How Do Expectations Influence Labour Supply? Evidence from a Framed Field Experiment

Lisa Stockley *

Draft Date: November 17, 2015

For the most current version of this paper, please visit:
http://individual.utoronto.ca/lisastockley/documents/LisaStockleyJMP.pdf

Abstract

Labour income is a critical resource for the world's poor, yet remarkably little is conclusive about how labour supply is determined. To reconcile disparate evidence regarding individuals' response to wage changes, a leading behavioural theory proposes that in addition to valuing the level of income, workers evaluate income as gains or losses with respect to their expectations. In this paper, I am the first to test this model of labour supply in a real effort framed field experiment. Specifically, I conduct a set of experiments among a sample of impoverished individuals involved in piece-rate work in Northeast Brazil. I manipulate workers' probabilistic beliefs about income and check if these expectations determine labour supply. In both experiments, I find that expectations do influence effort: if expectations are high, participants work less than if expectations are low. This pattern is inconsistent with the leading behavioural model's predictions and existing laboratory evidence, suggesting that how expectations influence effort may vary with context.

*Contact: Department of Economics, University of Toronto, 150 St. George Street, Toronto ON M5S 3G7. Phone: (647) 213-1798. Email: lisa.stockley@utoronto.ca. This project was supported by a grant from The Russell Sage Foundation (Project# 98-14-12) and a Mitacs Globalink Fellowship (Ref.# IT04817). IRB approval was granted by the University of Toronto and Innovations for Poverty Action. I thank Gustavo J. Bobonis, Marco Gonzalez-Navarro, and Tanjim Hossian for their supervision, advice and guidance. I also thank John List, Philip Oreopoulos, Nicola Lacetera, Mitchell Hoffman, Edward Miguel, Alfia Karimova, Juan Sebastian Morales, Graham Beattie and the participants at workshops at the University of Toronto, the 49th annual CEA conference, and the 2015 Southern Ontario Behavioural Decision Research Conference for their helpful comments. All errors and omissions are my own.

1 Introduction

Labour income is a critical resource for the world's poor, yet remarkably little is conclusive about how labour supply is determined. A large literature presents conflicting evidence about the way individuals respond to wage changes.[1] The standard model of inter-temporal labour supply suggests that workers respond positively to transitory wage shocks: they work more when wages are high and substitute to leisure when its price – the foregone wage – is low. However, estimated transitory wage elasticities are often negative or insignificant.[2] A leading explanation for this behaviour is that workers exhibit reference-dependent preferences. That is: (i) they make labour supply decisions over very narrow time horizons,[3] and (ii) they have some target level of income after which the marginal return to labour drops discontinuously.

To reconcile disparate evidence[4], Kőszegi and Rabin (2006, 2007, 2009) [henceforth KR] propose a model of *rational expectation based reference-dependent* labour supply. Their theory suggests that in addition to valuing the level of income, workers evaluate income as gains or losses with respect to their recently held probabilisitic beliefs (i.e., expectations) about that income. Since loss averse workers are more motivated to avoid feelings of loss than to acquire gains, the marginal return to effort drops discontinuously once accumulated income exceeds expected income. As a result, effort increases in expectations of income. To test the KR model, Abeler et al. (2011) conduct a laboratory experiment with university students. Using a lottery-based payment contract, they manipulate students' expectations of income and check if the expectations have an independent effect on effort. They find that if expectations are high, students work more than if expectations are low, and conclude that the KR model best fits the behaviour of their participants.

I am the first to test the KR model of labour supply in a real effort framed field experiment.[5]

[1] For a brief synopsis, see Chetty et al. (2011*b*). For a meta-analysis of quasi-experimental extensive margin elasticities, see Chetty et al. (2011*a*). For a meta-analysis of intensive margin elasticities from micro data see Chetty (2012).

[2] For instance, consider Mankiw, Rotemberg, and Summers (1985), Browning, Deaton, and Irish (1985), Altonji (1986), Laisney, Pohlmeier, and Staat (1992); Pencavel (1986); and Mulligan (1995).

[3] Read et al. (2000) defines isolating decisions as if they are not embedded in an stream of decisions as "narrow bracketing". This is closely linked to Thaler's (1985) notion of "mental accounting" – the way in which gamblers evaluate the outcomes of a particular day of gambling as independent of the outcomes on other days spent gambling

[4] e.g., Camerer et al. (1997), Oettinger (1999), Chou (2002), Fehr and Goette (2007), Farber (2008), Crawford and Meng (2011), Farber (2014), Dupas and Robinson (2014), and Agarwal et al. (2014)

[5] A framed field experiment is defined by Harrison and List (2004) as an experiment using a non-standard subject

Specifically, I conduct a pair of experiments to test the KR model's predictions among a sample of impoverished individuals involved in piece-rate work in Northeast Brazil. In the first experiment, I follow Abeler et al.'s (2011) design and manipulate workers' rational expectations of income with a lottery-based payment contract for an open-ended shift of work producing a familiar output. Once participants quit working, their payment is determined by a coin-flip: half the time, they receive their piece-rate earnings; otherwise, they receive a fixed payment. Under a neoclassical framework, participants' behaviour does not respond to the size of fixed payment as it does not contribute to the marginal return to effort. In contrast, if participants have KR preferences, their effort increases in their probabilisitic beliefs about income, and as such, will increase in the fixed payment.

In the second experiment, conducted three weeks after the first, I create a gap between participants' probabilisitic beliefs about income and the payment lottery they face in a second shift of work. Specifically, I return to a subsample of participants and offer them the opportunity to work an additional open-ended shift for a payment lottery. Among those who agree, some are offered the same lottery as during their first shift, while others are offered unexpectedly higher or lower piece-rate wages than the wages offered in the first round. Under a neoclassical framework, participants' current behaviour does not respond to their past piece-rate wage as it does not contribute to the marginal return to current effort. In contrast, if participants have KR preferences, their effort increases in the past wage, since recently held probabilisitic beliefs about income are determined by the payment lottery faced in the first shift.

In both experiments, I find that expectations do influence effort: if expectations are high, participants work less than if expectations are low. But this pattern of behaviour is the opposite of that predicted by the KR model and observed by Abeler et al. (2011) in the lab. In the first experiment, I estimate a negative wage elasticity of labour supply of -0.22 and a negative fixed payment elasticity of labour supply of -0.31 (both significant at the 10 percent level) that cannot be rationalized by the KR model. In the second experiment, the relationship between current and past wages and effort are imprecisely estimated for the whole sample. But for the subsample of

pool (i.e., not university students or any other low cost but low relevance population) in a field context where the subjects are familiar with the incentive scheme, commodity, or task parameters.

participants whose first and second shift circumstances were most similar,[6] I estimate a past wage elasticity of labour supply of -0.45 (significant at the 10 percent level). This finding suggests that those who have low expectations work more, not less, than their equally well paid counterparts.

Although participants do not behave as though they have KR preferences, their recently held probabilistic beliefs about income do affect their behaviour. In the first experiment, through both dimensions of the payment lottery, higher expected income reduces effort. If participants evaluate the money earned in this experiment in isolation from the ways they typically earn money, then this pattern is easily explained as an income effect. In the second experiment, the negative past wage elasticity may capture a notable feature in the distribution of responses – the propensity to keep labour supply constant. We might expect participants who experience wage changes to update their behaviour accordingly, but in this experiment I observe the opposite. Of those who experience wage changes, 21 percent keep their effort constant – significantly more than the 9 percent of participants who did not experience a wage change. This finding suggests two things: (i) the negative past wage elasticity captures the persistence of the effort supplied from the first shift; (ii) the heuristics[7] used to determine labour supply likely evolve with experience.

My results contribute to a growing literature that finds KR preferences are insufficient for explaining the labour supply behaviour of populations from the developing world. In the field, I find statistically different results than Abeler et al. (2011) find in the lab. These differences highlight that the influence of expectations on effort may be context specific, and emphasize a need for caution when extrapolating from WEIRD[8] behaviours to forecast the behaviour of workers in developing countries. Finally, workers' propensity to keep labour supply constant, especially when faced with wage changes, suggests that the heuristics[7] that determine effort evolve with experience.

The paper is organized as follows. The theoretical framework is explained in the following section. Section 3 outlines the experimental design, including an overview of the context where this experiment was conducted. Section four outlines the empirical methodology, Section 5 reports

[6] The shifts occurred on the same weekday or started at approximately the same time

[7] "Heuristics are efficient cognitive processes, conscious or unconscious, that ignore part of the information. Because using heuristics saves effort, the classical view has been that heuristic decisions imply greater errors than do 'rational' decisions as defined by logic or statistical models" – p.451 (Gigerenzer and Gaissmaier, 2011).

[8] Western Educated Industrialized Rich Democratic (Henrich, Heine and Norenzayan, 2010)

results, and Section 6 discusses them. Section 7 concludes my analysis.

2 Theoretical Framework

The theoretical framework follows Kőszegi and Rabin's (2006, 2007, 2009) model of rational expectation based reference dependence in a way that maps to the experimental design. The model captures a loss averse individual's "gain-loss utility" derived from standard consumption utility and reference points. Within this static model, I discuss two cases: the case where agents hold correct probabilisitic beliefs about income, which will map to the first shift of the experiment; and the case where these beliefs are wrong, which will map to the second shift of the experiment.

I begin with a model of labour supply where an agent chooses her optimal effort e while facing some income uncertainty. At the time she makes her effort choice, she knows that with probability p she will be paid some function of her effort $y(e)$, and with probability $(1 - p)$ she will receive a fixed payment f regardless of e. In this simple setting without income effects, her consumption utility $u(.)$ is separable and increasing in each component of income. She also experiences some positive and weakly increasing cost of effort $c(e)$ regardless of the income awarded. As such, her optimization problem is:

$$\max_{\{e\}} \mathbb{E}\big[U(e)\big] = pu(y(e)) + (1 - p)u(f) - c(e) \tag{1}$$

which leads to the optimal effort choice e^* that is independent of f and characterized by:

$$c\text{'}(e^*) = py\text{'}(e^*)u\text{'}(y(e^*)). \tag{2}$$

Next, I expand this model building on the basic intuition of Kahneman and Tversky's (1979) Prospect Theory – agents will additionally evaluate realized incomes as gains or losses with respect to some reference point. In this model, the reference point is the "probabilistic beliefs she held in the recent past about" incomes (Kőszegi and Rabin, 2006). Agents evaluate each potential income according to its expected utility, with the utility of each income being the average of how it feels relative to each possible realization of the reference point.[9]

Suppose this agent has recently held probabilistic beliefs about the income uncertainty, that may or may not have be correct. Suppose she believes that with probability $\tilde{p}$ she will be paid some function of her effort $\tilde{y}(e)$, and with probability $(1-\tilde{p})$ she will be paid a fixed payment $\tilde{f}$ regardless of e. As such, in addition to the expected consumption utility described in (1), the agent also experiences:

$$\mathbb{E}\big[V(e)\big] = p\Big[\tilde{p}\mu\big(y(e)-\tilde{y}(e)\big) + (1-\tilde{p})\mu\big(y(e)-\tilde{f}\big)\Big] + (1-p)\Big[\tilde{p}\mu\big(f-\tilde{y}(e)\big) + (1-\tilde{p})\mu\big(f-\tilde{f}\big)\Big].$$

The term in the first square bracket captures the average of how $y(e)$ feels with respect to each potential income, and similarly, the second square bracket captures the average of how f feels with respect to each potential income. In her total expected utility, each of these average feelings of gain or loss are scaled by the probability that particular realization of income will occur.

The function $\mu(.)$ transforms the difference between her realized income and her expectations into gains and losses of utility. This function needs to have two features: (i) $\mu(0)=0$ and (ii) for any $K>L>0$, $\big|\mu(K-L)\big| < \big|\mu(L-K)\big|$. The first ensures that if expectations are exactly met, they do not contribute to total utility. This means that when expectations perfectly match reality and there is no uncertainty, people will behave as if they are not reference-dependent. The second ensures that agents are more motivated to avoid losses than they are to acquire gains. This notion of loss aversion, introduced by Kahneman and Tversky (1979), is often captured by a loss aversion parameter $\lambda > 1$ used to weight the magnitude of losses over gains in total utility. Characterizing $\mu(.)$ with linear loss aversion:

$$\mu(s) = \begin{cases} s & \text{if } s \geq 0 \\ \lambda s & \text{if } s < 0 \end{cases}$$

captures all the features necessary for this analysis without adding unnecessary complexity.

If this agent cares about her level of income, as well as how it compares to her recently held

[9] More than this, the reference points are endogenously determined by the economic environment in the KR model. Agents in this model chose reference points such that, conditional on expectations, the agent's utility is maximized when the realizations of the outcome equal her expectations. The authors refer to this as a personal equilibrium.

expectations, then her optimization problem can be written as:

$$\max_{\{e\}} \mathbb{E}\big[U(e) + \eta V(e)\big]. \tag{3}$$

$U(e)$ is the consumption utility valuing the level of income described in (1), and $V(e)$ captures how feelings of perceived gain or loss impact total utility as described by (2). Thus, the parameter $\eta \geq 0$ can be thought of as the degree of reference dependence because it characterizes the weight given to the reference-dependent utility relative to her consumption utility (normalized to 1).

2.1 Case 1 - Expectations are Correct

If the agent's recently held expectations of the income uncertainty are correct (i.e, $\tilde{y}(e) = y(e)$, $\tilde{f} = f$, and $\tilde{p} = p$), then I can re-write her total expected utility as:

$$\mathbb{E}\big[U(e) + \eta V(e)\big] = pu(y(e)) + (1-p)u(f) - c(e) + \eta p(1-p)\Big[\mu\big(y(e) - f\big) + \mu\big(f - y(e)\big)\Big].$$

As such, she will choose her effort to satisfy:

$$c'(e^*) = \begin{cases} py'(e^*)u'(y(e^*)) + \eta p(1-p)\big(\lambda - 1\big) & \text{if } y(e^*) < f \equiv \text{ ``Avoiding Losses''} \\ \\ py'(e^*)u'(y(e^*)) - \eta p(1-p)\big(\lambda - 1\big) & \text{if } y(e^*) \geq f \equiv \text{ ``Acquiring Gains''}. \end{cases} \tag{4}$$

This expression shows that loss-averse, reference-dependent agents are more motivated to work when earnings are less than the fixed payment compared to when earnings are more than the fixed payment. The right-hand-side of this equation is the discontinuous marginal benefit of producing an additional unit of output e. If this agent is loss averse and reference-dependent (i.e., $\lambda > 1$ and $\eta > 0$), then the marginal benefit of each unit of effort is higher when her piece-rate earnings from working $y(e)$ are less than the fixed payment f. That is, she is more motivated to avoid perceived losses than she is to acquire gains. In contrast, if this agent is not reference-dependent or she is not loss averse (i.e., $\eta = 0$ or $\lambda = 1$), the decision equation (2) collapses to equation (1) and she makes her optimal effort choice independent of fixed payment f.

The shape of this discontinuous marginal benefit function generates two predictions about how expectations of income may influence the behaviour of KR agents. Since she is more motivated to work when $y(e) < f$ than when $y(e) \geq f$ because of loss aversion, anything that shifts the $y(e) = f$ threshold will affect behaviour. First, shifting the threshold to a higher level of effort increases the average marginal rate of return, and as such, the average level of effort also increases. Second, the discontinuous drop in marginal benefit that happens after $y(e) = f$ means that an agent is more likely to quit once she has produces enough output to cross this threshold.

To summarize the predictions about behaviour when expectations match the reality of the income lottery, if agents have KR preferences (i.e., $\eta > 0$ and $\lambda > 1$):

1. average effort increases as the $\bar{e}$ that satisfies $y(\bar{e}) = f$ increases
2. sharp increase in the probability of quitting once $y(e) \geq f$.

In contrast, if agents only have consumption utility (i.e., $\eta = 0$ or $\lambda = 1$) and they experience no income effects, then neither of the predictions hod. The average effort of such agents is independent of the $\bar{e}$ that satisfies $y(\bar{e}) = f$ and there is no sharp increase in the probability of quitting once $y(e) \geq f$ increases. If they do experience income effects, average effort decreases in expected wealth.[10]

2.2 Case 2 - Expectations are Incorrect

Suppose that the agent's recently held expectations about the fixed payment f and the probability p are correct, but she is wrong about her expectation of $y(e)$. That is, $y(e) \neq \tilde{y}(e)$.[11] I can re-write her total expected utility as:

$$\mathbb{E}\big[U(e)+\eta V(e)\big] = pu(y(e))+(1-p)u(f)-c(e)+\eta p\Big[p\big[\mu\big(y(e)-\tilde{y}(e)\big)\big]+(1-p)\big[\mu\big(y(e)-z\big)+\mu\big(z-\tilde{y}(e)\big)\big]\Big].$$

[10] Expected wealth is: $py(\bar{e}) + (1-p)f$.

[11] The complementary case to this example is that she is correct about $y(e)$ but wrong about f. The case where $z \neq \tilde{f}$ describes a model of reference dependence comparable to what Camerer et al. (1997) describe for NYC taxi drivers. This is the case of an external income target. If $y(e) = we$, this assumption generates the predictions of negative wage elasticities of supply, but an effort decision that is independent of f.

Her optimal effort decision relies on a marginal benefit function that is discontinuous at the points where $y(e) = \tilde{y}(e)$, $y(e) = z$, and where $z = \tilde{y}(e)$.[12] In the range of effort where she always experiences losses with respect to the fixed payment,[13] her optimal decision is characterized by:

$$c\text{'}(e^*) = \begin{cases} py\text{'}(e^*)u\text{'}(y(e^*)) + \eta p(1-p)\big(\lambda y\text{'}(e^*) - \tilde{y}\text{'}(e^*)\big) + \lambda\Big[\eta p^2\big(y\text{'}(e^*) - \tilde{y}\text{'}(e^*)\big)\Big] & \text{if } y(e^*) < \tilde{y}(e) \\ \\ py\text{'}(e^*)u\text{'}(y(e^*)) + \eta p(1-p)\big(\lambda y\text{'}(e^*) - \tilde{y}\text{'}(e^*)\big) + \Big[\eta p^2\big(y\text{'}(e^*) - \tilde{y}\text{'}(e^*)\big)\Big] & \text{if } y(e^*) \geq \tilde{y}(e) \end{cases}$$

and there are similar decision rules for the other ranges of effort[14] that all capture the same pattern: there is a sharp decrease in the marginal return to effort at the point where accumulated earnings $y(e)$ crosses the threshold of one of her expected incomes. This generates a comparable set of predictions about behaviour described for Case 1. The amount of effort provided increases as the thresholds where her accumulated earnings $y(e)$ equal her expectations shift to higher levels of output. Furthermore, because of the sharp decrease in marginal returns to labour after one of these thresholds is crossed, a worker is much more likely to quit once her accumulated earnings $y(e)$ excess one of her expected incomes.

For example, consider re-hiring an employee who expects, that like during her previous employment, she will be paid we or f with equal probability for producing output e in an open-ended shift of work. Upon agreeing to work, she experiences a wage shock. Her new payment contract now pays her $w'e$ or f with equal probability. Although her expectations are correct about the fixed payment and the probabilities of payment of each potential income, her expectations about the accumulated income do not match reality. That is, $we \neq w'e$. As she chooses her effort e, she is more motivated when avoiding losses with respect to both f and we. In the range of output where she always experiences losses with respect to the fixed payment,[15] her optimal output choice is characterized by:

[12]The fourth discontinuity is where $z = \tilde{f}$ which is always true because her expectations about f are correct.
[13]The range of output e where $y(e) < z$ and $\tilde{y}(e) < z$
[14]For instance, where $y(e) > z$ and $\tilde{y}(e) > z$, and those where f lies in the middle
[15]The range of output e where $we < f$ and $w'e < f$

$$c'(e^*) = \frac{w}{2}u'(w'e^*) + \frac{\eta}{4}\Big[\big(\lambda w - w'\big) + \lambda\big(w - w'\big)\Big] \tag{5}$$

and there are similar decision rules for the other ranges of effort[16] that all capture the same pattern: workers who receive a negative wage shock (i.e., $w' < w$) are more motivated to work than those who receive no shock (i.e., $w' = w$) because they are trying to avoid perceived losses. Similarly, workers who receive a positive wage shock (i.e., $w' > w$) are less motivated to work than those who receive no shock (i.e., $w' = w$) because they have already achieved the satisfaction of exceeding their expectations.

To summarize the predictions about behaviour when expectations do not match the reality of the income lottery, when we compare two agents earning the same wage w', if agents have KR preferences (i.e., $\eta > 0$ and $\lambda > 1$), their behaviour should reflect:

1. average effort increases in past wages: $e^*\big|_{w'<w} > e^*\big|_{w'=w} > e^*\big|_{w'>w}$

Contrastingly, if agents have only consumption utility (i.e., $\eta = 0$ or $\lambda = 1$), then their average effort will be independent of their past wages.

3 Experimental Design

I carry out the experiment in the relatively isolated interior of Northeast Brazil. The impoverished participants of this experiment live within a "vibrant and longstanding garment cluster" (Tendler, 2002) where there are very few economic alternatives to participating in the home-based production of textiles for the domestic market (Tilly et al., 2013). In 2013, approximately 500 000 people lived in the cities that encompass this cluster, with most directly or indirectly surviving on incomes from apparel production (Tilly et al., 2013). The high degree of informality in labour relations, the precariousness of the labour market, and the lack of alternative employment options have generated some of the lowest labour costs in the country (Almeida, 2008). In July 2014, the time of this experiment, the state mandated minimum salary for those with formal employment was 724BRL

[16]for instance, where $we \geq f$ and $w'e \geq f$, and those where f lies in the middle

per month (approximately $325USD). For most households in this population, this represented an upper bound on an individual's income. Only one third of the multigenerational households in the sample reported having even a single member earning an income this high, as most working aged individuals worked informally in the apparel sector or in seasonal agriculture.

The population of garment labourers work from their homes unsupervised. They are free to choose when to work, and once working, they are free to choose when to quit producing. They have the flexibility to inter-temporarily substitute away from work whenever the opportunity cost of their time is high. Their labour is unskilled, repetitive, and manual in nature and their incomes are directly linked to their output. Each of these features is mimicked in the experiment design. The decisions made in this experiment parallel the decisions participants make in their daily lives.

In this project, I randomized payment contracts to 366 adults in 43 neighbourhood clusters[17] for a shift of manual labour in their homes. Participants were recruited from a baseline survey[18] and offered, at most, the opportunity to work 2 shifts of this job. To minimize the chance that participants would learn about the various payment contracts offered to others, these contracts were randomized at the neighbourhood level.

The activity participants performed during this shift of work was to repetitively produce simple output. Production of the output required two inputs other than their labour: a deck of 4.25" x 5.5" cards, each of which had 10 randomly placed black dots[19] (see Figure 1), and a roll of labelling stickers (see Figure 2). Participants used the stickers to completely cover all the dots on each card. Once all dots were covered, the card was considered a produced unit of output.[20]

At the beginning of each shift of this experiment, participants were asked to produce as many units of output as they could in four minutes for a certain payment of 0.05BRL[21]/unit of output. This four minute activity allowed participants to develop their ability and to uncover the costs of

[17] Households that were walking distance from each other and at least 2km from the next neighbourhood cluster.

[18] Baseline household demographic information and localization information courtesy of Marco Gonzalez-Navarro, Gustavo J. Bobonis, Paul Gertler and Simeon Nichter's research regarding clean water access in the semi-arid regions of Brazil

[19] The cards were numbered and there were 100 unique designs that we ordered into identical decks for each participant. The decks of cards were sorted by hand in Brazil, and it was discovered ex post that the ordering was not always identical across decks – although the composition of cards was the same.

[20] Participants were given reminders and unlimited opportunities to fully correct for low quality work, resulting in very rare differences between the amount of output attempted and the amount of output completed

[21] 1 BRL $\approx$ 0.45USD at the time of the experiment - July/August 2014

producing one unit of output. Furthermore, as the surveyors paid participants in cash immediately following the four minutes of production, it helped establish the credibility of the surveyor as someone who will pay participants what is promised.

This particular production task was carefully chosen to meet several goals. First, the characteristics of the production process mimicked the characteristics of production in the local labour market. Second, effort was easily observed, generating a clear measure of labour supply. Third, the output produced was sufficiently different from the output produced in the local labour market, which ensured that participants did not mistake this experiment with something directly related to their livelihood[22] and developed no skills that may be useful outside of the experiment. Finally, it ensured that participants had no expectations about this experience other than what the researchers told them.

3.1 Project Timing and Work Schedule

The experiments took place in July and August of 2014.[23] A team of up to four locally hired professional surveyors travelled to the preselected individual's homes and invited them to participate in a shift of work.[24] Upon their first meeting, surveyors described this shift as participation in two immediate income generating activities: a mandatory four minutes of work in the first activity, and an open-ended amount of work in the second activity that would end at the participant's discretion without penalty. Less than one percent of shifts lasted more than an hour.

The survey team was in the field six days a week[25] from sun up to sun down. The first shift of the experiment was completed in 24 days, and the second shift took an additional 14 days of fieldwork. The average time between a participant's first and second shift was 23 calendar days.[26]

[22]That is, they do not think it is a job interview or a test by their current employers

[23]The first experiment took place after the completion of the locally hosted World Cup.

[24]Only 207 members of this preselected sample were found by the research team. Reported reasons by neighbours and other household members included moving, death, illness, and being at work. Upon discovering that a preselected individual could not be located, the research team attempted to replace this person with another from the same household. If no additional adult members of the household were available, the team would seek to replace within the same neighbourhood cluster. If that was not possible, the observation was dropped. In all, 159 individuals were replaced, created a total sample of 366 individuals

[25]Not Saturdays because this is the day that locals (including the surveyors) go to the market to do the weekly shopping and/or vending home produced goods like vegetables

[26]The minimum was 11 days, and the maximum was 41 days.

3.2 Payment Contract

The goal of the two shifts of the experiment was to link expectations of income to labour supply. In the second activity of both shifts, I used randomly assigned lottery-based payment contracts to manipulate expectations about income, then checked if these expectations had an effect on effort. In the first shift, the exogenous variation in expectations came only though the uncertainty in the payment lottery. In the second shift, the exogenous variation in expectations also came from unexpected changes to the payment lottery itself.

The second activity of each shift of the experiment linked labour supply to probabilistic beliefs over income using a lottery-based payment contract. Unlike the four minute practice round, in the second activity, participants were free to work for as long, or as little, as they liked.[27] Each completed unit of output was eligible to earn a piece-rate wage w. Instead of a certain payment when the participant decided to quit working, there was only a $p = 0.5$ chance she was paid her accumulated piece-rate earnings $y(e) = we$; otherwise, she was given a fixed-payment[28] f instead. This lottery was determined with a coin toss.

Although each participant was fully informed about the payment lottery when making her effort choices, the realization of her income was determined *after* she quit working. This unresolved uncertainty determined each participant's recently held probabilistic belief about income from participation. If she had KR preferences, then since the randomly assigned fixed payment f influenced her expectations about income, her effort would have respond the fixed payment.

Each randomly assigned payment contract contained either a high or low piece-rate wage, and a high or low fixed payment. The values of these payments are presented in Table 2. The combinations of these two wages and two fixed payments lead to four payment contracts offered throughout this experiment.

In the first shift of the experiment, all four potential payment contracts were randomly assigned to participants. Table 3 outlines this two wage × two fixed payment design. Included in Table

[27] There was a maximum amount of work determined by the amount of materials that the interviewer had brought to the site - the deck of 100 cards. In the first shift, 12/366 participants reached this maximum. In the second shift, 1/239 was constrained by this maximum.

[28] The wage w and fixed payments f was varied with experimental treatment assignment unbeknownst to the participants.

3 are the number of participants assigned to each contract and the threshold level of output $\bar{e}$ where the piece-rate earnings equal the fixed payment $w\bar{e} = f$. In this first shift, participants' only expectations about the payment lottery were correct. As such, the only reference points came from the uncertainty in the payment lottery.

In the second shift of the experiment, the survey team revisited all individuals assigned to the high fixed payment treatment arms from the first shift, and offered them the opportunity to work for us again. After completing the four minute practice activity, the surveyor then offered the participant a potentially new payment contract for the second activity. Unlike the first shift, only two potential payment contracts were randomly assigned to participants. All participants were offered the high fixed payment again, but some also received an unexpected wage shock. Half of the workers who earned the low piece-rate wage in the first shift continued to earn the low wage, while the other half received an unexpected increase in their piece-rate wage. Similarly, half of the workers who earned the high piece-rate wage in the first shift continued to earn the high wage, while the other half received an unexpected decrease in their piece-rate wage. Table 4 outlines this two current wage $\times$ two past wage design.

If participants' beliefs about income in the second shift depended on the payment contract they were offered in the first shift, then the unexpected wage shocks led to a deviation between their beliefs and the payment lottery. If the participants had KR preferences, then current effort would have responded to expectations of income, and as such, the past wage rate.

4 Data

The first experiment of the project includes 366 individuals from 43 neighbourhoods, and the second experiment includes only a subsample from the first – 239 individuals from 40 neighbourhoods. Therefore my sample consists of 605 count observations of labour supply. Since the payment contract randomization takes place at the neighbourhood-shift level, the individual observations are not independent, and so the unit of analysis is the neighbourhood-experiment behaviour. The outcome of interest is the amount of output produced in the second activity.

The 366 individuals in the survey came from 186 households, with no more than four adults coming from the same household. Table 1 presents self-reported characteristics of these participants at the time of the first experiment. The typical participant was a woman in her early 40s.

In both shifts of experiment, the four minute practice activity identified the relative productivity of participants. In the first shift, the median participant completed 1.5 cards per minute during this task, with the slowest producing 0.5 cards per minute and the fastest at 3.25 cards per minute. In the second shift, participants worked statistically faster than in the first,[29] suggesting that some learning-by-doing likely took place during the first shift. The median participant completed 1.75 cards per minute, the slowest producing 0.25 cards per minute and the fastest producing 3.75 cards per minute.

In the second activity of the first shift, the average amount of output produced was 20.0 cards and the average amount of time worked was 12.5 minutes. Table 5 reports the unconditional mean and standard deviation of output produced by payment contract treatment cell. There are no statistical differences in the unconditional average behaviour due to treatment. Participants did less work in the second shift than they did in the first.[30] The average amount of output produced was only 16.6 cards and the average amount of time worked was 9.four minutes. Table 6 reports the unconditional mean and standard deviation of output produced by payment contract treatment cell. Again, there are no statistical differences in the unconditional average behaviour due to treatment.

5 Empirical Methodology

5.1 First Shift Analysis

The data collected during the first shift is used to examine the effects of the payment contract on labour supply. To check if probabilistic beliefs of income affect the amount of output produced, I run the following regression:

[29] For individuals who participated in both shifts, the t-test of the difference in the mean of the second shift (1.85 cards/min) and the first shift (1.64 cards/min) is 0 has a p-value = 0.00

[30] For the 238 individuals who participated in both shifts, the t-test of the difference in the mean of the second shift (16.6 cards) and the first shift (20.4 cards) is 0 has a p-value = 0.00

$$ln(e_{ic}) = \alpha + \beta_f f_{ic}^{HIGH} + \beta_w w_{ic}^{HIGH} + \boldsymbol{X_{ic}\gamma} + \epsilon_c \quad (6)$$

The dependent variable $ln(e_{ic})$ is a measure of labour supply for individual i from neighbourhood cluster c. Since the amount of output produced is a count variable, I convert this measure of effort to its natural log[31] before including it in the OLS regression.[32] After conditioning on a vector of individual and shift characteristics $\boldsymbol{X_{ic}}$, the coefficients on the high fixed payment indicator f_{ic}^{HIGH} and the high wage indicator w_{ic}^{HIGH} identify if there is a relationship between expectations and labour supply.

If participants are KR agents, we'd expect effort to increase in the fixed payment and to decrease in the wage because increases in f and decreases in w both increases the threshold level of output $\bar{e}$ that satisfies $w\bar{e} = f$. Alternatively, if participants are not reference-dependent and there are no income effects, they will choose their effort independent of the fixed payment but increasing in the marginal return to effort. If there are income effects because these participants evaluate the budget generated by this experiment in isolation, then effort may decrease in expected wealth, and as such, decrease in both the fixed payment and the wage.

The Abeler et al. (2011) design is mostly closely replicated by the high wage treatment arm.[33] To replicate their analysis, I estimate:

$$ln(e_{ic}) = \alpha + \beta_f f_{ic}^{HIGH} + \boldsymbol{X_{ic}\gamma} + \epsilon_c \Big|_{w_{ic}^{HIGH}=1}. \quad (7)$$

If participants are KR agents, we expect effort to increase in the fixed payment as was reported by Abeler et al. (2011). Alternatively, if participants are not reference-dependent and there are no

[31]In the 5/366 cases where individuals produced 0 output, I have imputed an alternative output of 1 before taking the log

[32]Count variables are limited dependent variables because they are censored at 0 and not continuous. Common empirical approaches when faced with count dependent variables are to conduct OLS regressions on the natural log of the variable or to conduct a poisson or negative binomial regression (depending on the dispersion) of the raw variable. All reported results are robust to switching to either of these alternative methodologies.

[33]In the Abeler et al. (2011) experiment, students had to produce 15 units of output to make $y(e) = f$ in the low fixed payment lottery and 35 units of output to make $y(e) = f$ in the high fixed payment lottery. In my experiment, in the high wage treatment arm, participants had to produce 15 units of output to make $y(e) = f$ in the low fixed payment lottery and 30 units of output to make $y(e) = f$ in the high fixed payment lottery. The approximate amount of time to produce one unit of output was the same in both experiments.

income effects, they will choose their effort independent of the fixed payment . If there are income effects because these participants evaluate the budget generated by this experiment in isolation, then effort may decrease in expected wealth, and as such, decrease in the fixed payment.

The mechanism that links the fixed payment to effort for KR agents is through the movement of the threshold $\bar{e}$ that satisfies $w\bar{e} = f$. We expect effort to increase in the fixed payment in the above specifications because the threshold level of effort $\bar{e}$ increases in the fixed payment.[34] The smaller the piece-rate wage w, the larger a given change in the fixed payment affects the threshold level of effort $\bar{e}$.[35] To test this prediction, I estimate:

$$ln(e_{ic}) = \alpha + \beta_f f_{ic}^{HIGH} + \boldsymbol{X_{ic}\gamma} + \epsilon_c \Big|_{w_{ic}^{LOW}=0} \tag{8}$$

and compare the coefficient on the high fixed payment indicator β_f from equation (7) to equation (8). If the KR model correctly predicts behaviour, β_f from (8) will be greater than β_f from (7) and both will be greater than 0.

Finally, when an increase in wages is met with proportionate increase in the fixed payment, the threshold level of effort $\bar{e}$ that satisfies $w\bar{e} = f$ remains constant. Since the movement in $\bar{e}$ generates the prediction that effort decreases in the piece-rate wage, or $\beta_w < 0$, for KR agents in equation (6), if $\bar{e}$ is held constant as wages increase, there is no reference-dependent mechanism that could explain why $\beta_w < 0$. I estimate:

$$ln(e_{ic}) = \alpha + \beta_w w_{ic}^{HIGH} + \boldsymbol{X_{ic}\gamma} + \epsilon_c \Big|_{\bar{e}=30 \text{ units of output}} \tag{9}$$

Finally, to complement the main analysis, I plot the survival functions[36] that display the fraction of participants who continued to work after accumulating each level of piece-rate income. If participants are KR agents, we expect the probability of survival to drop sharply at $\bar{e}$, the threshold that satisfies $w\bar{e} = f$, after which there is a discontinuous drop in the marginal return to effort.

[34]That is, $\frac{d\bar{e}}{df} > 0$.

[35]That is, $\frac{d\bar{e}}{df dw} < 0$

[36]Kaplan-Meier curves that display the fraction of participants who continued to work after accumulating each level of piece-rate income.

Alternatively, if participants are not reference-dependent, there is no difference in the survival trend around this point.

to test this prediction. If participants are KR agents, we expect effort to increase in the piece-rate wage because the marginal consumption utility is positive and the only link between effort and the wages. Alternatively, if participants are not reference-dependent and there are no income effects, we should also observe effort increases in the wage. Finally, if there are income effects because these participants evaluate the budget generated by this experiment in isolation, then effort may decrease in expected wealth and, as such, effort decreases in the wage.

5.2 Second Shift Analysis

The data collected during the second shift of the experiment (all denoted with a straight prime) was used to examine the effects of the unanticipated wage shock on effort. In the main regression, I estimate:

$$ln(e'_{ic}) = \alpha + \beta_{w'} w'^{HIGH}_{ic} + \beta_w w^{HIGH}_{ic} + \boldsymbol{X'_{ic}\gamma} + \epsilon_c. \tag{10}$$

The dependent variable $ln(e'_{ic})$ is a measure of labour supply for individual i from neighbourhood cluster c from the second shift. Besides the vector of individual and interview characteristics $\boldsymbol{X'_{ic}}$, all the remaining regressors are indicator variables. The first regressor w'^{HIGH}_{ic} is an indicator that the current (second shift) wage is high. The second regressor w^{HIGH}_{ic} is an indicator that the past (first shift) wage is high. Any relationship between w^{HIGH}_{ic} and the dependent variable is consistent with expectations influence of labour supply.

If participants are KR agents, we expect that effort increases in the past piece-rate wage because high expectations of income lead to more effort. Alternatively, if participants are not reference-dependent, we expect current effort to be independent of past wages.

To identify if both positive and negative wage changes drive the relationship between past wages and current effort, I estimate:

$$ln(e'_{ic}) = \alpha + \beta_{w'} w'^{HIGH}_{ic} + \phi Neg\_Shock_{ic} + \psi Pos\_Shock_{ic} + \boldsymbol{X'_{ic}\gamma} + \epsilon_c \tag{11}$$

In this specification, I separate the effects of past wages into negative and positive wage shocks. Variable Neg_Shock_{ic} indicates that a negative wage shock has taken place because the current wage is less than the past wage. Variable Pos_Shock_{ic} indicates that a positive wage shock has taken place because the current wage is greater than the past wage. If participants are KR agents, we expect effort to increases in expectations of income. As such, we expect effort increases in a negative wage shock and decreases in a positive wage shock. Alternatively, if participants are not reference-dependent, we expect current effort to be independent of past wages.

6 Results

The relationship between the average amount of labour supplied in the first shift of the experiment and the payment contract described by Equations (6) - (9) are reported in Table 7. Estimation of Equation (6) is reported in Columns (1) and (2). I find that participants treated with the high wage do approximately 21 percent[37] less work than their low wage counterparts. I also find that the relationship between the fixed payment and effort is not statistically different from zero.[38] Although the negative relationship between the piece-rate wage and effort is consistent with KR preferences, the lack of relationship between effort and the fixed payment is not.

Estimation of Equation (7), the replication of the Abeler et al. (2011) design, is reported in Column (3). I find that participants treated with the high fixed payment do 36 percent[39] less work than their low fixed payment counterparts. To compare my fixed payment coefficient to that estimated by Abeler et al. (2011), I first impose the assumptions of constant elasticity of supply so that I can interpret both estimated coefficients as elasticities. I estimate a conditional fixed payment elasticity of labour supply equal to -0.31,[39] and Abeler et al. (2011) report a conditional fixed payment elasticity equal to +0.20[40] in the specification most analogous to my own. The

[37] $\beta_w = -0.23$ significant at 9 percent confidence; 95 percent confidence interval [- 0.49, + 0.03]

[38] $\beta_f = -0.05$ significant at 72 percent confidence; 95 percent confidence interval [-1.45, +1.35]

[39] $\beta_f = -0.31$ significant at 7 percent confidence; 95 percent confidence interval [- 0.62, + 0.008]

[40] (i) The estimate is significant at 2 percent confidence; 95 percent confidence interval [+ 0.02, + 0.38]) (ii) As is reported in Column 1 of Table 1 on p.479 of Abeler et al. (2011). (iii) The result from their unconditional specification a fixed payment elasticity of labour supply is +0.19 (significant at 3 percent confidence; 95 percent confidence interval [+ 0.005, + 0.37]) which just barely overlaps with the 95 percent confidence interval I estimate [- 0.62, + 0.008]

estimate of the fixed payment elasticity from Abeler et al. (2011) is statistically higher than what I find in my experiment.[41] The observed negative relationship between the fixed payment and effort in my experiment is inconsistent with KR preferences and the Abeler et al. (2011) results, but is consistent with agents experiencing income effects.

Estimation of Equation (8) is reported in Column (4). The estimated relationship between the fixed payment and effort in the low wage treatment arm is not statistically lower than the estimate from the high wage treatment.[42] Furthermore, the results from estimation of Equation (10) reported in Column (5) show that the wage elasticity of supply is persistently negative[43] even when threshold $\bar{e}$ is held fixed. Therefore, as such, there is no KR mechanism to explain why effort decreases in the piece-rate wage.

A snapshot of the first shift labour supply behaviour can be seen in the Kaplan-Meier survival functions plotted in Figure 3. On the vertical axis is the fraction of participants still working after earning the piece-rate income plotted along the horizontal axis. Figures 3 plots the first shift survival functions for the two fixed payment treatments separately. The first noticeable feature is that these two survival functions are not statistically different from each other,[44] reflecting that the fixed payment does not have an unconditional statistical effect on survival rates. Second, there is no sharp drop in the survival rates at the points where $we = f$ for either treatment group. Again, these patterns are not consistent with the behaviour of KR agents.

The relationship between labour supplied in the second shift of the experiment and the payment contract described by Equations (10) and (11) are reported in Table 8 for a subsample of participants.[45] I analyze the behaviour of participants whose first and second shifts were very similar – they started at the same time of the day, or on the same day of the week. The first column in Table 8 reports an estimate of the current wage elasticity of supply. Unlike the wage elasticity in

[41] The 95 percent confidence intervals do not overlap

[42] The 95 percent confidence interval of the estimate is [-0.49, +0.52]

[43] β_w = -0.167 significant at 42 percent confidence; 95 percent confidence interval [-0.56 , +0.23]

[44] A log rank test for equality of survival functions p-val = 0.681

[45] The behaviour of this subsample is more precisely estimated but is qualitatively no different from the behaviour reported for the whole sample. This subsample contains all the individuals who did their first and second shifts of work on the same weekday or whose shifts start times were within one hour of each other. If the opportunity costs of time are correlated with the day of the week or the time of the day, then the subsample whose shifts took place in the most similar contexts should have less noise in their responses than the full sample.

the first shift, the current elasticity is positive although imprecisely estimated.[46] In Column (2), I find that participants who received a raise work more than participants whose wage did not change, who work more than those who received a wage cut. I find that participants who are paid a high wage after previously experiencing a low wage do 58 percent[47] more work than those who did not experience a wage change. Those who are paid a low wage after previously experiencing a high wage do 57 percent[48] less work than those who did not experience a wage change. Columns (3) - (5) confirm that this relationship between past wages and current effort is driven by both positive and negative wage shocks. Although imprecisely estimated, negative shocks lead to weakly less effort,[49] and positive shocks lead to weakly more effort.[50] Although these results suggest that past wages influence current effort, this pattern is inconsistent with KR preferences.

7 Discussion

Neoclassical households treat income as fungible: a dollar is a dollar within the budget, no matter where is comes from.[51] But violations of this fungibility, especially in experimental contexts, is well noted. Although the consequences of a choice are rarely appreciated in isolation, a set of choices can be bracketed together more or less “narrowly” (Read et al., 2000). If experimental participants evaluate the income earned during an experiment in isolation from income generated by more traditional methods, it may explain why so few participants maximized the income generating potential of these shifts of the experiment. The financial marginal return to effort within the experiment was substantially higher than the financial marginal return to effort in alternative income generating activities. The average participant could have guaranteed to be paid at least the state mandated hourly wage with ≈22 minutes of effort in the low wage treatment arm,[52] yet only 13 of 605 observations of labour supply were top censored. This suggests that the piece-rate

[46] $\beta_{w'}$ = 0.29 significant at 34 percent confidence; 95 percent confidence interval [-0.28 , +0.86]
[47] $\beta_{w'}$ = 0.46 significant at four percent confidence; 95 percent confidence interval [-0.07 , +0.85]
[48] β_{w} = 0.45 significant at 10 percent confidence; 95 percent confidence interval [-0.94 , +0.05]
[49] ϕ = -0.52 significant at 14 percent confidence; 95 percent confidence interval [-1.18 , +0.16]
[50] ψ = 0.40 significant at 22 percent confidence; 95 percent confidence interval [-0.21 , +1.02]
[51] Adapted from Hastings and Shapiro (2013)
[52] although this number is not salient because the state mandated minimum wage is paid as a monthly salary.

workers consider the money earned in this experiment as different from that earned elsewhere, and "narrow bracket" their decision to earn income within this experiment.

The negative labour supply elasticities with respect to the wage rate and the fixed payment noted in the first shift of the experiment provide further evidence of "narrow bracketing". The income earned during the experiment represents a negligible change in total budget for a worker who is making her labour supply decisions over a long horizon and the income earning potential of the experiment is clearly transitory. As such, income effects should be small and the standard inter-temporal response to substitute towards experimental labour with its high income potential should prevail. Yet, we observe the opposite: when experimental wages are high, people work less. This pattern suggests that working in the experiment becomes more costly as the experimental income increases, something only likely if they see the budget created by the experiment as isolated from the budgets generated by their normal sources of income.

A competing explanation of the negative wage elasticity is an alternative form of reference dependence: external income targeting. External income targeting behaviour consists of participants having an income goal that is independent of treatment – for instance, to earn enough money to buy bread for dinner or medicine for a parent. In a supplementary analysis of heterogeneous treatment effects,[53] it is apparent that the negative wage elasticity found on average is driven by individuals who, ex post, report that they do not know how they will spend the money earned in this experiment.[54] If external income targeting was the reason for this behaviour, it should have been the opposite – people with a particular use for the money should have had the negative wage elasticity.

Just as how higher expected wealth reduced first shift effort, in the second shift of this experiment we again observe that expectations do influence effort but not as predicted by the KR model. Conditional on current wages, high past wages lead to less current effort. It is tempting to interpret this cross-sectional relationship as an individual's neoclassic response to wage changes – those who

[53]Not reported in this version of the paper (Nov 12, 2015), but will soon be added to an online appendix.

[54]It is an identical analysis to that which is reported in Table 7 but with additional controls for individuals reporting that they do know how they will spend their earnings and an interaction with the high wage treatment. For those who do know how they will spend their earnings, their wage elasticity is precisely 0 (i.e., the general wage elasticity and interacted elasticity are of equal magnitudes and opposite signs with at least marginal significance of p = 0.101).

experience raises update behaviour to work more, and those who experience wage cuts update to work less. Rather, in this experiment, the negative past wage elasticity captures a different feature of the effort responses – the propensity to not update behaviour and keep labour supply constant. A visual inspection of this pattern can be seen in Figure 4 which plots the difference between first and second shift output. Amongst those who do not experience a wage change, 9.4 percent of participants produced about the same amount of effort in the second shift as they did in the first.[55] Notably, amongst those who do experience a wage change, 21 percent of participants kept their effort constant, significantly more than those without wage shocks.[56] Participants were less likely to update behaviour when faced with wage changes. This suggests two things: (i) that the negative past wage elasticity captures the persistence of the effort supplied from the first shift when faced with wage shocks; (ii) that the heuristics used to determine labour supply likely evolve with experience.

In the experimental design, I had to make tradeoffs. By using the production of a foreign output rather than something more locally familiar in the experiment, I gained control over expectations but also risked that the experience was so new that participants used a labour supply heuristic that is particular to new experiences. Those who were given the same wage in the first and second shift may have gained information during the first shift they could use to update their behaviour in the second shift. In contrast, the participants who experienced wage changes between the first and second shifts again found themselves in a new income generating experience in the second. As such, they had less information to carry forward and thus many use the same heuristic again and supply the same amount of effort.

8 Conclusion

Understanding labour supply is crucial for public policy and it is a critical resource for the world's poor. To reconcile disparate evidence regarding individuals' response to wage changes, Kőszegi

[55] Plus or minus one unit of output.

[56] The marginal effect of a wage shock on the probability of $e' = e \pm 1$ is 0.11 (significant at 5 percent confidence) when estimated by probit with clustered standard errors.

and Rabin (2006, 2007, 2009) propose a model of rational expectation based reference-dependent labour supply. Their theory suggests that in addition to valuing the level of income, workers evaluate income as gains or losses with respect to their recently held probabilisitic beliefs about that income. As such, effort responds positively to beliefs about income because its marginal return drops discontinuously once accumulated income exceeds expected income.

In this paper, I test the model of rational expectation based reference-dependent labour supply in a real effort framed field experiment. Specifically, I conduct a set of experiments to test the KR model's predictions among a sample of impoverished individuals involved in piece-rate work in Northeast Brazil. In the first experiment, I replicate Abeler et al. (2011)'s design and manipulate workers' rational expectations of income with a lottery-based payment contract for an open-ended shift of work. In the second experiment, I create a gap between participants' probabilisitic beliefs about income and the payment lottery in a second shift of work using unanticipated wage shocks.

In both experiments, I find that expectations do influence effort, but not as predicted by the KR model. In the first experiment, higher expected wealth, through either dimension of the payment lottery, results in less effort – a pattern more consistent with income effects than KR preferences. In the second experiment, the inclination towards keeping labour supply constant, especially amongst those experiencing wage shocks, creates a negative relationship between previous wages and current effort. In sum, participants worked more when expectations were low than when they were high.

These results contribute to a growing literature finding that KR preferences are insufficient for explaining the labour supply behaviour of populations from the developing world. These differences highlight that the influence of expectations on effort may be context specific, and emphasize a need for caution when extrapolating from WEIRD[8] behaviours to forecast the behaviour of workers in developing countries. KR's assertion that current effort will "reflect the empirical patterns of average income" (Kőszegi and Rabin, 2006) is a tractable and informative tool for understanding individual specific labour supply. This pattern of behaviour is evident in this experiment. Nonetheless, the precise predictions of their model fail to explain how expectations influence the labour supply of these Brazilian piece-rate workers.

References

Abeler, Johannes, Armin Falk, Lorenz Goette, and David Huffman. 2011. "Reference Points and Effort Provision." *The American Economic Review*, 101(2): pp. 470–492.

Agarwal, Sumit, Mi Diao, Jessica Pan, and Tien Foo Sing. 2014. "Labor Supply Decisions of Singaporean Cab Drivers." *Available at SSRN 2338476*.

Almeida, Mansueto. 2008. "Understanding Incentives for Clustered Firms to Control Pollution: The Case of the Jeans Laundries in Toritama." 107. Ashgate Publishing, Ltd.

Altonji, Joseph G. 1986. "Intertemporal Substitution in Labor Supply: Evidence from Micro Sata." *The Journal of Political Economy*, S176–S215.

Browning, Martin, Angus Deaton, and Margaret Irish. 1985. "A Profitable Approach to Labor Supply and Commodity Demands Over the Life-cycle." *Econometrica: journal of the econometric society*, 503–543.

Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler. 1997. "Labor Supply of New York City Cabdrivers: One Day at a Time." *The Quarterly Journal of Economics*, 112(2): 407–441.

Chetty, Raj. 2012. "Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply." *Econometrica*, 80(3): 969–1018.

Chetty, Raj, Adam Guren, Dayanand S Manoli, and Andrea Weber. 2011*a*. "Does indivisible labor explain the difference between micro and macro elasticities? A meta-analysis of extensive margin elasticities." National Bureau of Economic Research.

Chetty, Raj, Adam Guren, Day Manoli, and Andrea Weber. 2011*b*. "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins." *The American Economic Review*, 101(3): 471–475.

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

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

Dupas, Pascaline, and Jonathan Robinson. 2014. "The Daily Grind: Cash Needs, Labor Supply and Self-Control." *Unpublished paper.*

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

Farber, Henry S. 2014. "Why You Can't Find a Taxi in the Rain and Other Labor Supply Lessons from Cab Drivers." National Bureau of Economic Research.

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

Gigerenzer, Gerd, and Wolfgang Gaissmaier. 2011. "Heuristic Decision Making." *Annual review of psychology*, 62: 451–482.

Harrison, Glenn W, and John A List. 2004. "Field Experiments." *Journal of Economic Literature*, 42(4): pp. 1009 – 1055.

Hastings, Justine S, and Jesse M Shapiro. 2013. "Fungibility and Consumer Choice: Evidence from Commodity Price Shocks." *The Quarterly Journal of Economics*, 128(4): 1449–1498.

Henrich, Joseph, Steven J Heine, and Ara Norenzayan. 2010. "Most People are not WEIRD." *Nature*, 466(7302): 29–29.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision Under Risk." *Econometrica: Journal of the Econometric Society*, 263–291.

Kőszegi, Botond, and Matthew Rabin. 2006. "A Model of Reference-Dependent Preferences." *The Quarterly Journal of Economics*, 121(4): pp. 1133–1165.

Kőszegi, Botond, and Matthew Rabin. 2007. "Reference-Dependent Risk Attitudes." *The American Economic Review*, 97(4): pp. 1047–1073.

Kőszegi, Botond, and Matthew Rabin. 2009. "Reference-dependent Consumption Plans." *The American Economic Review*, 99(3): 909–936.

Laisney, Francois, Winfried Pohlmeier, and Matthias Staat. 1992. *Estimation of Labour Supply Functions using Panel Data: A Survey.* Springer.

Mankiw, N Gregory, Julio J Rotemberg, and Lawrence H Summers. 1985. "Intertemporal Substitution in Macroeconomics." *The Quarterly Journal of Economics*, 100(1): 225–251.

Mulligan, Casey, et al. 1995. "The Intertemporal Substitution of Work—What Does the Evidence Say?" *University of Chicago, Population Research Center Discussion Paper Series*, 95(11).

Oettinger, Gerald S. 1999. "An Empirical Analysis of the Daily Labor Supply of Stadium Vendors." *Journal of political Economy*, 107(2): 360–392.

Pencavel, John. 1986. "Labor Supply of Men: a Survey." *Handbook of labor economics*, 1(Part 1): 3–102.

Read, Daniel, George Loewenstein, Matthew Rabin, Gideon Keren, and David Laibson. 2000. "Choice Bracketing." In *Elicitation of Preferences*. 171–202. Springer.

Tendler, Judith. 2002. "Small Firms, the Informal Sector, and the Devil's Deal." *IDS Bulletin*, 33(3).

Thaler, Richard. 1985. "Mental Accounting and Consumer Choice." *Marketing science*, 4(3): 199–214.

Tilly, Chris, Rina Agarwala, Sarah Mosoetsa, Pun Ngai, Carlos Salas, Hina Sheikh, Lucas Kerr, Marcelo Manzano, Christian Duarte, Kenneth Ng, et al. 2013. "Final Report: Informal Worker Organizing as a Strategy for Improving Subcontracted Work in the Textile and Apparel Industries of Brazil, South Africa, India and China." Institute for Research on Labor and Employment - University of California Los Angeles.

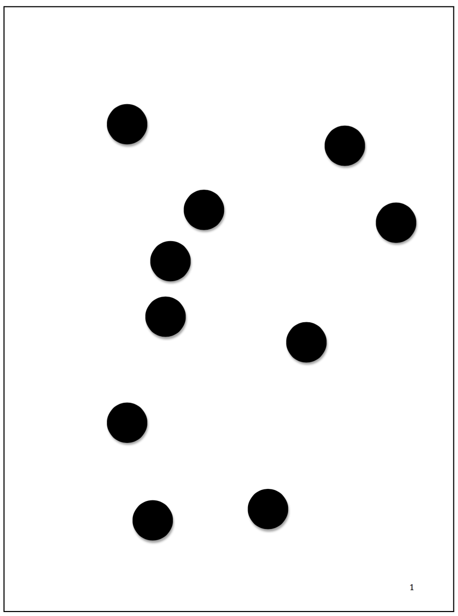
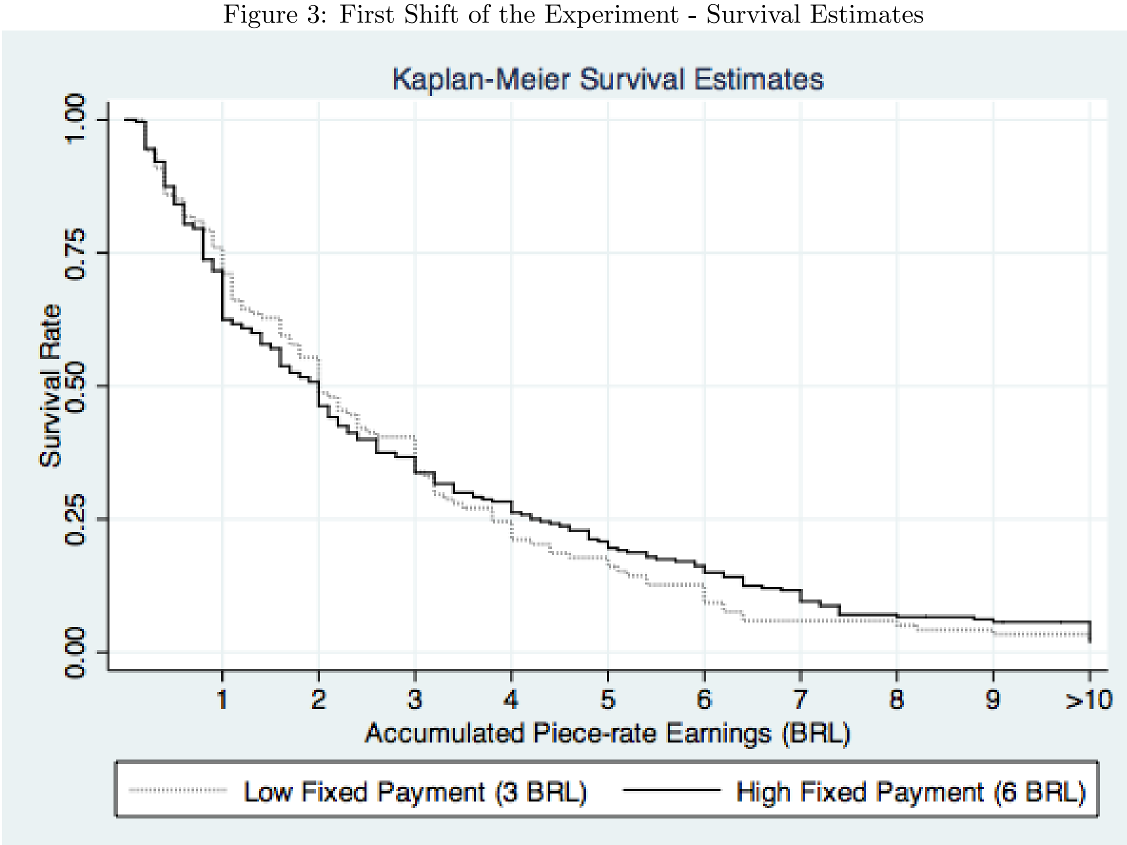


Figure 1: Card Example

Figure 2: Labelling Stickers

Figure 3: First Shift of the Experiment - Survival Estimates



Note: On the vertical axis is the fraction of participants still working after earning the piece-rate income plotted along the horizontal axis. The log rank test for equality of survival functions has a p-val = 0.681 (i.e., these lines are not different). If workers have KR preferences, there should be a sharp decrease in the survival rate where piece-rate earnings equal 3 for those assigned the low fixed payment, and similarly where piece-rate earnings equal 6 for those assigned the low fixed payment.

Figure 4: Labour Supply Consistency – Histograms of the Difference in Output Produced in the Second and First Shifts of the Experiment

Note: Bin width is 5 units of outputs. Current (second shift) output is e' and past (first shift) output is e. The left panel contains all individuals assigned the same wage in the first and second shifts. The right panel combines individuals who have both negative and positive wage shocks. Although the effect is (weakly) stronger for those experiencing negative wage shocks, both have the same pattern: effort less likely to change when a shock was experienced than when it was not. McCrary Test for Bunching p-value is 0.00 in both graphs. The marginal effect of a wage shock on the probability of $e' = e \pm 1$ is 0.11sym** when estimated by probit with clustered standard errors.

Table 1: Self-Reported Sample Characteristics of Participants

	First Shift Participants				Second Shift Participants			
	Low Wage		High Wage		Current Low Wage		Current High Wage	
	(1) Low Fixed Payment	(2) High Fixed Payment	(3) Low Fixed Payment	(4) High Fixed Payment	(5) Past Low Wage	(6) Past High Wage	(7) Past Low Wage	(8) Past High Wage
Age	41.34	45.10	43.76	44.26	43.58	45.52	46.72	42.98
	(15.83)	(18.80)	(15.86)	(15.04)	(20.33)	(14.93)	(17.05)	(15.16)
Male	0.47	0.39	0.39	0.44	0.35	0.42	0.43	0.46
	(0.50)	(0.49)	(0.49)	(0.50)	(0.48)	(0.50)	(0.50)	(0.50)
Employed	**0.50**	**0.34**	0.38	0.41	0.32	0.38	0.36	0.44
	(0.50)	(0.48)	(0.49)	(0.49)	(0.47)	(0.49)	(0.48)	(0.50)
– works from home	0.26	0.38	0.29	0.47	0.35	0.40	0.41	0.54
	(0.44)	(0.49)	(0.46)	(0.50)	(0.49)	(0.50)	(0.50)	(0.51)
– works for a piece-rate	0.52	0.41	0.35	0.38	0.39	0.39	0.42	0.36
	(0.51)	(0.50)	(0.49)	(0.49)	(0.50)	(0.50)	(0.51)	(0.49)
Subsistence Agriculture	0.13	0.18	**0.32**	**0.21**	**0.11**	**0.23**	0.26	0.19
	(0.34)	(0.39)	(0.47)	(0.41)	(0.31)	(0.43)	(0.44)	(0.39)
Minimum Salary	0.39	0.33	0.29	0.29	0.32	0.35	0.33	0.24
	(0.49)	(0.47)	(0.46)	(0.46)	(0.47)	(0.48)	(0.47)	(0.43)
Retirement Pension	0.26	0.29	0.19	0.22	0.26	0.23	0.31	0.20
	(0.44)	(0.45)	(0.39)	(0.41)	(0.44)	(0.43)	(0.47)	(0.41)
Bolsa Familia	0.15	0.13	**0.31**	**0.16**	**0.08**	**0.20**	0.18	0.12
	(0.36)	(0.33)	(0.46)	(0.37)	(0.27)	(0.40)	(0.39)	(0.33)
Observations	62	126	59	119	65	60	61	59

The left panel reviews first shift characteristics, and the right panel reviews the second shift. Reported values are the mean (sd) of each variable across the four treatment arms of each shift. All measures are self reported. Minimum Salary, Retirement Pension and Bolsa Familia all indicate that someone in the household receives these sources of income. It does not necessarily mean that the participant themselves is the one earning that income. The standard retirement pension is equal to the state mandated minimum salary. Bolsa Familia is a conditional cash transfer program Bolded values are statistically different from each other at 10 percent.

Table 2: Payment Contract Values

Treatment	Values	Notation in Empirical Methodology
Low Piece-rate Wage	0.10 BRL/output	$w^{HIGH} = 0$
High Piece-rate Wage	0.20 BRL/output	$w^{HIGH} = 1$
Low Fixed Payment	3 BRL	$f^{HIGH} = 0$
High Fixed Payment	6 BRL	$f^{HIGH} = 1$

Note: The variable w^{HIGH} is a "high wage" indicator. It equals 0 if $w = 0.10$ and 1 if $w = 0.20$. Similarly, the variable f^{HIGH} is a "high fixed payment" indicator. It equals 0 if $f = 3$ and 1 if $f = 6$. At the time of data collection, 1 BRL ≈ 0.45 USD.

Table 3: First Shift of the Experiment – Payment Contract Treatment Cells

		Piece-rate Wage w^{HIGH}	
		0	**1**
Fixed Payment f^{HIGH}	**0**	N = 62 $\bar{e} = 30$ units of output	N = 59 $\bar{e} = 15$ units of output
	1	N = 126 $\bar{e} = 60$ units of output	N = 119 $\bar{e} = 30$ units of output

Note: N is the number of participants assigned to the treatment cell; $\bar{e} = \frac{f}{w}$ is the threshold level of output where piece-rate earnings equal the fixed payment.

Table 4: Second Shift of the Experiment – Payment Contract Treatment Cells

		First Shift Piece-rate Wage w^{HIGH}	
		0	**1**
Second Shift Piece-rate Wage w'^{HIGH}	**0**	N = 61	N = 60
	1	N = 61	N = 56

Note: N is the number of participants assigned to the treatment cell. Each of these participants is from the High Fixed Payment treatment arm from the first shift. Of the 245 participants from the first shift, only 241 were located to participate in the second shift, and 3 declined to participate for reasons unrelated to this project.

Table 5: First Shift of Work – Average Output by Payment Contract Treatment Cells

		Piece-rate Wage w^{HIGH}	
		0	**1**
Fixed Payment f^{HIGH}	**0**	19.1(19.7)	19.2(15.5)
	1	21.8 (22.3)	18.8 (18.8)

Note: Reported values are the unconditional mean (standard deviation) units of output produced in each treatment cell of the first shift.

Table 6: Second Shift of Work – Average Output by Wage Shock Treatment Cells

		First Shift Piece-rate Wage w^{HIGH}	
		0	**1**
Second Shift Piece-rate Wage w'^{HIGH}	**0**	16.3 (13.2)	15.6 (13.7)
	1	16.5 (12.1)	17.1 (12.7)

Note: Reported values are the unconditional mean (standard deviation) units of output produced in each treatment cell of the second shift.

Table 7: First Shift Output as a function of the Payment Contract

			Replications†		
	Full Sample		High Wage	Low Wage	$\bar{e}$= 30 units
Dependent Variable: $\boldsymbol{ln(e_{ic})}$	(1)	(2)	(3)	(4)	(5)
High Wage Indicator (w_{ic}^{HIGH})	-0.229*	-0.227*	–	–	-0.167
	(0.132)	(0.131)			(0.201)
High Fixed Pmt Indicator (f_{ic}^{HIGH})	–	-0.0491	-0.308*	0.0462	–
		(0.134)	(0.161)	(0.242)	
Constant	1.438**	1.444**	0.851	1.891*	0.477
	(0.542)	(0.546)	(0.754)	(0.936)	(0.812)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	366	366	176	190	181
Dep. Var. Mean	2.506	2.506	2.458	2.550	2.434
Dep. Var. SD	1.077	1.077	1.093	1.063	1.104

Clustered standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
The dependent variable is the natural log of the amount of output produced in the first shift.
Each reported independent variable is an indicator (1/0) if the wage or fixed payment was high.
Because a high treatment has twice the value of the low treatment (i.e., a 100 percent increase),
and the dependent variable is logged, the coefficients can be interpreted as elasticities.
Controls: Productivity (cards per minute produced in the first activity), Age, Age-Squared,
Male, Time of Day FE, Day of Week FE, Interviewer FE, Same age as interviewer
indicator, Absolute age difference with interviewer, and Same sex as interviewer indicator.
† Replications of the Abeler et al. (2011) Experiment, most closely matched by the parameters
of the High Wage case.

Table 8: Second Shift Output as a function of Current and Past Wages

Dependent Variable: $ln(e'_{ic})$	(1)	(2)	(3)	(4)	(5)
Current Wage High ($w_{ic}^{'HIGH}$)	0.289	0.461**	0.185	0.127	0.0189
	(0.289)	(0.200)	(0.316)	(0.353)	(0.373)
Past Wage High (w_{ic}^{HIGH})	–	-0.445*	–	–	–
		(0.253)			
Current Wage < Past Wage (Neg_Shock_{ic})	–	–	-0.509	–	-0.518
			(0.342)		(0.335)
Current Wage > Past Wage (Pos_Shock_{ic})	–	–	–	0.395	0.402
				(0.321)	(0.313)
Constant	1.603**	1.518**	1.546**	1.572**	1.513**
	(0.604)	(0.606)	(0.606)	(0.609)	(0.609)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	66	66	66	66	66
Dep. Var. Mean	2.661	2.661	2.661	2.661	2.661
Dep. Var. SD	0.918	0.918	0.918	0.918	0.918

Clustered robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
The dependent variable is the natural log of the amount of output produced in the second shift. Each reported independent variable is an indicator (1/0). For both Current Wage High and Past Wage High, because high treatment has twice the value of the low treatment (i.e., a 100 percent increase), and the dependent variable is logged, the coefficients can be interpreted as elasticities. Controls: Productivity (cards per minute produced in the first activity), Age, Age-Squared, Male, Same age as interviewer indicator, Absolute age difference with interviewer, and Same sex as interviewer indicator. Estimation for the subsample whose first and second shifts took place on either the same day or had start times within ± 1 hour of each other. Results are the same pattern with the full sample, but the full sample is less precisely estimated.