Department of Economics and Finance

Gordon S. Lang School of Business and Economics

University of Guelph

Discussion Paper 2020-06

Working more for more and working more for less: Labor supply in the gain and loss domains

By:

C. Bram Cadsby
University of Guelph
bcadsby@uoguelph.ca

Fei Song
Ryerson University
fsong@ryerson.ca

Nick Zubanov
University of Konstanz
nick.zubanov@uni-konstanz.de

Department of Economics and Finance
Gordon S. Lang School of Business and Economics | University of Guelph
50 Stone Road East | Guelph, Ontario, Canada | N1G 2W1
www.uoguelph.ca/economics

Improve Life.

Working more for more and working more for less:
Labor supply in the gain and loss domains

C. Bram Cadsby*
Department of Economics and Finance
University of Guelph
Tel: 1-519-8244120 ext.53320
Email: bcadsby@uoguelph.ca

Fei Song
Ted Rogers School of Management
Ryerson University
Tel: 1-416-9795000 ext.7503
Email: fsong@ryerson.ca

Nick Zubanov
Department of Economics
University of Konstanz
Tel: 49-7531 88-4404
Email: nick.zubanov@uni-konstanz.de

*Corresponding Author.

Abstract

We examine the response of labor supply to short-run wage changes with and without a reference wage (RW) that we manipulate experimentally. We find that, in the absence of RW, labor supply increases monotonically with wage. In contrast, when RW is present, people work more both when wages rise and fall relative to RW. These findings suggest a kink in the labor-supply curve, consistent with income targeting by loss-averse individuals. However, the effects of income targeting are sensitive to context: in a treatment where wages could either rise or fall relative to RW, the kink in the labor-supply curve disappears.

Keywords: labor supply, short-run wage changes, reference wage, loss-aversion, experimental.

JEL: D91, J22, J31, M52

Acknowledgement: We thank James Amegashie for alerting us to this issue, and helpful comments by Gary Charness, David Dickinson, Philip Oreopoulos, Bradley Ruffle, Marie Claire Villeval, Irenaeus Wolff, and participants in the Micro seminar at the University of Konstanz.

INTRODUCTION

How wage rate affects the allocation of time between work and leisure is of central importance to many fields in economics, ranging from the economics of personnel management to public finance. Within the neoclassical model of labor supply, when the wage increases, the extra income creates an incentive to consume more leisure and hence work less (the *income effect*), but at the same time, the now higher opportunity cost of leisure induces substitution away from leisure towards more work (the *substitution effect*). With the income and substitution effects working in opposite directions, the response of labor supply to a change in wage depends on which of the two effects prevails. This is an empirical question, and the existing research has yet to converge on the answer to it.

Focusing on short-run changes in wages – a setting frequently occurring in empirical labor supply studies – the neoclassical model of labor supply would predict the substitution effect to prevail (e.g., Fehr and Goette, 2007, Section II). That is, labor supply should increase in response to higher wages, owing to inter-temporal substitution between consumption (funded by labor income) and leisure. Indeed, several field studies conducted in settings where workers had the flexibility to change hours in response to short-run wage fluctuations have confirmed this intuition, using data from diverse work contexts such as construction of the Trans-Alaskan pipeline (Carrington, 1996), vending at sports stadia (Oettinger, 1999), or harvesting lobsters (Stafford, 2015).

However, a seminal work by Camerer et al. (1997) found something quite different: New York City taxi drivers drove fewer hours on days when their hourly earnings were higher – a strong income effect with no evidence of intertemporal substitution between labor and leisure. The authors applied prospect theory (Kahneman and Tversky, 1979) to explain their results through *income targeting*: earning an income below a certain daily target is perceived as a loss, which loss-averse drivers try to avoid by working more hours if hourly earnings are low. The income target being reached more quickly on more lucrative days explains the negative relationship between wages and hours.

Though Camerer et al.'s (1997) study has been highly influential, its message in favor of the target income hypothesis has not been received with unanimous support. Some follow-up field studies found evidence for income targeting (e.g., Chou, 2002 for taxi drivers in Singapore; Chang and Gross, 2014 for fruit pickers in California), others found only weak or no evidence (Farber, 2005, 2008, 2015 using later and more complete NYC taxi driver records), while yet other studies reconciled those conflicting findings by allowing for income as well as work hour targets (Crawford and Meng, 2011 for NYC taxi drivers). It has been argued that the above field results are sensitive to imperfect measurement of labor supply, endogeneity of the wage, and sample selection issues (Oettinger, 1999; Farber, 2005; Stafford, 2015), which could, at least partly, explain their differing conclusions.

Experimental research helps address some of the imperfections of the field setting. In addition to exogenously manipulating the wage and controlling for participation, experiments allow i) more precise measurement of effort vs. hours, two dimensions of labor supply that may react differently to the same wage increase (Dickinson, 1999; Fehr and Goette, 2007), and ii) manipulating the income target exogenously instead of inferring it from the data using an assumed theoretical structure. Methodologically, our study belongs to this latter group. The income target may be manipulated by changing income expectations or the reference wage rate. Abeler et al. (2011), in a study closely related to ours, varied their experiment participants' income expectations by randomizing whether subjects would receive a pre-specified fixed payment or a piece rate based on their actual output. The outcome was randomly determined after the subjects finished the task with a 50% chance of each outcome. The prespecified fixed payment differed by treatment. The authors found that participants in the high fixed payment treatment worked more, and harder, than those in the low fixed payment treatment, and interpreted the observed difference in labor supply as evidence for loss-avoiding behavior consistent with the income targeting hypothesis.

Our approach is alternative, and complementary, to Abeler et al.'s (2011) in that we manipulate the income target by varying the presence and the level of the *reference wage* applied at the initial stage of our experiment. We then measure the effect of the reference wage on the amount of labor supplied at subsequently available wage rates, which we elicit using the strategy method (Brandts and Charness, 2011). Our design enables us to observe within-subject variation in labor supply in response to the wage being at, above or below the reference level, that is, in both the gain and loss domains with respect to the target income. Doing so allows us to generate and test a rich set of predictions from the labor supply model with an income target without requiring additional information such as individual loss aversion. Also, unlike setting income expectations using a lottery with varying payments, as in Abeler et al. (2011), our design is income risk-free, allowing us to disregard the possible effects of risk aversion on labor supply.

THEORETICAL FRAMEWORK AND TESTABLE PREDICTIONS

Consider an individual worker maximizing net income by choosing labor supply h given the reference wage rate $w_0 \geq 0$ per unit of labor and the cost of effort $\frac{\theta h^2}{2}$, where $\theta > 0$ is a scalar that captures the costliness of exerting effort. Maximizing the net income $w_0 \cdot h - \frac{\theta h^2}{2}$

with respect to h yields the optimal labor supply $h(w_0) = \frac{w_0}{\theta}$ and the corresponding gross income, which we refer to as the *target income* $T \equiv \frac{w_0^2}{\theta}$.[1]

Now suppose the wage rate changes from the reference level w_0 to another level w. We allow the worker to have loss-averse preferences, which reflects the abundant empirical evidence (Camerer, 2000; see also Wakker (2010) for a more recent survey). Loss aversion implies that falling below the target income by a given amount causes a greater worsening of utility than rising above the target by the same amount adds to it. Hence, the worker's problem becomes

$$\max_h U = \begin{cases} w \cdot h - T - \frac{\theta h^2}{2}, & w \cdot h \geq T \text{ (target income reached)} \\ \lambda \cdot (w \cdot h - T) - \frac{\theta h^2}{2}, & w \cdot h \leq T \text{ (target income NOT reached)} \end{cases}$$

where $\lambda \geq 1$ is the loss-aversion parameter ($\lambda = 1$ corresponds to loss-neutrality, a theoretically possible but empirically unlikely case). Separating the above problem into two constrained maximization problems, one for $w \cdot h \geq T$ and one for $w \cdot h \leq T$, and solving both simultaneously gives the utility-maximizing labor supply as a function of the going wage w:

(1) $$h(w) = \begin{cases} \frac{\lambda \cdot w}{\theta}, & w \leq \sqrt{\frac{T \cdot \theta}{\lambda}} \\ \frac{T}{w}, & \sqrt{\frac{T \cdot \theta}{\lambda}} \leq w \leq \sqrt{T \cdot \theta} = w_0 \\ \frac{w}{\theta} & w \geq \sqrt{T \cdot \theta} = w_0 \end{cases}$$

INSERT FIGURE 1 ABOUT HERE

Figure 1 plots the above function (thick line), revealing the "kink" in the labor supply curve that produces local non-monotonicity in labor supply as a function of wage for loss-averse workers

[1] We abstract from participation issues by assuming the reservation wage is low enough, which is plausible in our experimental setting, and borne out by the data, since no participant refused to supply labor at any wage rate.

$(\lambda > 1)$[2]. In particular, it draws a sharp contrast between labor supply responses to an increase in the wage above its previously established reference level, $w \geq w_0 = \sqrt{T \cdot \theta}$, which places the worker in the *gain domain*, and a decrease in the wage below its reference level, $w \leq w_0 = \sqrt{T \cdot \theta}$, which situates the worker in the *loss domain*.

In the gain domain, labor supply grows with wage at the same rate $\frac{1}{\theta}$ as without a reference wage or target income (thin line on Figure 1). On the other hand, in the loss domain, for relatively low $\mathrm{w} \leq \sqrt{\frac{\mathrm{T} \cdot \theta}{\lambda}}$, there is an extra incentive to work – to reduce the displeasure from falling below the target income level – as well as the usual substitution effect, so labor grows with wage at a higher rate, $\frac{\lambda}{\theta} > \frac{1}{\theta}$, than in the gain domain. Lastly, in the loss domain for the range of wage levels $\sqrt{\frac{\mathrm{T} \cdot \theta}{\lambda}} \leq \mathrm{w} \leq \sqrt{\mathrm{T} \cdot \theta}$, the optimal labor supply is $h(w) = \frac{T}{w}$, just enough to reach the target income *T*, reflecting the trade-off between the gains from reaching the target and the opportunity costs of doing so. As a result, less/more labor is supplied at a higher/lower wage in this range – the effect found in Camerer et al. (1997) and other studies.

Summarizing, the above theoretical framework generates the following predictions that can be tested experimentally.

Prediction 1: In the absence of a reference wage or target income, labor supply monotonically increases with wage level.

Prediction 2: In the presence of a reference wage and target income, labor supply increases with wage in the gain domain ($\mathrm{w} \geq \mathrm{w}_0 = \sqrt{\mathrm{T} \cdot \theta}$), at the same rate as without a reference wage.

---

[2] There is no kink for the loss-neutral ($\lambda = 1$). Structural estimation of the labor supply equation (Table 4) produces estimates of λ that are significantly larger than 1 and are on par with the findings of several other studies.

Prediction 3: In the presence of a reference wage and target income, labor supply increases (at least locally) as wage falls in the loss domain $\left(w \leq \sqrt{T \cdot \theta} = w_0\right)$.

Prediction 4: Given the wage level, labor supply is higher with than without a reference wage or target income, in the loss domain.[3]

THE EXPERIMENT

We test the above theoretical predictions by experimentally varying the initially paid reference wage for a given task as well the menu of possible subsequent wages.

Task. Our measure of labor supply at a given wage level is individual performance on an *arithmetic task* that involves adding sets of five two-digit numbers (e.g., Niederle and Vesterlund, 2007). This task was selected because outputs are mostly effort-based and little learning can take place in a short period of time (e.g., Else-Quest et al., 2010; Hyde et al., 1990). Moreover, it is unlikely that most people derive intrinsic satisfaction/utility from this task. The numbers for the arithmetic task were randomly generated on a computer, but the experiment itself was conducted using paper and pencil, and without calculators.

Wage levels. There were three piece-rate wage levels used in the experiment: $0.50, $1.00 and $2.00 per correctly solved problem (the sum of five two-digit numbers). In what follows, we refer to these as "low" (L), "medium" (M) and "high" (H) wage levels. However, the experiment participants were not given such labels for the three piece-rate wage levels. Instead, they were merely informed about the specific dollar amount of the relevant wages.

[3] This is similar to Abeler et al. (2011) who predict higher effort under higher target income and find supporting evidence. The difference between our study and theirs is that they keep piece rates constant across the treatments, whereas we vary both the reference piece-rate wage and piece rates.

Treatments. There were seven between-person treatments that varied in the wage schedule offered to the participants. Each treatment involved two stages: a two-minute warmup stage and the main experimental stage, which, importantly, did not have a set time limit. The warmup stage served to familiarize the participants with the task and to help them gauge its difficulty, and was kept short in order to avoid tiring the participants. The wage offered during the warmup stage, the level of which varied by treatment, served as the reference wage. Both task difficulty (parameter $\theta > 0$ in the model) and the reference wage offered in the warmup stage (w_0 in the model) are essential inputs for the decisions the participants would have to make at the beginning of the subsequent experimental stage.

At the beginning of the experimental stage, the participants were informed that there were different wage levels (varying by treatment), each equally likely to be applied to calculate their earnings. They were asked to think carefully and write down the number of problems they would commit themselves to working on under each of the equally likely wage levels. After the participants provided this information, the applicable wage level was determined for them by a random draw. To receive their payment, the participants had to correctly solve the number of problems they had committed to under the applicable wage level regardless of how long it took. Thus, the number of problems each participant committed to under different wage levels using the strategy method enabled us to identify the labor-supply response to wage changes controlling for individual-specific characteristics[4].

[4] Other measures of labor supply are possible, such as the time spent working on the problems or the solution accuracy rate. We feel those measures are not ideal because, in our setting, they rather seem to speak more to aspects of personality or work habits than work effort. Accuracy was checked by our experimental assistants. Problems were returned to participants to be redone if the submitted answers were not correct, so there was no gain from providing wrong solutions. The time spent working is hard to interpret as it may proxy both higher effort and lower effort intensity. Lastly, time and solution accuracy records that we kept on paper are not available at present, because of the access restrictions to the experimental site due to the COVID-19 pandemic.

Table 1 summarizes the details of the seven experimental treatments labelled according to their corresponding wage schedules. The first letter in the label denotes the wage level offered during the warmup stage (L, M or H, or X if no wage was offered), and the following letters denote the combinations of possible, and equally likely, wages in the experimental stage. For example, in treatment L-LM, each participant was offered a low (L) wage of $0.50 per correctly solved problem in the warmup stage and had either a low or medium ($1 per correctly solved problem) wage in the experimental stage determined by a random draw.

INSERT TABLE 1 ABOUT HERE

The seven experimental treatments represent three treatment types: with reference wage: unidirectional and bidirectional, and without reference wage. The focus of our interest was on the four unidirectional treatments. In two of these (treatments L-LM and M-MH), our intention was to elicit how participants would respond if faced with an increase in the wage relative to the lower reference wage offered in the warm-up stage. In two other treatments (treatments M-LM and H-MH), our purpose was to elicit how participants would respond when faced with a reduction in the wage relative to the higher reference wage offered in the warm-up stage. We added two treatments without reference wage to provide a contrast to our focal unidirectional treatments. In these treatments, no wage was offered in the warmup stage; instead, the participants received a $3.00 show-up fee, which was not contingent on the output produced in the warm-up stage. This was intended to neutralize any wealth effects that may have resulted from the earnings in the warm-up stages of the treatments with a reference wage. The labor supply decisions made in these treatments may be compared with the analogous decisions made in the corresponding unidirectional treatments in which a reference wage was present to highlight the impact of the reference wage. Finally, we ran a bidirectional treatment M-LMH.

Like two of the unidirectional treatments, M-LM and M-MH, this treatment offered a medium reference wage in the warm-up stage. However, in the subsequent experimental stage, participants were informed that there was an equal probability of the piece rate being high, medium, or low. Participants were then asked simultaneously to specify their labor-supply commitment in all three eventualities. The purpose was to examine whether people react differently when presented simultaneously with the possibility of gains and the possibility of losses rather than when shown only one or the other.

Participants and Procedures. A total of 247 undergraduate students (average age 20.7 years, standard deviation 2.4 years, 57% female) from a large Canadian university participated in the experiment and were randomly assigned to the seven treatments shown in Table 1. The treatments were randomized over morning and afternoon time slots and over days of the week.

We booked two separate lab spaces for this experiment. One was used as the "decision" room and the other was used as the "workstation" room. In order to avoid peer effects and information spillovers, each participant was scheduled to arrive at the lab one at a time and at least 20 minutes apart from any other participant. Thus, there was always only one participant in the "decision" room at any given time. During that 20-minute period, the participant would sign the consent form, read the instructions, complete the 2-minute warm-up round, read the instructions for the experimental round and make a decision on their committed labor supply. Then, he/she was given a die to throw to determine which wage rate was going to be implemented. At that point, the participant would be escorted to the "workstation" room, where they fulfilled their work commitments. In the work room, participants did see others entering and leaving, but they were not allowed to communicate with them, and no interaction between participants was detected. Another RA, who was the grader and was always stationed in the workstation room, would grade the work. If there were mistakes in the submitted work, they would need to be corrected by the participant. The participant then completed a post-experiment survey, which collected their age, gender and major of study at the University, received their payment privately, and left the premises. Each participant took part in the experiment only once. In this set-up, participants had total control over their choice between leisure, which was leaving the experiment and going home, and labor supply as they could choose the exact amount of labor supply.

RESULTS

Figure 2 and Table 2 summarize our key results[5]. While average labor supply in the warmup stage varies little by treatment, its variation in the experimental stage is stronger and is mostly consistent with our model predictions. We outline the results below.

INSERT FIGURE 2 AND TABLE 2 ABOUT HERE

Result 1: People work more for a higher wage in the absence of reference wage.

In the reference wage-free treatment X-LM, labor supply was 3.66 higher for the medium wage than for the low wage (p=0.003). In the other reference wage-free treatment X-MH, the higher wage elicits an increase in labor supply by 1.87 (p=0.066). Higher labor supply at higher wages in both treatments supports Prediction 1 and implies that in the absence of a reference wage the substitution effect in labor supply dominates the income effect – at least in the short run and within our experimental setting. Lastly, the increase in labor supply at higher versus lower wages in these two treatments is not significantly different (3.66 vs. 1.87, t=1.21, p=0.231).

[5] Participants' age, gender, and major of study are balanced along all observable dimensions in our experiment and do not have any significant impact on participants' behaviour in the experiment, in particular the response of labor supply to wage changes.

Result 2: People work more for a higher wage in the gain domain (wage above the reference level).

In treatments L-LM and M-MH, the wages in the warm-up stage were low and medium, respectively. Then, in the subsequent experimental stage, participants were asked how much they would work for a higher wage as well as for the previously experienced lower reference wage. Specifically, the two possible wage rates were low and medium in treatment L-LM and medium and high in treatment M-MH. The within-person differences in labor supply at the higher and lower wage rates in the experimental stage are 3.18 and 5.92 in treatments L-LM and M-MH, respectively, both significantly different from zero (p<0.01). The willingness to work significantly more when wages increase from the previously experienced reference wage level supports Prediction 2.

Prediction 2 also states that, in the gain domain, labor supply increases with wage at the same rate as in the absence of a reference wage. While labor supply reactions to wage changes in the treatments X-LM and L-LM are quite comparable, treatments X-MH and M-MH produce significantly different results: 1.87 vs. 5.92 (p=0.006). Still, labor supply positively reacts to wage increases in both treatments.

Result 3: People work more for a lower wage in the loss domain (wage below the reference level).

In treatments M-LM and H-MH, the wages in the warm-up stage were medium and high respectively. Then, in the subsequent experimental stage, participants were asked how much they would work for a lower wage as well as for the previously experienced reference wage. Specifically, the two possible wage rates were low and medium in treatment M-LM and medium and high in treatment H-MH. The within-person differences in labor supply at the higher and lower wage rates in the experimental stage are -5.00 and -4.36 in treatments M-LM and H-MH, respectively, both statistically significant (p<0.01). The willingness to work significantly more when wages decrease from the previously experienced reference wage level supports Prediction 3.

Together, Results 2 and 3 imply that there is a kink in the labor supply curve at the previously experienced reference wage level, and that the income and substitution effects shaping labor supply are domain-specific. Wages falling below the reference point trigger income targeting, as people try to mitigate the disutility from earning less than the target income by working more. This behaviour stems from loss aversion: the utility from lower financial losses relative to an expected reference point is traded for disutility from extra work effort. On the other hand, in the gain domain, when wages rise above the reference point, loss aversion no longer affects effort and substitution effects dominate income effects. That is, people respond to the increased return to labor by sacrificing leisure to produce more when higher rewards to work are available.

Result 4: People tend to work more for a given wage in the loss domain.

We now turn from within- to between-participant, across-treatment comparisons. Starting with labor supply at the low wage, we find that, in treatment M-LM, in which earnings at the low wage would be in the loss domain, average labor supply (20.561) is higher than that in treatments X-LM and L-LM featuring the same menu of wages (average 17.720), although this difference is not statistically significant (p=0.385). A similar exercise for labor supply under the medium wage reveals that labor supply in treatment H-MH, 24.513, when the income under the medium wage is in the loss domain, is higher than the average labor supply in the otherwise identical treatments X-MH and M-MH (average 15.971), a statistically significant difference

(p<0.01). Labor supply at a given wage being higher when the income it brings is in the loss domain is consistent with Prediction 4.

Result 5: Income targeting is not complete.

Although there is income targeting in the loss domain (people working more for a lower wage to bring income closer to the target, Result 3), income nonetheless drops significantly as the wage falls. Table 3 presents the average earnings that would have occurred at each wage level in the experimental stage based on the elicited labor supply commitments in the two unidirectional-down treatments, M-LM and H-MH, where income targeting occurs. In these treatments, where participants make labor-supply decisions for the reference wage as well as for a lower wage, the dominance of income effects over substitution effects in the loss domain partially offsets the fall in income, but does not eliminate it. The difference in income earned given labor supply commitments under the two potential piece-rate wages is still both economically and statistically significant, indicating income targeting is not complete in either treatment (p<0.001 in both cases).

INSERT TABLE 3 ABOUT HERE

The incompleteness of income targeting provides a variation in labor supply in the loss domain that enables us to separately estimate the structural parameters of the theoretical labor supply equation (1), the loss aversion parameter λ and the difficulty of effort parameter θ with our data.[6] Table 4 reports the estimation results. The implied degree of loss aversion (2.4 to 2.7)

---

[6] Specifically, θ is identifiable from changes in labor supply, Δh, in response to changes in wages from the reference level, $\Delta w = w_1 - w_0$ from equations $\Delta h = \Delta w \frac{1}{\theta}$, which holds in the gain domain, and $\Delta h = \frac{T}{w_1} - \frac{T}{w_0} = \frac{w_0^2}{\theta \cdot w_1} - \frac{w_0}{\theta}$, which holds in the loss domain for the observations with complete income targeting. λ is estimable from equation $\Delta h = \frac{\lambda \cdot w_1}{\theta} - \frac{w_0}{\theta}$, which holds in the loss domain for the observations with incomplete income targeting (see Equation 1 and Figure 1 and the accompanying discussion). Observing from the data $\Delta h, \Delta w$, the domain, and whether

is significantly above 1 (*p*<0.001) and is similar in magnitude to the commonly reported values of just over 2. What is more interesting, however, is that people's choices under the reference wage, and hence the implied loss aversion coefficient, change depending on the distribution of the wages above and below the reference wage, which we discuss next.

Result 6: In the bi-directional treatment M-LMH, people choose to work more for a higher wage across all wage levels.

In treatment M-LMH, participants received the medium wage in the warmup stage, followed by an equal probability of a low, medium, or high wage in the experimental stage. This treatment was designed to examine whether the apparent kink in the labor supply curve that emerged when comparing the unidirectional-up and unidirectional-down treatments would continue to manifest itself in a bidirectional treatment, which could result in either gain or loss relative to the reference wage and implied target income. It did not. Labour supply increased by 4.10 (*p*<0.001) moving from the low to the medium wage and by 3.20 (*p*<0.001) moving from the medium to the high wage, indicating the dominance of substitution over income effects both above and below the reference wage.

While the labor supply changes in response to a higher wage are similar in magnitude to those observed in the gain domain in the other treatments, the apparent absence of the dominant income effect in the loss domain (low vs. medium wage with the medium wage as the reference point) as previously observed in the unidirectional treatments is puzzling. One possible explanation for this empirical puzzle is a preference for consistency (Cialdini et al., 1995;

---

income targeting was complete or not in the loss domain, we estimate the empirical equivalent of all three of the above equations put together: $\Delta h = \Delta w \frac{1}{\theta} \cdot \left(1 - I(loss)\right) + \left(\frac{w_0^2}{\theta \cdot w_1} - \frac{w_0}{\theta}\right) \cdot I(loss) \cdot I(complete)i + \left(\frac{\lambda \cdot w_1}{\theta} - \frac{w_0}{\theta}\right) \cdot I(loss) \cdot \left(1 - I(complete)\right)$, where $I(loss) = 1$ if income is in the loss domain or 0 if income is in the gain domain and $I(complete) = 1$ if income targeting is complete or 0 if income targeting is incomplete. We use nonlinear least squares estimator.

Guadagno and Cialdini, 2010; Falk and Zimmermann, 2013), that is, a desire to "respond to … situations in a manner consistent with prior attitudes, behaviors, and commitments, particularly when … consistency is salient to them" (Guadagno and Cialdini, 2010, p. 153). A nonmonotonic labor-supply response to wage changes may appear less consistent than a monotonic one, producing an observationally similar relationship between labor supply and wages as in the treatments with no reference wage. Another possible explanation is provided by the salience theory of choice under risk (Bordalo et al., 2012): more alternatives in treatment M-LMH, including the status quo, make the loss less salient, shifting the focus from offsetting the loss by working more to substitution between work and leisure in response to changing wages as observed in the treatments with no reference wage. Both of the above explanations are *ex-post*, and require more research to further explore their potential impact on labor-supply decisions as wages change. What is clear from our results, however, is that the income and substitution effects shaping labor supply are not only domain-specific but also are sensitive to other behavioral framings of the labor-supply choice problem.[7]

DISCUSSION AND CONCLUSION

Labor supply reactions to wages is among the most fundamental questions of labor economics. As the world of work moves away from a fixed nine-to-five schedule with rigid wages to more flexible arrangements, such as gig employment where both hours and rates fluctuate, this question becomes ever more important. Our work contributes to the existing literature on labor supply that allows for the presence of income targeting (e.g. Camerer et al., 1997; Crawford and Meng, 2011; Farber, 2008; Fehr and Goette, 2007; Abeler et al., 2011). In

[7] Note that risk aversion is unlikely to be an explanation here. First, risk preference cannot explain the lack of the kink in labor supply. Second, while risk preferences do affect choices, including labor supply, when outcomes are uncertain at the time of making the choice (Cadsby, Song, Zubanov, 2019), in our study people specify their labor supply *conditional* on a specific realization of the wage. Therefore, given the wage, there is no income risk.

these studies, with the exception of Abeler et al. (2011), the income targets were not observed and had to be inferred structurally. By experimentally manipulating clearly set reference wage levels and eliciting labor supply responses with the within-person strategy method, our study provides an ideal setting for examining the short-run impact of wages on labor supply with and without income targets[8].

To summarize our findings, in the absence of a reference wage, people are willing to work more for higher wages (Result 1). This is consistent with the predictions of the neoclassical model of labor supply for the case of a temporary change in wages. However, when a reference wage is introduced, people work more for higher wages in the gain domain, at a wage above the reference level (Result 2). In contrast, in the loss domain, with the wage below the reference level, people work more for *lower* wages (Result 3). We interpret these findings as consistent with income targeting, when people perceive income below the target as a loss and try to mitigate it by working more (Results 4 and 5). The rather dramatic kink in the labor supply curve that the above results reveal is sensitive to the context in which labor supply decisions are made, however. If wages may either rise or fall relative to the reference point, so that workers must simultaneously consider the gain and loss domains, the kink disappears and more labor is supplied at higher wages above or below the reference point (Result 6).

To conclude, our results suggests that income targets may matter for labor supply decisions by causing people to work more in order to mitigate the loss from income falling below the target level. However, income targets are not omnipresent and may vary, both in

[8] Brandts and Charness (2011) surveyed the literature regarding whether the strategy method, in which a responder makes conditional decisions for each possible information set, leads to different experimental results than does the more standard direct-response method, in which the responder learns the action of the first mover and then chooses a response. They found that any effect identified by using the strategy method is also always found by using the direct-response method, indicating that the strategy method is a conservative approach to identify any likely effects, which are likely to be at least as large in a direct-response environment.

magnitude and importance, depending on the earnings history and the context in which labor supply decisions are made. People earning more in the past will not always clock in more hours when their wages go down, even if they are loss-averse. Further research should look more closely at how income targets emerge, as well as consider other behavioural aspects of labor-supply decision making.

References

Abeler, J., Falk, A., Huffman, D., Goette, L. (2011). Reference point and effort provision. *American Economic Review,* 101, 470-492.

Bordalo, P., Gennaioli, N., Shleifer, A. (2012) Salience theory of choice under risk. *Quarterly Journal of Economics*, 127(3), 1243-1285.

Bracha, A., Gneezy, U. and Loewenstein, G. (2015). Relative pay and labor supply. *Journal of Labor Economics*, 33, 297-315.

Brandts, J., and Charness, G., (2011). The strategy versus the direct-response method: A first survey of the experimental comparisons. Experimental Economics, 375-398.

Cadsby, C. B., Song, F., and Zubanov, N. (2019). The "Sales Agent" problem: Effort/leisure allocation under performane pay as behavior towards risk. *Economic Inquiry* 57 (4), 1997-2016.

Camerer, C., Babcock, L., Loewenstein, G., Thaler, R. (1997). Labor supply of New York City cabdrivers: One day at a time. *Quarterly Journal of Economics*, 112(2), 407–441.

Carrington, W. (1996). The Alaskan labor market during the pipeline era. *Journal of Political Economy*, 104, 186-218.

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

Chou, Y.K. (2002) Testing alternative models of labour supply: Evidence from taxi drivers in Singapore. *The Singapore Economic Review*, 47(1), 17-47.

Camerer, C. (2000) Prospect theory in the wild: Evidence from the field. In *Choices, Values, and Frames*, D. Kahneman and A. Tversky (eds.), Cambridge University Press.

Cialdini, R.N., Trost, M.R., Newsom, J.T. (1995) Preference for consistency: the development of a valid measure and the discovery of surprising behavioral implications. *Journal of Personality and Social Psychology*, 69(2), 318-328.

Crawford, V., Meng, J. (2011). New York City cabdrivers' labor supply revisited: Reference-dependent preferences with rational-expectations targets for hours and income. *American Economic Review,* 101, 1912-1932.

Dickinson, D. (1999). An experimental examination of labor supply and work intensity. *Journal of Labor Economics*, 17, 638-670.

Else-Quest, N., Hyde, H, Linn, M. 2010. Cross-national patters of gender differences in mathematics: A meta-analysis. Psychological Bulletin, 136, 103-127.

Falk, A., Zimmermann, F. (2013) A taste for consistency and survey response behavior. *CESifo Economic Studies*, 59(1), 181-193.

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 lessons from cab drivers. *Quarterly Journal of Economics*, 130(4), 1975-2026.

Fehr, E., Goette, L. (2007). Do workers work more if wages are high? Evidence from a randomized field experiment. *American Economic Review,* 97, 298-317.

Guadagno, R.E., Cialdini, R.N. (2010) Preference for consistency and social influence: A review of current research findings. *Social Influence*, 5(3): 152-163.

Kahneman, D., and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47: 263-291.

Kőszegi, B., Rabin, M. (2006). A model of reference-dependent preferences. *Quarterly Journal of Economics*, 121, 1133-1165.

Lazear, E., Shaw, K., Stanton, C. (2016). Making do with less: Working harder during recessions. *Journal of Labor Economics*, 34, 333-360.

Niederle, M., Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics*, 122, 1067-1101.

Oettinger, G. (1999). An empirical analysis of the daily labor supply of stadium vendors. *Journal of Political Economy*, 107, 360-392.

Stafford, T. M. (2015). What do fishermen tell us that taxi drivers do not? An empirical investigation of labor supply. *Journal of Labor Economics*, 33(3): 683–710.

Wakker, P. (2010) *Prospect theory: For risk and ambiguity*, Cambridge University Press.

Figure 1: Labor supply (h) as a function of wage (w) with and without reference wage w_0 and target income ($T = \frac{w_0^2}{\theta}$).

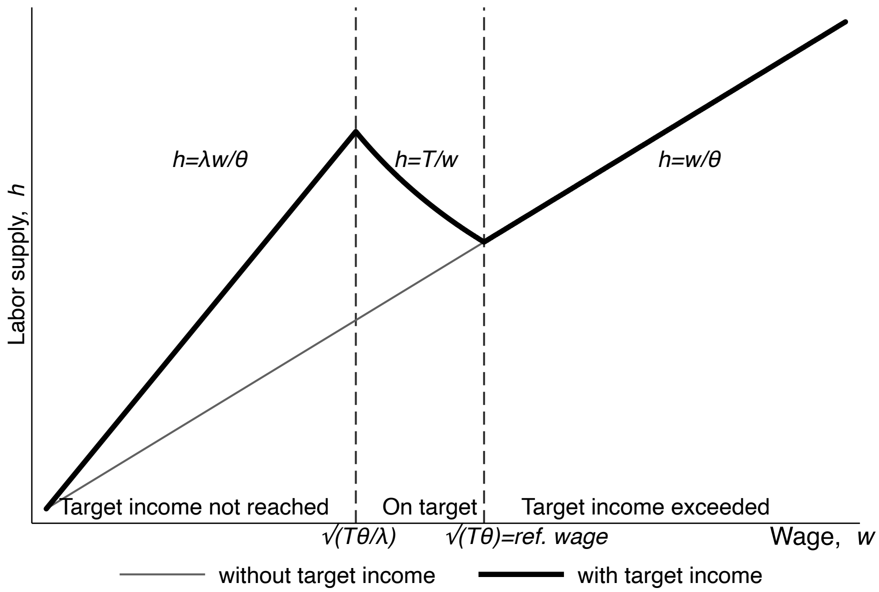

Figure 2: Mean labor supply by treatment and wage level

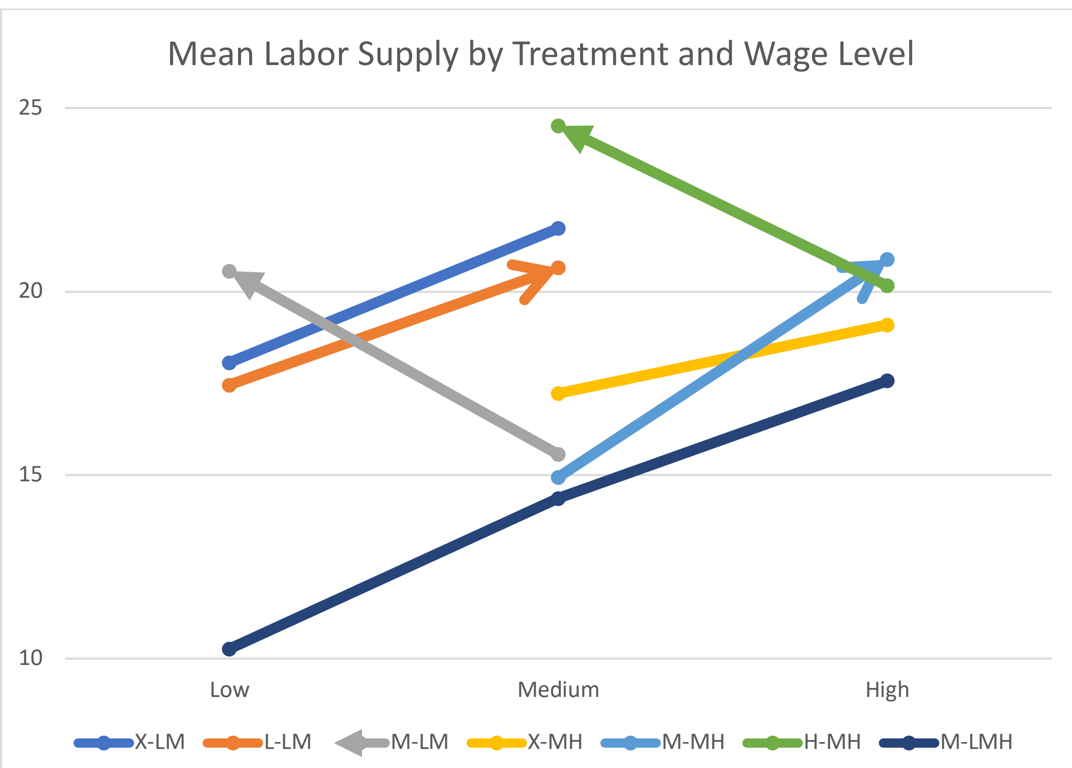

Table 1: Experimental treatments summary

	Treatment Type	Specific Treatment	Reference wage (paid in the warmup stage)	Experimental stage wage levels (equiprobable, shaded areas)		
				Low	Medium	High
1	No Reference Wage	X-LM	None			n/a
2	Unidirectional Up	L-LM	Low			n/a
3	Unidirectional Down	M-LM	Medium			n/a
4	No Reference Wage	X-MH	None	n/a		
5	Unidirectional Up	M-MH	Medium	n/a		
6	Unidirectional Down	H-MH	High	n/a		
7	Bidirectional	M-LMH	Medium			

Notes : The first letter in the treatment label denotes the wage level offered during the warmup stage (L, M or H, or X if no wage was offered), and the following letters denote the combinations of possible, and equally likely, wages in the experimental stage.

Table 2: Mean labor supply by treatment and wage level

	Treatment	Warmup Stage	Experimental stage, by wage level Low	Medium	High	Changes in labor supply with wage L to M	M to H
1	X-LM	2.552	18.069	21.724		3.655***	
	n=29	(0.270)	(2.153)	(2.207)		(1.106)	
2	L-LM	2.744	17.462	20.641		3.179***	
	n=39	(0.274)	(1.794)	(1.865)		(0.868)	
3	M-LM	2.439	20.561	15.561		-5.000***	
	n=41	(0.204)	(2.960)	(2.126)		(1.184)	
4	X-MH	2.839		17.226	19.097		1.871*
	n=31	(0.275)		(2.395)	(2.319)		(0.980)
5	M-MH	3.026		14.947	20.868		5.921***
	n=38	(0.198)		(1.773)	(2.224)		(1.041)
6	H-MH	3.077		24.513	20.154		-4.359***
	n=39	(0.253)		(1.893)	(1.637)		(1.115)
7	M-LMH	2.533	10.267	14.367	17.567	4.100***	3.200***
	n=30	(0.218)	(1.493)	(1.808)	(2.022)	(0.782)	(0.839)
Average across treatments		2.753	16.950	18.470	19.551	1.065	1.514
p-value test of equality across treatments		0.359	0.018	0.001	0.704	0.000	0.000

Notes: The first letter in the treatment label denotes the wage level offered during the warmup stage (L, M or H, or X if no wage was offered), and the following letters denote the combinations of possible, and equally likely, wages in the experimental stage. Standard deviations are in parentheses. *** (**) [*] indicates changes are significant at the 1% (5%) [10%] level in a two-sided *t*-test with regression standard errors clustered by participant ID (the level at which various treatments were applied).

Table 3: Average Earnings in the experimental stage in the Unidirectional-Down Treatments

Treatment	Low wage ($0.50)	Medium wage ($1.00)	High wage ($2.00)
M-LM, n=41	$10.28	$15.56	
H-MH, n=39		$24.51	$40.30

Table 4: Estimated loss aversion (λ) and difficulty of effort (θ) parameters from the theoretical labor supply equation (1)

Parameter	All data	All data, with controls	Excluding the data from the bi-directional treatment M-LMH
Loss aversion λ	2.443	2.400	2.713
	(0.224)	(0.221)	(0.229)
Difficulty of effort θ	0.235	0.244	0.218
	(0.030)	(0.031)	(0.031)

Notes: The parameters are estimated with nonlinear least squares method. Standard errors clustered by participant ID are in parentheses.

Appendix: Experimental Instructions

Instructions

Thank you for participating today.

All of your responses in this study will remain completely anonymous. It is important that during this experiment you do not talk or make any noise that might disrupt others around you. If you have any questions, please raise your hand and the experimenter will answer your questions individually.

During this experiment you will be asked to add up sets of five double-digit integers such as the following.

98	42	69	50	78	

The first round is a 2-minute warm-up round for you to get familiar with the task. It will be followed by an experimental round. Both the warm-up and the experimental rounds will be used to calculate your earnings as explained below. You are not allowed to use a calculator, but may write numbers down on scratch paper provided by us. The numbers are randomly drawn and each problem is presented as above.

You will have a *Workbook* that will contain all of your work. Your task in each round is to solve such summing problems. Your earnings in this experiment will depend on your performance and/or the specific compensation method applied to the warm-up and the experimental rounds. Once we begin the experiment, you will not be able to look ahead to future pages or to go back to previous pages.

To ensure confidentiality, just write down your participant number on each page of the *Workbook*. Please do not write your name on any of these materials.

Please make sure that you completely understand the instructions for the experiment. Once again, remember not to make any noises that might disturb others around you. If you have any questions, raise your hand and we will answer your questions individually.

Warm-up Round:
Please write all of your work in this *Workbook* and turn pages only when instructed to do so.

The next round is a warm-up round and it will last for 2 minutes.

Please solve the problems in the order presented (i.e. You are NOT allowed to skip problems).

You will earn \$1.00 for each problem you solve correctly.

Thus, your total earnings for this round will be: \$1.00 × the number of problems solved correctly.

Sample Experimental Round:

For this round, there are 2 possible pay levels below:

Method A	**Method B**
$0.50/problem solved	$1.00/problem solved

Each of these two options has an equal chance of being chosen for your actual work output and payment. You will roll a 6-sided die. If the die-roll outcome is an odd number (i.e. 1, 3 or 5), then Method A will be chosen for both work output and payment. If the die-roll outcome is an even number (i.e. 2 , 4 or 6), then Method B will be chosen for both work output and payment.

Please write down the number of problems you'd like to commit yourself to work on under each method. Please note that these are binding commitments. When one of these two methods is chosen at random, you are committed to solve the number of problems you committed yourself to solve in order to receive payment. They must be solved correctly. If there are any incorrect answers, you will be given the opportunity to redo those problems. You will then be paid according to the payment method randomly chosen by the die roll.

Now please take a minute to make your decisions on how many problems you'd like to solve under each method:

	Method A	**Method B**
	$0.50/problem solved	$1.00/problem solved
I will solve	**_________problems**	**_________problems**