0
interviewer	664	0
inventory specialist	4,717	0
it	42,336	1
lab specialist	7,096	0
legal	14,231	0
logistics associate	6,099	0
logistics manager	9,396	0
loss prevention	4,561	0
maintenance	5,666	0
management consulting	95,320	2
marketing manager	89,198	2
mechanical engineer	50,681	1
medical technician	18,287	0
merchandiser	13,896	0
military	3,050	0
mobile developer	1,768	0
model	697	0
nursing	26,512	1
operations	12,898	0

Continued on next page

Group	Number	Percent
optician	1,209	0
other	19,522	0
patient care technician	13,701	0
personal trainer	3,165	0
pharamacist	15,204	0
pharmacy technician	10,383	0
physician	13,643	0
physician advisor	1,305	0
pilot	1,985	0
police & security officers	10,935	0
procurement	3,393	0
producer	7,086	0
product manager	31,051	1
product support	7,106	0
production associate	3,987	0
professor	45,191	1
program coordinator	22,470	1
program manager	27,251	1
programmer developer	305,766	7
project coordinator	758	0
project manager	88,422	2
psychologist	2,071	0
purchasing specialist	10,646	0

Continued on next page

Group	Number	Percent
quality assurance	51,186	1
real estate broker	3,985	0
recruiter	21,163	1
regulatory affairs manager	1,908	0
research assistant	25,040	1
researcher	79,493	2
retail representative	86,298	2
revenue manager	11,655	0
sales representative	150,091	4
sap developer	11,396	0
scientist	26,918	1
seo strategist	1,846	0
server	11,185	0
sharepoint developer	5,636	0
skilled labor	11,166	0
social media	5,741	0
social worker	10,339	0
software architect	18,268	0
software engineer	396,165	9
solution specialist	8,518	0
stock clerk	14,883	0
store manager	49,966	1
student	38,474	1

Continued on next page

Group	Number	Percent
supply chain specialist	13,246	0
systems administrator	109,585	3
systems technician	18,079	0
tax specialist	14,192	0
teacher	83,059	2
technical advisor	2,702	0
technical consultant	7,986	0
technical coordinator	1,601	0
technical manager	26,336	1
technical sales	9,362	0
technical staff	14,153	0
technical support	22,371	1
technology specialist	19,524	0
teller	21,385	1
therapist	33,395	1
trainer	11,829	0
underwriter	11,154	0
unskilled labor	15,129	0
valuation associate	1,143	0
veterinary	1,881	0
wall street	23,918	1
Total	4,181,092	100

A.4 Chapter 2 Estimation Results

Table A.14. Income Elasticity of Overall Rating Estimates

VARIABLES	(1)	(2)	(3)	(4)
Log Total Real Annual Compensation ($1000s)	0.083***	0.111***	0.093***	0.118***
	(0.010)	(0.009)	(0.007)	(0.006)
Years of Relevant Experience		-0.003***		-0.005***
		(0.000)		(0.000)
Current Job		0.213***		0.170***
		(0.007)		(0.006)
Age		-0.015***		-0.009***
		(0.001)		(0.001)
Age^2		0.000***		0.000***
		(0.000)		(0.000)
HIGHEST EDUCATION				
Bachelors		0.016*		0.011
		(0.009)		(0.009)
High School		-0.009		0.006
		(0.009)		(0.010)
JD		0.027		0.019
		(0.024)		(0.028)
Masters		0.039***		0.004
		(0.014)		(0.009)
MBA		-0.027**		-0.047***
		(0.013)		(0.012)
MD		-0.040		-0.075
		(0.051)		(0.060)
PHD		0.077**		-0.049***
		(0.031)		(0.017)
GENDER				
Male		0.013**		0.009***
		(0.006)		(0.003)
Prefer Not to State		-0.164***		-0.157***
		(0.013)		(0.013)
Unknown		-0.285**		-0.306**
		(0.136)		(0.133)
Constant	0.729***	0.868***		
	(0.044)	(0.052)		
Observations	425,927	167,202	358,773	132,520
R-squared	0.011	0.079	0.241	0.304
Additional Fixed Effects	No	No	Yes	Yes

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

Note: Dependent variable is log Overall Rating. Robust standard errors clustered by major occupation group in parentheses.

Table A.15. Income Elasticity of Career Opportunities Rating Estimates

VARIABLES	(1)	(2)	(3)	(4)
Log Total Real Annual Compensation ($1000s)	0.122***	0.171***	0.095***	0.131***
	(0.009)	(0.007)	(0.007)	(0.006)
Years of Relevant Experience		-0.005***		-0.006***
		(0.000)		(0.000)
Current Job		0.195***		0.146***
		(0.006)		(0.005)
Age		-0.016***		-0.012***
		(0.001)		(0.001)
Age^2		0.000***		0.000***
		(0.000)		(0.000)
HIGHEST EDUCATION				
Bachelors		-0.010		-0.015**
		(0.007)		(0.007)
High School		0.001		0.009
		(0.009)		(0.009)
JD		-0.041		0.008
		(0.031)		(0.026)
Masters		-0.011		-0.038***
		(0.014)		(0.009)
MBA		-0.071***		-0.089***
		(0.013)		(0.010)
MD		-0.086*		-0.127**
		(0.050)		(0.056)
PHD		0.010		-0.079***
		(0.031)		(0.017)
GENDER				
Male		0.004		0.001
		(0.005)		(0.004)
Prefer Not to State		-0.142***		-0.131***
		(0.012)		(0.013)
Unknown		-0.151		-0.242
		(0.111)		(0.174)
Constant	0.497***	0.624***		
	(0.039)	(0.043)		
Observations	409,572	161,937	344,743	128,174
R-squared	0.021	0.080	0.226	0.280
Additional Fixed Effects	No	No	Yes	Yes

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

Note: Dependent variable is log Career Opportunities Rating. Robust standard errors clustered by major occupation group in parentheses.

Table A.16. Income Elasticity of Compensation & Benefits Rating Estimates

VARIABLES	(1)	(2)	(3)	(4)
Log Total Real Annual Compensation ($1000s)	0.211***	0.242***	0.209***	0.232***
	(0.007)	(0.006)	(0.007)	(0.010)
Years of Relevant Experience		-0.002***		-0.002***
		(0.000)		(0.000)
Current Job		0.097***		0.060***
		(0.005)		(0.004)
Age		-0.010***		-0.009***
		(0.001)		(0.001)
Age^2		0.000***		0.000***
		(0.000)		(0.000)
HIGHEST EDUCATION				
Bachelors		-0.042***		-0.035***
		(0.007)		(0.007)
High School		-0.012*		-0.001
		(0.007)		(0.008)
JD		-0.101***		-0.058***
		(0.030)		(0.021)
Masters		-0.066***		-0.067***
		(0.012)		(0.009)
MBA		-0.097***		-0.095***
		(0.011)		(0.011)
MD		-0.121**		-0.141***
		(0.047)		(0.051)
PHD		-0.049*		-0.085***
		(0.027)		(0.015)
GENDER				
Male		-0.010*		-0.013***
		(0.005)		(0.003)
Prefer Not to State		-0.079***		-0.074***
		(0.012)		(0.011)
Unknown		0.005		-0.125
		(0.119)		(0.111)
Constant	0.214***	0.346***		
	(0.030)	(0.041)		
Observations	410,037	162,105	345,094	128,304
R-squared	0.074	0.094	0.288	0.322
Additional Fixed Effects	No	No	Yes	Yes

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

Note: Dependent variable is log Compensation & Benefits Rating. Robust standard errors clustered by major occupation group in parentheses.

Table A.17. Income Elasticity of Senior Leadership Rating Estimates

VARIABLES	(1)	(2)	(3)	(4)
Log Total Real Annual Compensation ($1000s)	0.070***	0.108***	0.065***	0.096***
	(0.010)	(0.009)	(0.009)	(0.007)
Years of Relevant Experience		-0.006***		-0.006***
		(0.001)		(0.001)
Current Job		0.239***		0.203***
		(0.009)		(0.007)
Age		-0.017***		-0.011***
		(0.001)		(0.001)
Age^2		0.000***		0.000***
		(0.000)		(0.000)
HIGHEST EDUCATION				
Bachelors		0.018**		0.001
		(0.009)		(0.008)
High School		-0.022**		-0.007
		(0.010)		(0.010)
JD		0.076***		0.063**
		(0.025)		(0.026)
Masters		0.041***		0.004
		(0.014)		(0.010)
MBA		0.017		-0.013
		(0.013)		(0.013)
MD		-0.054		-0.113*
		(0.058)		(0.062)
PHD		0.071**		-0.041**
		(0.031)		(0.018)
GENDER				
Male		0.024***		0.023***
		(0.007)		(0.004)
Prefer Not to State		-0.104***		-0.101***
		(0.016)		(0.016)
Unknown		-0.204		-0.202
		(0.142)		(0.203)
Constant	0.620***	0.754***		
	(0.045)	(0.058)		
Observations	406,325	160,863	341,859	127,233
R-squared	0.006	0.076	0.218	0.280
Additional Fixed Effects	No	No	Yes	Yes

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

Note: Dependent variable is log Senior Leadership Rating. Robust standard errors clustered by major occupation group in parentheses.

Table A.18. Income Elasticity of Work-Life Balance Rating Estimates.

VARIABLES	(1)	(2)	(3)	(4)
Log Total Real Annual Compensation ($1000s)	0.091***	0.093***	0.073***	0.088***
	(0.013)	(0.013)	(0.008)	(0.006)
Years of Relevant Experience		-0.002***		-0.003***
		(0.001)		(0.000)
Current Job		0.164***		0.138***
		(0.007)		(0.006)
Age		-0.013***		-0.009***
		(0.001)		(0.001)
Age^2		0.000***		0.000***
		(0.000)		(0.000)
HIGHEST EDUCATION				
Bachelors		0.003		-0.011
		(0.010)		(0.010)
High School		-0.053***		-0.020*
		(0.009)		(0.010)
JD		0.114***		0.090**
		(0.033)		(0.036)
Masters		0.055***		-0.005
		(0.017)		(0.011)
MBA		0.077***		0.034**
		(0.022)		(0.015)
MD		-0.057		-0.113**
		(0.053)		(0.053)
PHD		0.097***		-0.037*
		(0.028)		(0.019)
GENDER				
Male		0.037***		0.040***
		(0.007)		(0.003)
Prefer Not to State		-0.028**		-0.039**
		(0.014)		(0.016)
Unknown		-0.089		-0.091
		(0.090)		(0.147)
Constant	0.695***	0.884***		
	(0.054)	(0.060)		
Observations	410,071	162,144	345,135	128,329
R-squared	0.012	0.052	0.248	0.304
Additional Fixed Effects	No	No	Yes	Yes

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

Note: Dependent variable is log Work-Life Balance Rating. Robust standard errors clustered by major occupation group in parentheses.

Table A.19. Income Elasticity of Cultural Values Rating Estimates

VARIABLES	(1)	(2)	(3)	(4)
Log Total Real Annual Compensation ($1000s)	0.074***	0.100***	0.067***	0.081***
	(0.012)	(0.010)	(0.007)	(0.006)
Years of Relevant Experience		-0.003***		-0.004***
		(0.000)		(0.000)
Current Job		0.224***		0.182***
		(0.008)		(0.007)
Age		-0.017***		-0.010***
		(0.001)		(0.001)
Age^2		0.000***		0.000***
		(0.000)		(0.000)
HIGHEST EDUCATION				
Bachelors		0.002		-0.006
		(0.008)		(0.010)
High School		-0.043***		-0.019*
		(0.010)		(0.011)
JD		0.037		0.057
		(0.039)		(0.044)
Masters		0.023		-0.016
		(0.014)		(0.011)
MBA		-0.016		-0.036**
		(0.015)		(0.017)
MD		-0.081		-0.065
		(0.071)		(0.087)
PHD		0.068**		-0.056**
		(0.030)		(0.022)
GENDER				
Male		0.017***		0.018***
		(0.006)		(0.004)
Prefer Not to State		-0.181***		-0.170***
		(0.024)		(0.032)
Unknown		-0.353**		-0.265
		(0.137)		(0.202)
Constant	0.752***	0.950***		
	(0.050)	(0.054)		
Observations	358,861	137,815	298,238	106,783
R-squared	0.007	0.069	0.252	0.315
Additional Fixed Effects	No	No	Yes	Yes

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

Note: Dependent variable is log Cultural Values Rating. Robust standard errors clustered by major occupation group in parentheses.

Table A.20. Linear Probability of Overall Rating 3, 4 or 5

VARIABLES	(1)	(2)	(3)	(4)
Total Real Annual Compensation ($1000s)	0.001***	1.001***	0.001***	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Residual Compensation ($1000s)	0.001***	1.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Residual Compensation within Employer ($1000s)	-0.001**	0.999***	-0.001***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.677***			
	(0.020)			
Observations	227,009	224,940	90,880	88,225
R-squared	0.008	0.034	0.081	0.402
Additional Controls	No	Yes	Yes	Yes
Demographic Controls	No	No	Yes	Yes
Career Opportunities	No	No	No	Yes

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

Note: Robust standard errors clustered by major occupation group in parentheses. I exclude residuals in the top and bottom 1% of their respective distributions. Additional controls include sector, state, and major occupation group. Demographic controls include years of relevant experience, an indicator for if the job is the respondant's current job, level of education, age, age-squared, and reported gender. Career opportunities denotes inclusion of dummy variables for individual rating of "Career Opportunities" on a 1-5 scale.

A.5 Chapter 3 Subject Demographics

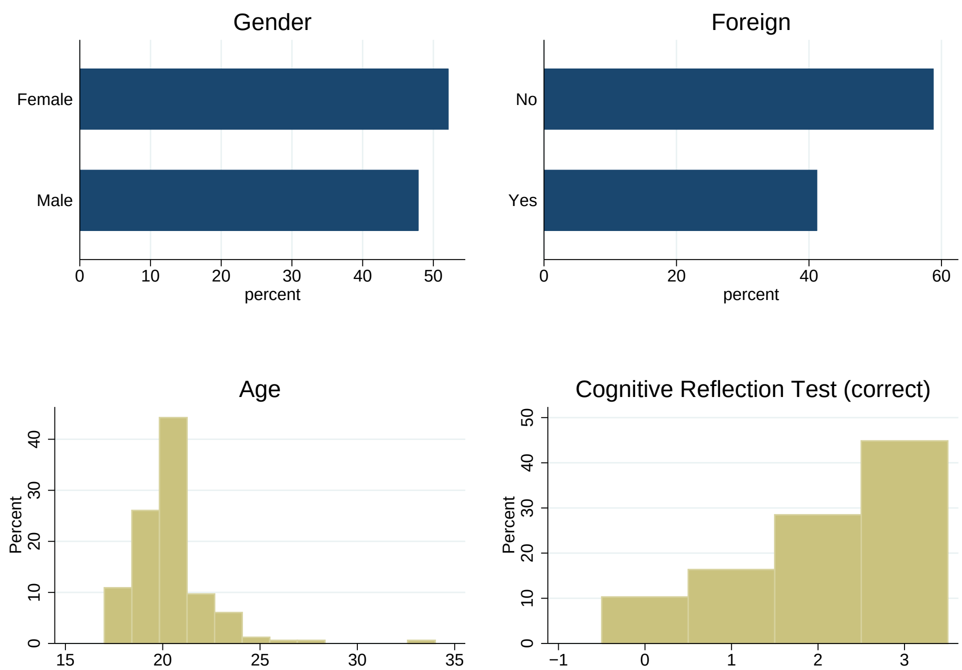

Figure A.1. Subject demographics

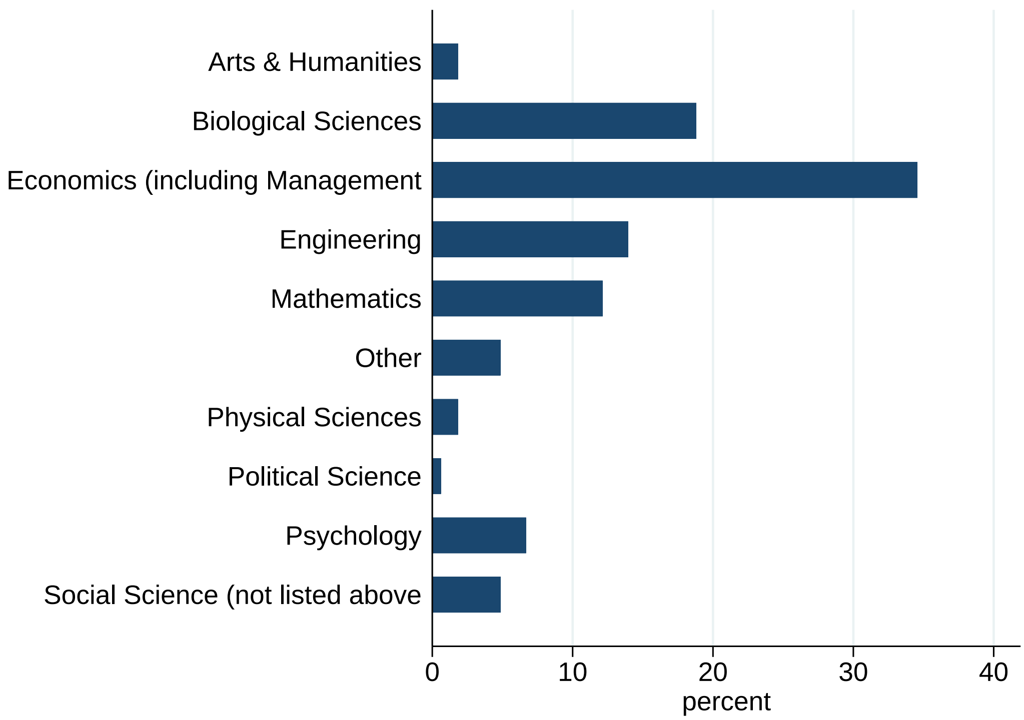

Figure A.2. Distribution of major fields of study for subjects

A.6 Chapter 3 Robustness

Table A.21. OLS estimates for time spent working on tasks with ρ estimated by dropping one budget to make risk choices GARP consistent

VARIABLES	(1) t	(2)	(3)
Low Variance Wage	2.56	3.99	3.28
	(3.92)	(4.25)	(4.32)
High Variance Wage	-4.00	-3.06	-4.12
	(4.18)	(4.39)	(4.55)
Timer	10.25***	10.28***	9.91**
	(3.43)	(3.55)	(3.86)
ρ	-22.66**	-24.49**	-23.59**
	(9.40)	(10.00)	(10.60)
Low Variance Wage $\times$ ρ	17.75	17.86	19.44
	(12.18)	(11.76)	(13.20)
High Variance Wage $\times$ ρ	10.12	18.34	16.48
	(12.75)	(13.57)	(13.97)
Timer $\times$ ρ	23.61**	24.94**	24.48**
	(10.61)	(10.96)	(11.34)
Low Variance Wage $\times$ Timer $\times$ ρ	-13.97	-13.44	-15.50
	(13.20)	(12.62)	(13.63)
High Variance Wage $\times$ Timer $\times$ ρ	-7.58	-15.20	-13.45
	(14.33)	(15.22)	(15.43)
Observations	153	153	153
R-squared	0.12	0.17	0.19
Demographic Controls	No	Yes	Yes
Day of Week Controls	No	No	Yes
Timer + Risk Effect = 0 (p-value)	0.226	0.397	0.374

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

Note: Robust standard errors in parentheses. Demographic controls include gender, age, major, and an indicator for if the subject went to high school in the United States. Timer + Risk Effect = 0 displays a p-value for an F-test of the null hypothesis that ρ is uncorrelated with the outcome for each treatment with the inclusion of a timer.

Table A.22. Tobit estimates for time spent working on tasks with ρ estimated by dropping one budget to make risk choices GARP consistent

VARIABLES	(1)	(2)	(3)
Low Variance Wage	2.24	4.09	3.42
	(6.26)	(6.38)	(6.40)
High Variance Wage	-6.19	-5.01	-6.51
	(6.26)	(6.39)	(6.53)
Timer	15.46***	15.22***	13.67***
	(5.05)	(4.93)	(5.22)
ρ	-35.73**	-38.93**	-37.78**
	(15.07)	(15.68)	(15.70)
Low Variance Wage × ρ	26.34	26.38	27.90
	(18.43)	(17.18)	(17.98)
High Variance Wage × ρ	22.09	37.08*	34.94*
	(18.50)	(20.04)	(19.91)
Timer × ρ	36.92**	39.49**	40.06**
	(17.47)	(17.65)	(17.51)
Low Variance Wage × Timer × ρ	-16.74	-17.25	-20.79
	(21.70)	(20.21)	(20.40)
High Variance Wage × Timer × ρ	-16.45	-31.42	-30.45
	(22.08)	(23.87)	(23.64)
Observations	153	153	153
Demographic Controls	No	Yes	Yes
Day of Week Controls	No	No	Yes
Timer + Risk Effect = 0 (p-value)	0.413	0.551	0.505

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

Note: Robust standard errors in parentheses. Demographic controls include gender, age, major, and an indicator for if the subject went to high school in the United States. Timer + Risk Effect = 0 displays a p-value for an F-test of the null hypothesis that ρ is uncorrelated with the outcome for each treatment with the inclusion of a timer.

Table A.23. Logit estimates for probability of working on task for 60 minutes with ρ estimated by dropping one budget to make risk choices GARP consistent

VARIABLES	(1)	(2)	(3)
Low Variance Wage	-0.05	-0.06	-0.19
	(0.46)	(0.49)	(0.51)
High Variance Wage	-0.23	-0.31	-0.51
	(0.48)	(0.49)	(0.54)
Timer	1.16***	1.08**	0.85*
	(0.44)	(0.46)	(0.46)
ρ	-2.95**	-2.69**	-2.36*
	(1.45)	(1.37)	(1.21)
Low Variance Wage × ρ	1.67	-1.15	-2.89
	(1.94)	(3.03)	(3.47)
High Variance Wage × ρ	2.62	3.22*	2.75
	(1.72)	(1.81)	(1.74)
Timer × ρ	2.97*	2.62*	2.52*
	(1.57)	(1.48)	(1.35)
Low Variance Wage × Timer × ρ	-1.20	1.54	2.97
	(2.09)	(3.10)	(3.53)
High Variance Wage × Timer × ρ	-2.23	-2.86	-2.66
	(1.90)	(1.96)	(1.92)
Observations	153	145	141
Demographic Controls	No	Yes	Yes
Day of Week Controls	No	No	Yes
Timer + Risk Effect = 0 (p-value)	0.553	0.844	0.902

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

Note: Robust standard errors in parentheses. Demographic controls include gender, age, major, and an indicator for if the subject went to high school in the United States.

Table A.24. OLS estimates for total tasks completed with ρ estimated by dropping one budget to make risk choices GARP consistent

VARIABLES	(1)	(2)	(3)
Low Variance Wage	3.14	4.33	4.24
	(4.64)	(4.88)	(4.90)
High Variance Wage	-4.95	-3.71	-4.69
	(4.84)	(4.92)	(5.11)
Timer	8.83**	9.02**	10.03**
	(3.98)	(4.14)	(4.35)
ρ	-27.10***	-29.78***	-29.95***
	(8.00)	(8.98)	(9.33)
Low Variance Wage × ρ	19.42	21.60*	26.65*
	(12.49)	(12.85)	(13.52)
High Variance Wage × ρ	17.25	24.36*	21.88
	(12.09)	(13.06)	(13.26)
Timer × ρ	24.73**	26.17**	26.30**
	(10.01)	(10.76)	(10.88)
Low Variance Wage × Timer × ρ	-11.18	-11.71	-16.26
	(14.06)	(14.06)	(14.32)
High Variance Wage × Timer × ρ	-12.10	-18.19	-14.94
	(14.21)	(14.98)	(15.22)
Observations	153	153	153
R-squared	0.11	0.17	0.20
Demographic Controls	No	Yes	Yes
Day of Week Controls	No	No	Yes
Timer + Risk Effect = 0 (p-value)	0.350	0.446	0.290

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

Note: Robust standard errors in parentheses. Demographic controls include gender, age, major, and an indicator for if the subject went to high school in the United States. Timer + Risk Effect = 0 displays a p-value for an F-test of the null hypothesis that ρ is uncorrelated with the outcome for each treatment with the inclusion of a timer.

Table A.25. OLS estimates for various measures of the distribution of task completion time with ρ estimated by dropping one budget to make risk choices GARP consistent

VARIABLES	(1) Avg. Task Time	(2) Median Task Time	(3) Max Task Time	(4) Min Task Time	(5) Task Time Std. Dev.	(6) Task Time 25th Pctle	(7) Task Time 75th Pctle
Low Variance Wage	-0.13	-0.19	0.54	-0.11	0.05	-0.09	-0.26
	(0.17)	(0.17)	(0.44)	(0.12)	(0.11)	(0.12)	(0.21)
High Variance Wage	0.15	0.11	0.36	0.17	0.11	0.17	0.11
	(0.21)	(0.21)	(0.48)	(0.19)	(0.14)	(0.19)	(0.25)
Timer	-0.45	-0.49*	0.27	-0.55*	0.04	-0.53*	-0.51*
	(0.29)	(0.29)	(0.53)	(0.29)	(0.13)	(0.28)	(0.29)
ρ	0.81	0.93	-0.73	0.96**	0.00	0.87*	0.92
	(0.57)	(0.58)	(0.84)	(0.48)	(0.22)	(0.52)	(0.62)
Low Variance Wage × ρ	-1.52*	-1.78*	0.53	-1.83**	0.05	-1.68**	-1.68*
	(0.92)	(0.90)	(1.76)	(0.82)	(0.44)	(0.85)	(0.96)
High Variance Wage × ρ	0.48	0.36	1.88	0.23	0.32	0.24	0.49
	(1.11)	(1.13)	(1.41)	(1.12)	(0.33)	(1.13)	(1.11)
Timer × ρ	-0.59	-0.76	1.11	-0.80*	0.10	-0.70	-0.67
	(0.58)	(0.61)	(0.88)	(0.45)	(0.24)	(0.49)	(0.67)
Low Variance Wage × Timer × ρ	1.03	1.38	-1.50	1.48**	-0.34	1.31*	1.16
	(0.85)	(0.85)	(1.69)	(0.72)	(0.42)	(0.75)	(0.93)
High Variance Wage × Timer × ρ	-0.86	-0.74	-1.67	-0.58	-0.34	-0.61	-0.94
	(1.22)	(1.25)	(1.64)	(1.24)	(0.41)	(1.25)	(1.24)
Constant	-0.61	-0.50	-2.76	-0.33	-0.63	-0.43	-0.25
	(1.32)	(1.29)	(2.96)	(1.20)	(0.80)	(1.19)	(1.44)
Observations	151	151	151	151	146	151	151
R-squared	0.25	0.26	0.16	0.28	0.13	0.28	0.25
Demographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day of Week Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Timer + Risk Effect = 0 (p-value)	0.314	0.221	0.271	0.244	0.369	0.203	0.352

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

Note: Robust standard errors in parentheses.

Bibliography

Abeler, B. J., Falk, A., Goette, L., and Huffman, D. (2011). Reference Points and Effort Provision. *American Economic Review*, 101(2):470–492.

Allen, E. J., Dechow, P. M., Pope, D. G., and Wu, G. (2014). Reference-Dependent Preferences: Evidence from Marathon Runners.

Andreoni, J. and Harbaugh, W. T. (2015). Unexpected Utility : Testing Expected Utility and Its Alternatives using Convex Risk Budgets.

Barth, E., Bryson, A., Davis, J. C., and Freeman, R. (2016). It's Where You Work: Increases in the Dispersion of Earnings across Establishments and Individuals in the United States. *Journal of Labor Economics*, 34(S2):S67–S97.

Block, M. and Heineke, J. (1973). The Allocation of Effort under Uncertainty: The Case of Risk-Averse Behavior. *Journal of Political Economy*, 81(2):376–385.

Böckerman, P. and Ilmakunnas, P. (2009). Job disamenities, job satisfaction, quit intentions, and actual separations: Putting the pieces together. *Industrial Relations*, 48(1):73–96.

Borjas, G. (1980). The relationship between wages and weekly hours of work: The role of division bias. *Journal of Human Resources*, 15(3):409–423.

Brickman, P. and Campbell, D. T. (1971). Hedonic Relativism and Planning the Good Society. *Adaption-level theory*, pages 287–305.

Brown, G. D., Gardner, J., Oswald, A. J., and Qian, J. (2008). Does Wage Rank Affect Employees' Wellbeing? *Industrial Relations*, 47(3):355–389.

Browning, M. and Crossley, T. F. (2001). The Life-Cycle Model of Consumption and Saving. *Journal of Economic Perspectives*, 15(3):3–22.

Bygren, M. (2004). Pay reference standards and pay satisfaction: What do workers evaluate their pay against? *Social Science Research*, 33(2):206–224.

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

Cappelli, P. and Sherer, P. D. (1988). Satisfaction, Market Wages, & Labor Relations: An Airline Study. *Industrial Relations: A Journal of Economy and Society*, 27(1):56–73.

Chamberlain, A. (2015). Does Company Culture Pay Off? Analyzing Stock Performance of Best Places to Work Companies. Technical Report March.

Chetty, R. (2006). A New Method of Estimating Risk Aversion. *American Economic Review*, 96(5):1821–1834.

Chiu, W. H. and Eeckhoudt, L. (2010). The effects of stochastic wages and non-labor income on labor supply : update and extensions. *Journal of Economics*, (100):69–83.

Chou, Y. K. (2002). Testing Alternative Models of Labour Supply: Evidence From Taxi Drivers in Singapore. *The Singapore Economic Review*, 47(01):17–47.

Clark, A., Oswald, A., and Warr, P. (1996). Is job satisfaction U-shaped in age? *Journal of Occupational and Organizational Pshychology*, 69:57–81.

Clark, A. E. (2001). What really matters in a job? Hedonic measurement using quit data. *Labour Economics*, 8(2):223–242.

Clark, A. E., Frijters, P., and Shields, M. A. (2008). Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles. *Journal of Economic Literature*, 46(461):95–144.

Clark, A. E., Kristensen, N., and Westergard-Nielsen, N. (2009). Job satisfaction and co-workers wages: status or signal? *Economic Journal*, 119(2008):430–447.

Clark, A. E. and Oswald, A. J. (1996). Satisfaction and comparison income. *Journal of Public Economics*, 61(3):359–381.

Correia, S. (2015). REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects.

Crawford, V. P. and 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(August):1912–1932.

Easterlin, R. A. (1974). Does Economic Growth Improve the Human Lot? Some Empirical Evidence. In *Nations and Households in Economic Growth*, pages 89–125. Elsevier.

Ekern, S. (1980). Increasing Nth Degree Risk. *Economics Letters*, 6:329–333.

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

Farber, H. S. (2008). Reference-Dependent Preferences and Labor 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 Labor Supply Lessions from Cab Drivers. *Quarterly Journal of Economics*, pages 1975–2026.

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

Ferrer-i Carbonell, A. (2005). Income and well-being: An empirical analysis of the comparison income effect. *Journal of Public Economics*, 89(5-6):997–1019.

Freeman, R. B. (1978). Job satisfaction as an economic variable. *American Economic Review*, 68(2):135–141.

Friedman, M. and Savage, L. (1948). The Utility Analysis of Choices Involving Risk. *Journal of Political Economy*, 56(4):279–304.

Gao, W. and Smyth, R. (2010). Job satisfaction and relative income in economic transition: Status or signal? The case of urban China. *China Economic Review*, 21(3):442–455.

Gneezy, U., Goette, L., Sprenger, C., and Zimmermann, F. (2016). The Limits of Expectations-Based Reference Dependence. *Journal of the European Economic Association*.

Green, F. (2010). Well-being, job satisfaction and labour mobility. *Labour Economics*, 17(6):897–903.

Grossberg, A. J. (1989). Labor Supply under Real Wage Uncertainty: A New Look at the Intertemporal Subsitution Hypothesis. *Southern Economic Journal*, 55(4):974–986.

Hicks, J. (1963). *The Theory of Wages*. St. Martin's Press, New York.

Jann, B. (2015). COEFPLOT: Stata module to plot regression coefficients and other results.

Jouini, E., Napp, C., and Nocetti, D. (2013). Economic consequences of Nth-degree risk increases and Nth-degree risk attitudes. *Journal of Risk and Uncertainty*, (47):199–224.

Judge, T. A., Thoresen, C. J., Bono, J. E., and Patton, G. K. (2001). The Job Satisfaction-Job Performance Relationship: A Qualitative and Quantitative Review. *Psychological Bulletin*, 127(3):376–407.

Kahneman, D., Alan, K., Schkade, D., Schwarz, N., and Stone, A. (2006). Would You Be Happier If You Were Richer. *Science*, 312:1908–1910.

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

Knight, F. H. (1964). *Risk, uncertainty and profit*. Kelly.

Koszegi, B. and Rabin, M. (2006). A Model of Reference-Dependent Preferences. *Quarterly Journal of Economics*, CXXI(November):1133–1165.

Krueger, A. B. and Schkade, D. a. (2008). The Reliability of Subjective Well-Being Measures. *Journal of Public Economics*, 92(8-9):1833–1845.

Leah-Martin, V. (2016). When to Quit: Narrow Bracketing and Reference Dependence in Taxi Drivers.

Leah-Martin, V. and Andreoni, J. (2016). Adaptive convex risk budgets in oTree.

Lévy-Garboua, L., Montmarquette, C., and Simonnet, V. (2007). Job satisfaction and quits. *Labour Economics*, 14(2):251–268.

Low, H. (2005). Self-Insurance in a Life-Cycle Model of Labor Supply and Savings. *Review of Economic Dynamics*, 8(4):945–975.

Milgrom, P. and Segal, I. (2002). Envelope Theorems for Arbitrary Choice Sets. *Econometrica*, 70(2):583–601.

Parker, S. C., Belghitar, Y., and Barmby, T. (2005). Wage Uncertainty and the Labour Supply of Self-Employed Workers. *The Economic Journal*, 115(502):C190–C207.

Pfeifer, C. and Schneck, S. (2012). Relative wage positions and quit behavior: Evidence from Linked Employer-Employee Data. *Industrial and Labor Relations Review*, 65(1):126–147.

Rabin, M. (2000). Notes and comments. *Econometrica*, 68(5):1281–1292.

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

Sacks, D. W., Stevenson, B., and Wolfers, J. (2010). Subjective Well-Being , Income , Economic Development and Growth. *CESIFO Working Paper*.

Shefrin, H. and Statman, M. (1985). The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence. *The Journal of Finance*, XL(3):777–790.

Shields, M. A. and Ward, M. (2001). Improving nurse retention in the National Health Service in England: The impact of job satisfaction on intentions to quit. *Journal of Health Economics*, 20(5):677–701.

Smart, M. (2016). Preventing Bias in Glassdoor Employer Reviews Author : Morgan Smart Abstract :.

Song, J., Price, D., Guvenen, F., Bloom, N., and von Wachter, T. (2015). Firming Up Inequality. Technical report, National Bureau of Economic Research, Cambridge, MA.

Turkyilmaz, A., Akman, G., Ozkan, C., and Pastuszak, Z. (2011). Empirical study of public sector employee loyalty and satisfaction. *Industrial Management & Data Systems*, 111(5):675–696.

Tversky, A. and Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice-. *Science*, 211(4481):453–458.

Tversky, A. and Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics*, 106(4):1039–1061.

Wanous, J. P., Reichers, A. E., and Hudy, M. J. (1997). Overall Job Satisfaction: How Good Are Single-Item Measures? *Journal of Applied Psychology*, 82(2):247–252.