

Toward an Understanding of Reference-Dependent Labor Supply: Theory and Evidence from a Field Experiment

Steffen Andersen* Alec Brandon Uri Gneezy John A. List

June 3, 2022

Abstract: We experimentally study the effects of transitory incentives and income predicted by theories of reference-dependent labor supply. Visiting an open-air market in India, we induce experimental variation in incentives by offering vendors a supplemental wage over the course of two days and in income by, one morning, administering a substantial overpayment for a good sold in the market. We find that incentives have a staggered effect on hours worked: Initially vendors do not respond, but by the second day their hours worked increases. In response to the income shock, we find that both hours worked and a measure of effort are unaffected. Collectively these findings suggest vendors supply their labor according to a model where reference dependence is moderated by experience with an incentive regime or by the time at which income is accumulated. *JEL* Codes: H2, J21, J22.

* Stefano DellaVigna, Justin Holz, Casey Mulligan, David Novgorodsky, Zongjin Qian, and George Wu provided helpful comments and suggestions. Andersen: Department of Economics, Copenhagen Business School, Porcelænshaven 16A, 1, DK-2000 Frederiksberg, Denmark (e-mail: sa.eco@cbs.dk); Brandon: Carey Business School, Johns Hopkins University, 100 International Drive, Baltimore, MD 21231 (e-mail: alec.brandon@jhu.edu); Gneezy: Rady School of Management, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093 (e-mail: ugneezy@ucsd.edu); List: Department of Economics, University of Chicago, 1126 E. 59th St., Chicago, IL 60637, and NBER (e-mail: jlist@uchicago.edu).

1. Introduction

Theories of reference-dependent preferences predict that labor market participation can respond negatively to transitory incentives (Camerer et al., 1997). Underlying this prediction are income effects that, once a worker reaches a reference level of earnings, can dominate an incentive's substitution effect. While a large collection of studies on incentive and income effects are consistent with this prediction, few of these studies are experimental. Experimental estimates of these effects can overcome concerns with observational and quasi-experimental estimates, such as endogeneity, and help advance understanding of reference-dependent preferences in labor supply.

We experimentally estimate incentive and income effects by designing, conducting, and analyzing two field experiments on the labor market participation of vendors in an open-air market in India. These vendors face a similar decision problem to the samples of bike messengers, farmers, fishermen, rickshaw drivers, taxicab drivers, salespeople, and undergraduates studied in the literature on reference-dependent labor supply. Much like these samples, the vendors in our experiments have flexible work scheduling with which they can respond to the highly variable earnings they face from one day to the next.

To estimate the effects of transitory incentives, we conducted the Market Survey Experiment with 250 vendors. The vendors in this experiment were randomly assigned to a treatment or control group. While vendors in the control group were left untouched, over the course of two consecutive days, treatment group vendors had the hourly return on their labor market participation increased by INR 10, 30, or 60. Relative to the prevailing wage, these transitory incentives were substantial. Conversations with vendors suggested that, in a typical hour, vendors earned INR 10, which indicates that our smallest wage supplement increased

hourly earnings by 100 percent. We frame the wage supplements as an incentive for hourly completion of the Market Survey, which tasked treatment vendors with keeping a tally of the number of customers visiting their shop in each hour. In response, we collected hourly observations of whether treatment and control vendors participated in the market.

We estimate income effects with 79 vendors in the Betel Nut Experiment. In this experiment, vendors selling betel nuts were selected and randomly assigned to a treatment or control group. Control group vendors were again left untouched, while one morning treatment vendors had their income altered by a westerner paying INR 500 for a pack of betel nuts that typically sells for INR 17. Such an overpayment was approximately a week's worth of earnings. We were able to administer such a large overpayment because vendors in the market did not post prices and instead haggled over prices with potential customers. In response, we collected hourly observations of whether treatment and control vendors participated in the market and a measure of the effort exerted in haggling with potential customers. To elicit this measure of effort, we hired a team of confederates and incentivized their haggling with vendors over betel nut transactions. In total, our confederates elicit the initial price offer and final price of 255 transactions with 55 of the vendors in the Betel Nut Experiment.

We motivate the design and analyses of these experiments by developing an econometric model to characterize each of our two parameters of interest: The effect of incentives on hours worked and the effect of income on the decision to quit working for the day. Observational and quasi-experimental estimates of these parameters are frequently cited as evidence of reference-dependent labor supply (see, e.g., DellaVigna 2009; O'Donoghue and Sprenger, 2018). Our econometric models formalize the link between these parameters and the variation an experiment must induce to identify these parameters. The models also formalize the threats to the

identification of our parameters of interest. We show that a mean independence assumption is needed to overcome these threats. Intuitively, this assumption places three restrictions on the non-incentive and non-income determinants of labor supply behavior: First, these labor supply determinants must be mean independent of future changes to incentives and income; second, these determinants must be mean independent of prior changes to incentives and income; and third, these determinants must be mean independent of the labor supply determinants of other vendors.

Using data generated by our experiments, we scrutinize the validity of each of these restrictions. The first restriction is evaluated by comparing labor supply outcomes of treatment and control group vendors observed before the two experiments were administered. In the Market Survey Experiment, the second restriction is avoided by omitting labor supply outcomes after the incentives ended. However, there is no such workaround for the Betel Nut Experiment. This is because our overpayment could introduce dynamic selection bias. Such a bias would arise if treatment vendors inferred that a second overpayment from a westerner could be earned if they remain active in the market. This concern motivates our elicitation of vendor effort, which is less likely than labor market participation to be directly affected by beliefs of a future overpayment. By comparing the effect of the overpayment on quitting to the effect on elicited transaction prices, the data can address whether dynamic selection bias has confounded our findings. To address the third restriction, we examine whether control vendor labor supply is altered by spillover effects induced by variation in the share of nearby vendors receiving the wage supplement in the Market Survey Experiment. When we undertake each of these exercises, we find no evidence of the data violating these restrictions.

We report two sets of findings. First, in the Market Survey Experiment, we find a staggered hours worked response to the transitory incentives. On the first day that treatment vendors received the wage supplement, we find that they reduce their hours worked, although at a level that is statistically indistinguishable from a null effect. This null response is driven by treatment vendors quitting earlier in the day. However, on the second day of the experiment, the hours worked by treatment vendors is highly responsive to the wage supplements. We find that each INR of the wage supplement increases time in the market by approximately 1 minute and this increase is highly statistically significant. Driving this effect is that treatment vendors show up earlier and quit later in the day. Additional analyses of Market Survey Experiment help rule out explanations for the staggered response to the incentives such as the credibility of the incentive varying by day and vendors needing time to plan their labor supply response to the incentives. Second, in the Betel Nut Experiment, we find that, after treatment vendors received the overpayment, they quit working at rates that are statistically indistinguishable from control vendors. We also find that, after the overpayment, treatment vendors did not exert less effort in their haggling with our confederates over transaction prices, which suggests that the quit rate response to the overpayment was not confounded by dynamic selection bias.

These findings offer contributions to two strands of research on reference dependence. The first is the experimental literature on the incentive and income effects predicted by reference-dependent models of labor supply. Most of the experimental research on reference-dependent labor supply focuses on framing effects (Abeler et al., 2011; Gill and Prowse, 2012; Hossain and List, 2012; Levitt et al., 2016; Brownback and Sadoff, 2020; Pierce et al., 2020; Fryer et al., Forthcoming; see Ferraro and Tracy, Forthcoming for an overview). Two important exceptions are Fehr and Goette (2007) and Dupas et al. (2020). Fehr and Goette (2007)

randomize a two-week long increase of 25 percent in the commission rate earned by 42 bicycle messengers. While these bicycle messengers work longer hours in response to the incentive, they also reduce their effort and, consistent with reference dependence, these responses are moderated by an elicited measure of loss aversion. Dupas et al. (2020) recruit 259 bicycle taxi drivers in Kenya to participate in periodic lotteries conducted by their research staff. They find that winning a lottery, which, on average, pays out a day's worth of earnings, has a statistically indistinguishable effect on labor supply. Relative to these studies, our findings and experimental design are distinguished by the staggered response to the incentives, the scale of the incentives and income shock, the natural framing of the income shock, and the approach used to rule out dynamic selection bias in the response to the income shock.

Second, our findings contribute to the theoretical literature on reference dependence. Understanding the relationship between labor supply, incentives, and income is crucial for understanding labor market dynamics (Lucas and Rapping, 1969). Yet conventional models of labor supply cannot explain the staggered response to incentives we find in the Market Survey Experiment. Furthermore, leading reference-dependent models of labor supply cannot explain the non-response to the income shock that we find in the Betel Nut Experiment (e.g., Köszegi and Rabin, 2006). Two alternative approaches to modeling reference-dependent labor supply are attractive. The first is motivated by experiments on experience effects in reference dependence (List, 2003; 2004; 2011; Engelmann and Hollard, 2010) and explains the staggered response in the Market Survey Experiment with reference dependence that attenuates when vendors gain experience with the wage supplements. The second is motivated by observational evidence on intraday adaption of reference levels of earnings (Thakral and Tô, 2021). Such adaptation explains the null effect in the Betel Nut Experiment as an artifact of the time at which the

overpayments were administered. Because treatment vendors received the overpayments early in their workdays, their reference earnings level could adapt.

We also offer contributions to other strands of the literature. The staggered response to incentives in the Market Survey Experiment is a novel contribution to a large body of empirical research on incentive effects in labor supply (Camerer et al., 1997; Oettinger, 1999; Chou, 2002; Goette et al., 2004; Farber, 2015; Stafford, 2015; Goldberg, 2016; Sheldon, 2016; DellaVigna et al., 2017; He et al., 2018; Chen et al., 2019; Chen et al., 2020; Richards, 2020; Angrist et al., 2021). Furthermore, the results of the Betel Nut Experiment contribute to the literature that measures daily income effects with observational research designs (Rizzo and Zeckhauser, 2003; Farber, 2005; 2015; Crawford and Meng, 2011; Nguyen and Leung, 2013; Giné et al., 2017; MacDonald and Mellizo, 2017; Ran et al., 2014; Agarwal et al., 2015; Morgul and Ozbay, 2015; Martin, 2017; Hammarlund, 2018; Schmidt, 2018; Thakral and Tô, 2021; Dodini, 2022). These studies consistently find income effects that cannot be detected in the Betel Nut Experiment.

The remainder of this paper is organized as follows. In Section 2 we formalize our strategy for identifying the effects of incentives and income on labor supply outcomes. Section 3 describes the design of our field experiments and evaluates the restrictions imposed on the data by our identification strategy. In Section 4 we present our findings and Section 5 concludes.

2. Theory

In this section, we formalize our strategy for identifying the effects of transitory incentives and income. In so doing, we highlight threats to identification and consider how data can be used to evaluate these threats.

2.1 Setting and Notation

Consider a sample of vendors who operate their own businesses in an open-air market. Each day, when the market is open, vendors are free to decide their hours of work. Let $j \in \{1, \dots, J\}$ index each vendor in the sample, $s \in \{1, \dots, S\}$ index each day, and $t \in \{1, \dots, T\}$ index time over the course of each day. Throughout, we assume that t indexes each hour of the day.

The outcomes of interest are daily hours worked and the intradaily decision to quit working for the day. Let Y_{js} denote the daily hours of work decision for vendor j on day s and let D_{jst} indicate whether vendor j quits working on day s at time t .

These outcomes may be influenced by a transitory incentive and the accumulation of daily income. Let I_{js} denote the incentive to work for vendor j on day s . Implicit in this representation is that vendors face different work incentives on different days, but the incentive to work is stable over the course of a given day. Furthermore, let M_{jst} denote the income accumulated by vendor j on day s at time t .

2.2 Parameters of Interest

Our objective is to identify two parameters of interest. These parameters are motivated by models of reference-dependent labor supply, which predict that transitory increases in the rate of hourly earnings can negatively affect hours worked and accumulated income positively affects the probability of quitting. These propositions can be modeled with the following equations,

$$Y_{js} = \alpha I_{js} + U_{js}, \quad (1)$$

where U_{js} captures residual (i.e., non-incentive) determinants of hours worked, and,

$$D_{jst} = \beta M_{jst} + V_{jst}, \quad (2)$$

where V_{jst} captures residual (i.e., non-income) determinants of quitting work for the day.

Our parameters of interest are the coefficients α and β . Relating these coefficients to the predictions of reference-dependent models of labor supply is straightforward. While conventional economic models predict $\alpha > 0$ and $\beta \approx 0$ (e.g., Lucas and Rapping, 1969), models of reference dependence predict that $\alpha < 0$ is possible and that $\beta > 0$ (Camerer et al., 1997; Farber, 2005; 2008; 2015; Köszegi and Rabin, 2006; DellaVigna, 2009; Crawford and Meng, 2011; Thakral and Tô, 2021). Some models of reference-dependent labor supply also propose specific mechanisms that underly the formation of reference earnings. Such mechanisms can be incorporated into the models in equations 1 and 2 by allowing the coefficients to vary by vendor, day, time, or other relevant features. We consider this possibility in greater detail in Section 4.3.

2.3 Assumptions and Identification

We identify our parameters of interest by experimentally varying a supplement to hourly earnings and a shock to daily income. The success of this strategy depends on the validity of mean independence assumptions, which we present formally in this subsection. We also consider the restrictions these assumptions place on vendor labor supply and strategies to empirically evaluate these restrictions.

Let W_{js} denote the magnitude of a wage supplement and X_{jst} denote the magnitude of a shock to daily income. These variables are related to incentives and income according to,

$$I_{js} = W_{js} + \tilde{I}_{js}, \quad (3)$$

where \tilde{I}_{js} denotes non-experimental work incentives, and

$$M_{jst} = X_{jst} + \tilde{M}_{jst}, \quad (4)$$

where \tilde{M}_{jst} denotes non-experimental daily income accumulation.

To identify our parameters of interest, we must invoke a mean independence assumption between the residual determinants of labor supply and an experimental wage supplement or income shock. Such an assumption to identify the effect of incentives on hours worked, α , requires that,

$$E[\tilde{U}_{js} | O_{js}, W_{j1}, \dots, W_{js}] = E[\tilde{U}_{js} | O_{js}] \text{ for } s \in \{1, \dots, S\}, \quad (5)$$

where $E[\cdot]$ is the expectations operator, $\tilde{U}_{js} = U_{js} + \alpha \tilde{I}_{js}$ captures the residual determinants of labor supply, and O_{js} is a vector of observables that controls the variation that identifies α . For example, if O_{js} includes a fixed effect for each vendor, then within vendor variation will identify α . To identify the effect of income on the decision to quit working, β , we must make a similar assumption for the income shock variable,

$$E[\tilde{V}_{jst} | O_{jst}, X_{j11}, \dots, X_{jST}] = E[\tilde{V}_{jst} | O_{jst}] \text{ for } s \in \{1, \dots, S\} \text{ and } t \in \{1, \dots, T\} \quad (6)$$

where $\tilde{V}_{jst} = V_{jst} + \beta \tilde{M}_{jst}$ captures the residual determinants of labor supply and O_{jst} is a vector of observables.

These mean independence assumptions place several restrictions on the relationship between residual determinants of labor supply and an experimental wage supplement or income shock. First, residual determinants of labor supply must be mean independent of contemporaneous and future realizations of the wage supplement and income shock. An experiment can overcome this threat by randomly assigning whether a sample of vendors receives the wage supplement and whether a sample of vendors receives the income shock. The data generated by an experiment can be used to evaluate this restriction. Under equations 5 and 6, labor supply outcomes observed prior to the receipt of a wage supplement or income shock should be balanced across recipients and non-recipients. Statistically significant differences in pre-receipt labor supply outcomes would reject this restriction.

Second, the residual determinants of labor supply must be mean independent of prior realizations of the wage supplement and income shock. This restriction can be easily avoided for the identification of incentive effects, α , by omitting labor supply outcomes observed after the wage supplements are offered. However, no such workaround is possible for the identification of income effects on the subsequent decision to quit working for the day, β . Instead, we adopt the following strategy to evaluate whether this restriction holds for the income shock. We design the administration of the income shock to minimize the risk of violating this restriction and we generate auxiliary data to evaluate the effects of income on labor supply outcomes that are less susceptible to violations of this restriction. Under equation 6, we should find consistency across our analyses of the quitting and auxiliary data. We discuss the design of the income shock's administration and auxiliary data collection in greater detail in Section 3.2.

Third, the residual determinants of labor supply must be mean independent of the wage supplement and income shock received by other vendors. While this restriction on spillovers is left implicit in equations 5 and 6, it is still necessary for the identification of our parameters of interest. To evaluate this restriction, we examine whether the vendors who form the control groups in our wage supplement experiment are influenced by the share of vendors near them who receive the supplement. Finding statistically significant differences in the hours worked by control group vendors as the share of vendors receiving the supplement increases would reject this restriction (Miguel and Kremer, 2004).

Under these restrictions, identification of our parameters of interest is straightforward. The effect of incentives, α , is identified by the coefficient δ in,

$$Y_{js} = \delta W_{js} + \kappa' O_{js} + \tilde{U}_{js}, \quad (7)$$

and the effect of income accumulation, β , is identified by the coefficient γ in,

$$D_{jst} = \gamma X_{jst} + \kappa' O_{jst} + \tilde{V}_{jst}. \quad (8)$$

Next we describe the experimental design that allows for the estimation of our parameters of interest.

3. Design

To estimate our parameters of interest, we design and conduct a transitory wage supplement and income shock experiment with 329 vendors in the Burra Bazar in Shillong, India (for more on this setting see Andersen et al., 2018).¹ The open-air market is the city's main poultry, meat, vegetable, produce, and merchandise market. Vendors in this market work as independent contractors. They are free to set their hours of work and their hourly earnings are a proportion of their hourly revenues. Prices in the market are not posted. Instead, vendors haggle with customers to determine the price of each transaction, which are made with cash. Next, we discuss the design of our wage supplement and income shock experiments in greater detail. Both discussions also provide empirical tests of the mean independence assumptions that underly our identification strategy of incentive and income effects.

3.1 Market Survey Experiment

We experimentally vary a transitory wage for 250 vendors with the Market Survey Experiment. While this experiment was conducted over the course of two sessions in 2006 and 2007 and there were slight differences in administration between the two sessions (e.g., number of days observed), this subsection describes the major features common to both sessions and a more

¹ See also Andersen et al. (2008), Gneezy et al. (2009), Andersen et al. (2011), Hoffman et al. (2011), Andersen et al. (2013), and Andersen et al. (2018) for studies conducted in this region more generally.

thorough description of each session is provided in Online Appendix A. Across both sessions, vendors in the experiment were randomly assigned to a treatment or control group. Vendors in the control group were left untouched, while treatment vendors were offered a financial incentive for each hour in which the Market Survey was completed. To complete the Market Survey in a given hour, vendors had to record a tally of customers who visited their shop and turn the survey in when research staff made their hourly visit to collect the survey. Completing this survey was intended to require minimal effort. We did not evaluate the quality of surveys that were completed and, as a result, we do not make use of the collected responses. Research staff attempted to collect these tallies from treatment vendors in each hour that the market was open. Each tally that was collected earned the vendor the incentive for that hour. Treatment vendors received one of three incentive levels for each hour that the Market Survey was completed: INR 10, INR 30, or INR 60. While no administrative data exists to calculate a baseline of hourly earnings against which these incentives can be benchmarked, conversations between research staff and vendors suggested that hourly earnings were around INR 10.

Before running the Market Survey Experiment, vendors were selected for the sample by research staff. To facilitate the administration of the experiment, vendors were selected in 14 distinct geographic clusters. Treatment and control status was randomly assigned within each cluster, or stratum, and every treatment vendor in a stratum was offered the same level of incentive. That is, treatment vendors in the same geographic cluster were never offered a different incentive. All told, 112 of the vendors were assigned to the control group, while 138 vendors were assigned to the treatment group. 60 of the treatment vendors were offered the INR 10 incentive, 51 were offered the INR 30 incentive, and 27 were offered the INR 60 incentive.

Time in the Market Survey Experiment was divided into three mutually exclusive periods: Pre-Incentive, Incentive, and Post-Incentive. During the Pre-Incentive Period research staff hourly recorded whether vendors were present in the market. At the end of the final day of the Pre-Incentive Period, treatment vendors were recruited to complete the Market Survey over two consecutive days. If a treatment vendor was absent from the market on the final day of the Pre-Incentive Period, then recruitment occurred when the vendor returned to the market. Because completion of the survey required negligible effort and the incentives were so lucrative, only two vendors declined participation. Treatment and control vendors then began a two-day long Incentive Period. During this period, research staff continued their hourly observation of whether vendors were present in the market and attempted an hourly collection of each treatment vendor's Market Survey. Treatment vendors were paid the incentive for each Market Survey collected when they quit working for the day. Then the Post-Incentive Period began, during which research staff continued hourly observation of whether vendors were present in the market.

The data observed in the Market Survey Experiment can easily be related to the empirical model in equation 7 that we seek to estimate. Hours worked, Y_{js} , is measured by summing the number of times vendor j was observed by research staff on their hourly visits on day s . The wage supplement variable, W_{js} , is 0 for all vendors during the Pre-Incentive Period. During the Incentive Period, this variable remains 0 for control vendors, whereas the variable switches to 10, 30, or 60 for treatment vendors depending on the incentive level offered for completing the Market Survey. If a treatment vendor declined the incentive offer or was absent from the market on the final day of the Pre-Incentive Period, then their wage supplement variable is still set to the incentive offered. The vector of observables, O_{js} , dictated by the design of the Market Survey

Experiment is a fixed effect for each stratum. This vector can also be augmented to determine the sources of variation that identify the effect of the wage supplement. For example, populating O_{js} with vendor fixed effects would only use within vendor variation to identify our parameter of interest.

The data observed in the Market Survey Experiment also allow us to evaluate the mean independence assumption that underlies our strategy for identifying incentive effects. In Section 2.3 we discussed three restrictions this assumption places on vendor labor supply. Table 1 presents an evaluation of the first restriction. This restriction requires that, before the wage supplements are offered, the residual determinants of labor supply are balanced across the control and treatment vendors receiving the different wage supplements. To evaluate this restriction, we regress several labor supply outcomes constructed from data observed in the Pre-Incentive Period on the incentive level offered in the Incentive Period, with fixed effects included for each stratum. Inference is conducted with clustered standard errors that are robust to heteroskedasticity and vendor autocorrelation.

Table 1 shows that the Pre-Incentive Period data does not reject the first restriction of the mean independence assumption. This is shown for the probability of a vendor working, total hours worked, start time of day, total break hours taken, and quit time of day. Across each of these outcomes, we find that the treatment levels are statistically indistinguishable from the control group at the five percent level of significance. This provides an empirical basis for the first restriction of the mean independence assumption needed to identify the labor supply effect of incentives.

Tables 1 also provides an empirical description of the vendors in the Market Survey Experiment. On average, the vendors in our sample work nearly every day that we collect data in

the market, they work nearly 8.3 hours per day, show up to the market just before 10 AM, take very few breaks, and exit the market just after 6 PM. These estimates highlight the margins over which vendors could respond to the wage supplements. Intensive margin responses were somewhat constrained, but possible, as the market opened between 6 and 8 AM and closed at 7 PM. Extensive margin responses, though, were highly constrained because vendors nearly always worked. While these features of the market likely diminished the extent to which vendors could work more in response to the incentives, they allow our analysis to speak more directly to theories of reference-dependent labor supply, as such theories are silent on the intensive margin of incentive effects.

In Table 2, we provide an evaluation of the third restriction imposed by the mean independence assumption that underlies our identification strategy.² This restriction requires that the wage supplement offered to a treatment vendor does not impact the residual determinants of control vendor labor supply. We evaluate this restriction by examining whether, during the Incentive Period, control vendors alter their labor supply as the share of vendors who are treated increases in each stratum. Such an examination is conducted by regressing total hours worked, start time of day, total break hours taken, and quit time of day on the share of vendors in a stratum receiving the treatment, with stratum fixed effects also included. We omit a test with the probability of a vendor working as the outcome because the share treated is perfectly collinear with the strata fixed effects. Inference is conducted with clustered standard errors that are robust to heteroskedasticity and vendor autocorrelation. Table 2 shows that, consistent with the mean independence assumption, the share of vendors receiving the treatment did not have a

² See Section 2.3 for a discussion of the bearing of the second restriction on the Market Survey Experiment.

statistically significant effect on the labor supply outcomes of control vendors observed during the Incentive Period.

3.2 Betel Nut Experiment

We experimentally vary a shock to the income of 79 vendors with the Betel Nut Experiment. While our intended sample size was 85 vendors, 6 were absent from the market on the morning in which the overpayment was administered. These vendors met the same eligibility criteria as the vendors in the Market Survey Experiment and one additional criterion: They all sold a non-perishable good called betel nuts. Vendors in the experiment were randomly assigned to a treatment or control group. One morning in 2006, the 20 vendors assigned to treatment had their earnings altered by a member of our research staff who paid INR 500 for a pack of betel nuts, while the 59 vendors assigned to control received no such change to their earnings. Paying INR 500 for a pack of betel nuts was a substantial overpayment, as the typical transaction for betel nuts was approximately INR 17.

This overpayment was likely to have caused treatment vendors to determine that they reached their daily level of reference income. As discussed in Section 3.2, conversations with vendors indicated that the typical hourly rate of earnings was INR 10, which suggests that, in isolation, the overpayment constituted nearly a week of earnings. Furthermore, the context in which the overpayment was received was likely to have caused treatment vendors to incorporate the income shock into their mental account of daily earnings. Unlike Dupas et al. (2020), which uses a lottery conducted by research staff to evaluate reference-dependent labor supply, the income shock in our experiment is received in the same context as their typical daily earnings. While it is still possible that vendors omit the overpayment from their mental account of daily

earnings, such an omission is difficult to reconcile with theories of mental accounting. For example, Lian (2021) models mental accounting in labor supply as a consequence of workers simplifying the complex intertemporal decision problems they face. If the vendors in our experiment assign earning from different transactions to different mental accounts then, instead of simplifying, they are adding complexity to their intertemporal labor supply decisions.

In response to the overpayment, we collected hourly observations of whether treatment and control vendors were in the market. These observations were collected the day of the overpayment and the day after the overpayment. These observations allow us to observe the data necessary to estimate the empirical model of income effects in equation 8. The decision to quit working variable, D_{jst} , is constructed by creating a vector that starts at 0 in the first hour that each vendor is observed in the market and switches to 1 until the vendor exits the market for the day. The income shock variable, X_{jst} , takes on one of two values. All vendors in the experiment start the day with the income shock variable equal to 0. Once the income shock is administered, this variable switches to 500 for treatment vendors and remains at 0 for control vendors. The data we observe also allow us to populate the vector of observables, O_{jst} , with controls for when the vendor started their day and time spent on break during the day.

These data also allow us to evaluate the first restriction imposed by the mean independence assumption required for the identification of our parameter of interest. Recall that in Section 2.3, we discussed how the first restriction of the mean independence assumption in equation 6 requires that residual labor supply determinants are mean independent of future receipt of the income shock. We can evaluate this by testing whether, on average, treatment and control group vendors started their days at different times. Column 1 of Table 3 presents the results of this test, which finds that, on average, control group vendors started their day at 9:21,

while treatment group vendors started at 9:03. Consistent with the mean independence assumption, the p -value of 0.11 on the 18 minute difference between start times is not statistically significant at traditional thresholds.

These data, however, do not allow us to evaluate the second restriction required by the mean independence assumption. This restriction holds that residual determinants of labor supply are mean independent of prior receipt of the overpayment. A major concern is that this restriction will not hold because, after receiving the overpayment, treatment vendors will perceive that there is a greater chance of receiving an additional overpayment if they remain in the market. To diminish this concern with dynamic incentives, we had the only westerner on our research staff administer the overpayments. Because the odds of receiving an additional overpayment from a westerner scale with the probability of a westerner entering the market, we hoped this design decision would limit concerns with dynamic incentives, as westerners were rarely seen in the market.

We also collected additional labor supply outcomes that are less susceptible to concerns with dynamic incentives. To do so, we hired six locals to act as confederates. These confederates, who were kept blind to vendor treatment status, were tasked with eliciting a measure of the effort exerted by vendors. To elicit such a measure, our confederates were periodically assigned a vendor and endowed with INR 30 to haggle over the price of a pack of betel nuts. Upon approaching a vendor, our confederates were instructed to first elicit an initial price offer for one of two types of betel nuts. The confederates were then free to haggle with the vendor until a final price was agreed upon. Haggling was incentivized, with the confederates earning the difference between their endowment and the agreed upon price. In total, confederates executed 258 transactions with 58 vendors in the experiment: 19 from treatment and 39 from

control. These transactions occurred on three days: Before the overpayment, the day of the overpayment (after the overpayment had been administered), and the day after the overpayment. Columns 2 and 3 of Table 3 provide a description of the data collected before the day of the overpayment.

Regressing the initial offers and final transaction prices on receipt of the overpayment allows us to estimate income effects that are less susceptible to concerns with dynamic incentives. This claim can be substantiated with a reference-dependent model of labor supply and bargaining effort where receiving the treatment increases vendor beliefs of an additional overpayment. While such a model has an ambiguous prediction on the effect of an overpayment on probability of exiting the market for the day, the prediction on effort is unambiguous. Relative to control, treatment vendors are predicted to exert less effort. These predictions are developed more formally in Online Appendix B.

3.3 Power Analysis and Pre-Analysis Plan

The Market Survey and Betel Nut Experiments were run in 2006 and 2007. As a result, our experiments pre-date the advent of norms of power and pre-analysis plans before an experiment is conducted. For example, power analyses were first formalized in the journal *Experimental Economics* in 2011 (Sadoff et al., 2011) and the AEA RCT Registry did not launch until 2012. We eschew the practice of ex-post power analyses and posting to a registry (Abrams et al., 2020) and instead encourage readers to view our analyses much like they would the analysis of any experiment run before 2011 or 2012.

4. Results

In this section, we present estimates of our parameters of interest and then discuss their implications for theories of labor supply.

4.1 Market Survey Experiment

Figure 1 illustrates the estimated effect of each wage supplement on hours worked per day, with brackets indicating the 95 percent confidence interval on each estimate. These estimates are obtained by regressing total hours worked per day in the Pre-Incentive and Incentive Periods on an indicator for each level of the wage supplement and day of the Incentive Period, with strata fixed effects also included. Starting at the top panel of Figure 1, we see that, contrary to conventional economic models and consistent with models of reference-dependent labor supply, on the first day of the Incentive Period estimated effects are negative for two of the three levels of the wage supplement and none of the estimates are statistically distinguishable from a null effect. Moving to the bottom panel, which presents the same estimates for the second day of the Incentive Period, we see a very different response. Each wage supplement has a positive effect on hours worked, the estimates are monotonic in the size of the wage supplement, and, in isolation, the effect of the largest wage supplement rejects the null of hypothesis of no labor supply response.

We next present estimates of our first parameter of interest, the effect of incentives on hours worked, in Table 4. Given the heterogeneity in response to the wage supplements over the two days of the Incentive Period, we report the incentive effect for each day. The vector of observables used in Column 1 is a fixed effect for each stratum, while in column 2 the vector of observables is a fixed effect for each vendor. Across both columns we see that each INR of the

wage supplement reduced the hours worked by vendors on the first day of the Incentive Period, although not at a level that can be statistically distinguished from a null effect, and increased hours worked by vendors on the second day of the Incentive Period. This increase in hours worked that we find on the second day of the Incentive Period is highly statistically significant regardless of the fixed effects used.

To put the magnitude of these effects into perspective, we calculate the implied intertemporal elasticity of labor supply. To do so, we assume that, in the absence of the wage supplement, vendors earn 10 INR and work 8.4 hours per day. Under these assumptions, we find that the -0.002 to -0.003 effect of each INR on hours worked on the first day of the Incentive Period implies an elasticity of -0.002 to -0.003 and the 0.013 and 0.015 effect of each INR on the second day of the Incentive Period implies an elasticity of 0.013 to 0.015. The magnitude of these elasticities is small relative to other estimates of the intertemporal elasticity of labor supply in the literature. For example, Chetty et al. (2011) reports an elasticity of 0.6 to 0.8 from a meta-analysis of the quasi-experimental literature. This divergence is likely the consequence of features specific to our setting. One such feature was the limited flexibility our vendors had to adjust on the extensive margin. Experimental studies on incentives find much larger responses on the extensive margin. For example, Goldberg (2016) finds an elasticity of 0.15 in a Malawian day labor market when extensive margin incentives are varied, and Chen et al. (2020) find that 91 percent of the effect of an experimental wage multiplier on Uber is on the extensive margin of labor supply. Yet the vendors in our setting had little flexibility on this margin. Regardless of the wage supplement, nearly every vendor worked on the days we collected data in the market. Ultimately, we focus on the relationship between the direction of our estimated effects and the

comparative static of reference-dependent models, and we discourage readers from extrapolating the specific magnitude of the implied elasticities to other settings.

Additional empirical analyses help isolate the margins of labor supply that determine these incentive effects on hours worked. Online Appendix Table C3 decomposes the effect of incentives into components attributable to the decision to work, when the day is started, break time, and when the day is ended. Therein, we find that on the first day of the Incentive Period, the incentive effect on hours worked is driven by a statistically significant effect on the time vendors ended their day, with less precisely estimated evidence of vendors also showing up to work later. Moving to the second day of the Incentive Period, the effect of the incentive on both these margins flip. Each INR of the wage supplement causes vendors to show up earlier and exit later the second day of the Incentive Period. Regardless of day, Online Appendix Table C3 also shows that incentives do not significantly alter the probability of a vendor working or the time they take on break.

In the final column of Online Appendix Table C3, we also consider the effect of our wage supplement on hours worked in the Post-Incentive Period. Therein we report that each INR of the wage supplement caused vendors to work 0.005 fewer hours per day in the Post-Incentive Period and this effect is not statistically distinguishable from a null effect. We also consider the effect of accumulated incentives on the vendor decision to quit working for the day. Online Appendix Table C4 reports that the day of the wage supplement also influenced the effect on the probability of quitting for the day. On the first day of the supplement, treatment vendors were more likely than control vendors to quit working for the day. This flips on the second day, with quitting marginally discouraged by the wage supplements. These findings are broadly consistent with the findings in Table 4 and Online Appendix Table C3.

One potential explanation for our estimated incentive effects is that, on the first day of the Incentive Period, vendors did not view the wage supplements as credible because they were not paid their earnings until the end of the first day. To examine this possibility, we exploit a quirk in the administration of the first session of the Market Survey Experiment. In this session, treatment vendors in each stratum were recruited in two waves. If the credibility of the incentives in the Market Survey Experiment was a determinant of our findings, then we should observe as a different response across the two waves because by the time the second wave started, treatment vendors in the first wave had received their payments for each Market Survey they completed. Online Appendix Table C5 shows that there was no such dynamic between the two waves. As a result, we conclude that our estimated incentive effects reflect vendor preferences and constraints, not the administration of our experiment.

Another potential explanation is that, because treatment vendors exited the market earlier on the first day of the wage supplement, they are better rested for a long day of work on the second day of the wage supplement. We test for this explanation by replicating Table 4 with a control for the hours worked by a vendor on the prior day. Online Appendix Table C6 shows that once lagged hours worked are controlled for, the same staggered response persists: Day 1 of the wage supplement has an effect on hours worked that is statistically indistinguishable from a null and on Day 2 the supplement has a highly statistically significant positive effect on hours worked.

A final potential explanation our data can address is the possibility that vendors respond to the wage supplements on the second day and not the first because of the additional time they had to plan their response. A quirk in the administration across the two sessions of the Market Survey Experiment allows us to evaluate this explanation. In the first session, treatment vendors

were informed about the wage supplements the evening before the supplements could be earned. In the second session, treatment vendors were informed two calendar days before the supplements could be earned. Online Appendix Table C7 shows that this differential timing did not influence staggered response to the wage supplements.

4.2 Betel Nut Experiment

Figure 2 plots the share of vendors in treatment and control who were active in the market on the day of the overpayment. When the overpayments were distributed around 10 AM, all treatment and control vendors were in the market. Shortly after the overpayment, five to ten percent of the treatment vendors exited the market, only to return around 4 PM. Then at 5 PM control vendors started to quit working for the day at a marginally faster rate than treatment vendors. By 7 PM all the treatment and control vendors quit working for the day.

Table 5 uses the data plotted in Figure 2 to estimate our second parameter of interest: The effect of accumulated income on the decision to quit working for the day. Columns 1 and 2 report the coefficients from a regression of the binary quitting decision on an indicator for treatment. Therein we see that the baseline quitting rate was approximately 9.7 percent and, depending on whether start time is controlled for, treatment had a -0.04 to 0.02 percentage point effect on the quit rate. These estimates are small in magnitude and are statistically indistinguishable from the null effect predicted by a conventional economic model. Columns 3 and 4 estimate equation 8 by regressing the decision to quit working for the day on a variable that measures the magnitude of the overpayment. Therein we see that each INR of the overpayment had an exceedingly small effect, in percentage points, on the baseline quitting rate.

Much like the estimates in columns 1 and 2, the estimates in columns 3 and 4 are also statistically indistinguishable from a null effect.

A basic concern with the estimates in Table 5 is that treatment vendors may remain in the market because the overpayment causes them to believe that continuing to work will net them an additional overpayment. To evaluate an outcome that is less likely to be confounded by such beliefs, we collected auxiliary data on the effort exerted by vendors in their haggling over the price of betel nuts. Table 6 reports the effect of the overpayment on these outcomes. Columns 1 and 2 show that, regardless of whether vendor fixed effects are included, vendors initial offer for a pack of betel nuts is INR 18. After the overpayment, treatment vendor offers increase by an average of INR 2 to 3. Moving to columns 3 and 4, we find similar effects of the overpayment on the final transaction price. While these effects on offers are not statistically distinguishable from a null effect, such a null is consistent with the effects reported on quitting and this consistency suggests that dynamic selection bias did not confound our estimates in Table 5.³

4.3 Implications

The data generated by our two experiments are not consistent with either model considered in Section 2.2. While a conventional model of labor supply predicts vendors will increase their hours worked on both days of the Market Survey Experiment, $\alpha > 0$, we cannot reject the null hypothesis of no response to the wage supplements on the first day of the Market Survey Experiment. Additional analyses fail to find evidence that this null first day effect can be attributed to a conventional model that also features learning, either via observation of the behavior of nearby vendors who start receiving the wage supplement a day earlier or via a longer

³ In Online Appendix Table 7, we show that auxiliary data observed the day after the overpayment is consistent with the results in Table 6.

time horizon to respond to the first day of the wage supplement. Similarly, theories of reference-dependent preferences predict that a shock to income will cause vendors to exit the market early, $\beta > 0$, but our data cannot reject the null hypothesis of no response in quit rates and elicited effort.

Allowing heterogeneity in reference dependence is an attractive approach to explaining our findings. The first potential dimension of this heterogeneity is across vendors and is described by the experience vendors have with the wage supplements in the Market Survey Experiment. Such experience effects have been found in studies on reference-dependent preferences more broadly. For example, Engelmann and Hollard (2011) find that exchange asymmetries predicted by reference-dependent preferences disappear after a 15 minute period in which undergraduates are forced to make trades in a lab experiment (see also List, 2003; 2004; 2011 for evidence on experience gained over longer time horizons). This type of heterogeneity can be incorporated into equation 1 by allowing the parameters of interest to vary by vendor,

$$Y_{js} = \alpha_j I_{js} + U_{js}, \quad (1')$$

where α_j covaries positively with a measure of vendor experience. This positive covariance between α_j and experience can explain why, on the first day of the Market Survey Experiment when vendors have no experience with the wage supplements, they do not work a longer day and why, on the second day once experience has been gained, they increase their hours worked. This positive covariance can also explain why Farber (2015) finds labor supply elasticities increase with experience and why recent studies on incentive effects with rideshare drivers do not find a staggered response (e.g., Chen et al., 2019; Chen et al.; 2020): The drivers in these samples already have a tremendous amount of experience with the incentive regime that is varied. Comparing the effect of an incentive vendors are used to receiving, such higher earnings due to a

demand shock, to an unusual incentive like the Market Survey could provide a test of this explanation.

The other dimension of heterogeneity is intertemporal. A recent observational study on New York City taxicab drivers, Thakral and Tô (2021), proposes that this type of heterogeneity can be described by the following dynamic process. Around the start of a day, workers respond to surprises by updating their reference level of income. But as a day continues, it is harder and harder for workers to adjust their reference levels of income. Incorporating this type of heterogeneity into equation 2 can be accomplished by allowing the parameters of interest to vary by time,

$$D_{jst} = \sum_{k=1}^t \beta_k M_{jsk} + V_{jst}, \quad (2')$$

where β_t grows in magnitude as time in the day, t , grows. This formulation offers an explanation for the data in the Betel Nut Experiment. Because we administer the overpayments early in the day, under equation 2', treatment vendors are able adapt their reference level of earnings and stay at work. Overpayments administered later in the day would then provide a test of the dynamic formation of reference levels of earnings.

5. Conclusion

How do labor markets respond to transitory shifts in demand? Conventional economic models predict that labor market participation responds positively to the incentives created by transitory demand shifts (Lucas and Rapping, 1969). Underlying this prediction is an assumption about the nature of income effects proposed by Friedman (1957). This assumption holds that, absent credit constraints, income effects are determined by the extent to which they alter lifetime earnings. As a result, when transitory peaks in demand create incentives that do not meaningfully alter

expected lifetime earnings, the quantity of labor supplied is predicted to increase. Twenty-five years ago, Camerer et al. (1997) argued that this prediction is not guaranteed to hold if workers have reference-dependent preferences. Under such preferences, small incentives can create big income effects that inhibit the supply of labor when demand peaks. While a large collection of studies report evidence that is consistent with this prediction, little of this evidence is experimental.

We experimentally study the effects of transitory incentives and income on the labor market participation of 329 vendors in an open-air market in India. The effects we estimate fail to conform with the predictions of either model of labor supply. In the Market Survey Experiment, we find a staggered response to high-stakes incentives. On the first day of the wage supplement, consistent with reference dependence, treatment vendors reduce their hours worked. While this effect is statistically indistinguishable from a null effect, such an effect can be explained by reference dependence but not a conventional model of labor supply. By the second day of the wage supplement, though, treatment vendors behave according to a conventional labor supply model. Higher levels of incentives on this day lead to more hours of work and this effect is highly statistically significant. In the Betel Nut Experiment, we find a null effect on labor market participation in response to a high-stakes overpayment received to treatment vendors. We also fail to find evidence that this null response is driven by treatment vendors remaining in the market in hopes of earning an additional overpayment. While this null response is consistent with the predictions of a conventional labor supply model, theories of reference-dependent preferences predict otherwise.

These findings can be explained by two variants of reference-dependent labor supply. The first has vendors assign a large weight to their reference-dependent preferences when first

responding to an incentive arrangement. Once experience with an incentive arrangement is gained, though, vendors diminish the weight placed on reference dependence and behave according to their conventional economic desires. The second variant allows for intraday updating of reference levels of earnings (Thakral and Tô, 2021). Under this theory, our overpayment did not influence treatment vendors because it was conducted in the morning, which allowed vendors to update their earnings referents. A useful exercise for future research is to experimentally introduce the variation suggested by these competing explanations into the experimental designs developed in this study.

References

- Abeler, Johannes, Armin Falk, Lorenz Goette, and David Huffman, "Reference Points and Effort Provision," *American Economic Review*, 101 (2), 470–492, 2011.
- Abrams, Eliot, Jonathan Libgober, and John A. List, "Research Registries: Facts, Myths, and Possible Improvements," NBER Working Paper No. 27250, 2020
- Agarwal, Sumit, Mi Diao, Jessica Pan, and Tien Foo Sing, "Are Singaporean Cabdrivers Target Earners?" Working Paper, 2015.
- Andersen, Steffen, Erwin Bulte, Uri Gneezy, and John A. List, "Do Women Supply More Public Goods Than Men? Preliminary Experimental Evidence from Matrilineal and Patriarchal Societies," *American Economic Review*, 98(2): 376-81, 2008.

Andersen, Steffen, Seda Ertac, Uri Gneezy, Moshe Hoffman, and John A. List, "Stakes Matter in Ultimatum Games," *American Economic Review*, 101(7): 3427-39, 2011

Steffen Andersen, Seda Ertac, Uri Gneezy, John A. List, Sandra Maximiano, "Gender, Competitiveness, and Socialization at a Young Age: Evidence From a Matrilineal and a Patriarchal Society," *Review of Economics and Statistics*, 95(4): 1438–1443, 2013.

Andersen, Steffen, Seda Ertac, Uri Gneezy, John A. List, and Sandra Maximiano, "On the Cultural Basis of Gender Differences in Negotiation," *Experimental Economics*, 21: 757-778, 2018.

Angrist, Joshua D., Sydnee Caldwell, and Jonathan V. Hall, "Uber versus Taxi: A Driver's Eye View," *American Economic Journal: Applied Economics*, 13(2): 272–308, 2021

Brownback, Andy and Sally Sadoff, "Improving College Instruction through Incentives," *Journal of Political Economy*, 128(8): 2925-2972, 2020.

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

Chen, M. Keith, Peter E. Rossi, Judith A. Chevalier, and Emily Oehlsen, "The Value of Flexible Work: Evidence from Uber Drivers," *Journal of Political Economy*, 127 (6), 2735–2794, 2019.

Chen, Kuan-Ming, Claire Ding, John A. List, and Magne Mogstad, "Reservation Wages and Workers' Valuation of Job Flexibility: Evidence from a Natural Field Experiment," NBER Working Paper No. 27807, 2020.

Chetty, Raj, Adam Guren, Day Manoli, and Andrea Weber, "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins," *American Economic Review*, 101(3): 471-75, 2011.

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

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

DellaVigna, Stefano, "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature*, 47 (2), 315-72, 2009.

DellaVigna, Stefano, Attila Linder, Balazs Reizer, and Johannes F. Schmider, "Reference-Dependent Job Search: Evidence from Hungary," *Quarterly Journal of Economics*, 132(4): 1969-2018, 2017.

Dodini, Samuel, "Making Reference-Dependent Preferences: Evidence from Door-to-Door Sales," Working Paper, 2022.

Dupas, Pascaline, Jonathan Robinson, and Santiago Saavedra, "The Daily Grind: Cash Needs and Labor Supply," *Journal of Economic Behavior and Organization*, 177:399-414, 2020.

Engelmann, Dirk and Guillaume Hollard, "Reconsidering the Effect of Market Experience on the 'Endowment Effect'," *Econometrica*, 78(6):2005-2019, 2010.

Farber, Henry S., "Is Tomorrow Another Day? The Labor Supply of New York City Cab-drivers," *Journal of Political Economy*, 113(1), 46–82, 2005.

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

Farber, Henry S., "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, 2015.

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

Ferraro, Paul J. and J. Dustin Tracy, "A reassessment of the potential for loss-framed incentive contracts to increase productivity: a meta-analysis and a real-effort experiment," *Experimental Economics*, Forthcoming.

Friedman, Milton, *A Theory of the Consumption Function*, Princeton University Press, 1957

Fryer, Roland G., Steven D. Levitt, John List, and Sally Sadoff, "Enhancing the Efficacy of Teacher Incentives through Loss Aversion: A Field Experiment," *American Economic Journal: Economic Policy*, Forthcoming.

Gill, David, and Victoria Prowse, "A Structural Analysis of Disappointment Aversion in a Real Effort Competition." *American Economic Review*, 102(1):469-503, 2012.

Giné, Xavier, Monica Martinez-Bravo, and Marian Vidal-Fernández, "Are labor supply decisions consistent with neoclassical preferences? Evidence from Indian boat owners," *Journal of Economic Behavior and Organization*, 142:331–347, 2017.

Gneezy, Uri, Kenneth L. Leonard, and John A. List, "Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society," *Econometrica*, 77(5):1637-1664, 2009.

Goette, Lorenz, David Huffman, and Ernst Fehr, "Loss Aversion and Labor Supply," *Journal of the European Economic Association*, 2(2-3):216–228, 2004.

Goldberg, Jessica, “Kwacha Gonna Do? Experimental Evidence about Labor Supply in Rural Malawi,” *American Economic Journal: Applied Economics*, 8(1):129-49, 2016.

Hammarlund, Cecilia, “A trip to reach the target? – The labor supply of Swedish Baltic cod fishermen,” *Journal of Behavioral and Experimental Economics*, 75:1–11, 2018.

He, Shu, Liangfei Qiu, and Xusen Cheng, “Wage Elasticity of Labor Supply in Real-Time Ridesharing Markets: An Empirical Analysis,” Working Paper, 2018.

Hoffman, Moshe, Uri Gneezy, and John A. List, "Nurture affects gender differences in spatial abilities," *Proceedings of the National Academy of Arts and Sciences*, 108(36):14786-14788, 2011.

Hossain, Tanjim and John A. List, “The Behavioralist Visits The Factory: Increasing Productivity Using Simple Framing Manipulations,” *Management Science*, 58(12): 21-51, 2012.

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

Levitt, Steven D., John A. List, Susanne Neckermann, Sally Sadoff, "The Behavioralist Goes to School: Leveraging Behavioral Economics to Improve Educational Performance," *American Economic Journal: Economic Policy*, 8(4):183-219, 2016.

Lian, Chen, "A Theory of Narrow Thinking," *Review of Economic Studies*, 88(5): 2344-2374, 2021.

List, John A., "Does Market Experience Eliminate Market Anomalies?" *Quarterly Journal of Economics*, 118(1):41-72, 2003.

List, John A., "Neoclassical Theory Versus Prospect Theory: Evidence from the Marketplace," *Econometrica*, 72(2):615-625, 2004.

List, John A., "Does Market Experience Eliminate Market Anomalies? The Case of Exogenous Market Experience," *American Economic Review*, 101 (3): 313-17, 2011.

Lucas Jr., Robert E. and Leonard A. Rapping, "Real Wages, Employment, and Inflation," *Journal of Political Economy*, 77(5): 721-754, 1969.

MacDonald, Daniel and Philip Mellizo, "Reference dependent preferences and labor supply in historical perspective," *Journal of Behavioral and Experimental Economics*, 69(C): 117-124, 2017

Martin, Vincent, "When to quit: Narrow bracketing and reference dependence in taxi drivers," *Journal of Economic Behavior and Organization*, 144:166–187, 2017.

Miguel, Edward and Michael Kremer, "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica*, 72(1):159-217, 2004.

Morgul, Ender Faruk and Kaan Ozbay, "Revisiting Labor Supply of New York City Taxi Drivers: Empirical Evidence from Large-Scale Taxi Data," Transportation Research Board 94th Annual Meeting, 2015.

Nguyen, Quang and Pingsun Leung, "Revenue targeting in fisheries," *Environment and Development Economics*, 18(5):559–575, 2013.

O'Donoghue, Ted and Charles Sprenger, "Reference-Dependent Preferences," *Handbook of Behavioral Economics: Applications and Foundations*, 1, 1–77, 2018.

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

Pierce, Lamar, Alex Rees-Jones, and Charlotte Blank, "The Negative Consequences of Loss-Framed Performance Incentives," NBER Working Paper No. 26619, 2021.

Ran, Tao, Walter R. Keithly, and Chengyan Yue, "Reference-Dependent Preferences in Gulf of Mexico Shrimpers' Fishing Effort Decision," *Journal of Agricultural and Resource Economics*, 39(1):19–33, 2014.

Richards, Timothy J., “Income Targeting and Farm Labor Supply,” *American Journal of Agricultural Economics*, 102(2):419–438, 2020.

Rizzo, John A. and Richard J. Zeckhauser, "Reference Incomes, Loss Aversion, and Physician Behavior," *Review of Economics and Statistics* 85(4):909–922, 2003.

List, John A., Sally Sadoff, and Mathis Wagner, “So you want to run an experiment, now what? Some simple rules of thumb for optimal experimental design,” *Experimental Economics*, 14(4):439-457, 2011.

Schmidt, Marc-Antoine, “The Daily Labor Supply Response to Worker-Specific Earnings Shocks,” Working Paper, 2018

Sheldon, Michael, “Income Targeting and the Ridesharing Market,” Working Paper, 2016.

Stafford, Tess M., “What Do Fishermen Tell Us That Taxi Drivers Do Not? An Empirical Investigation of Labor Supply,” *Journal of Labor Economics*, 33(3):683–710, 2015.

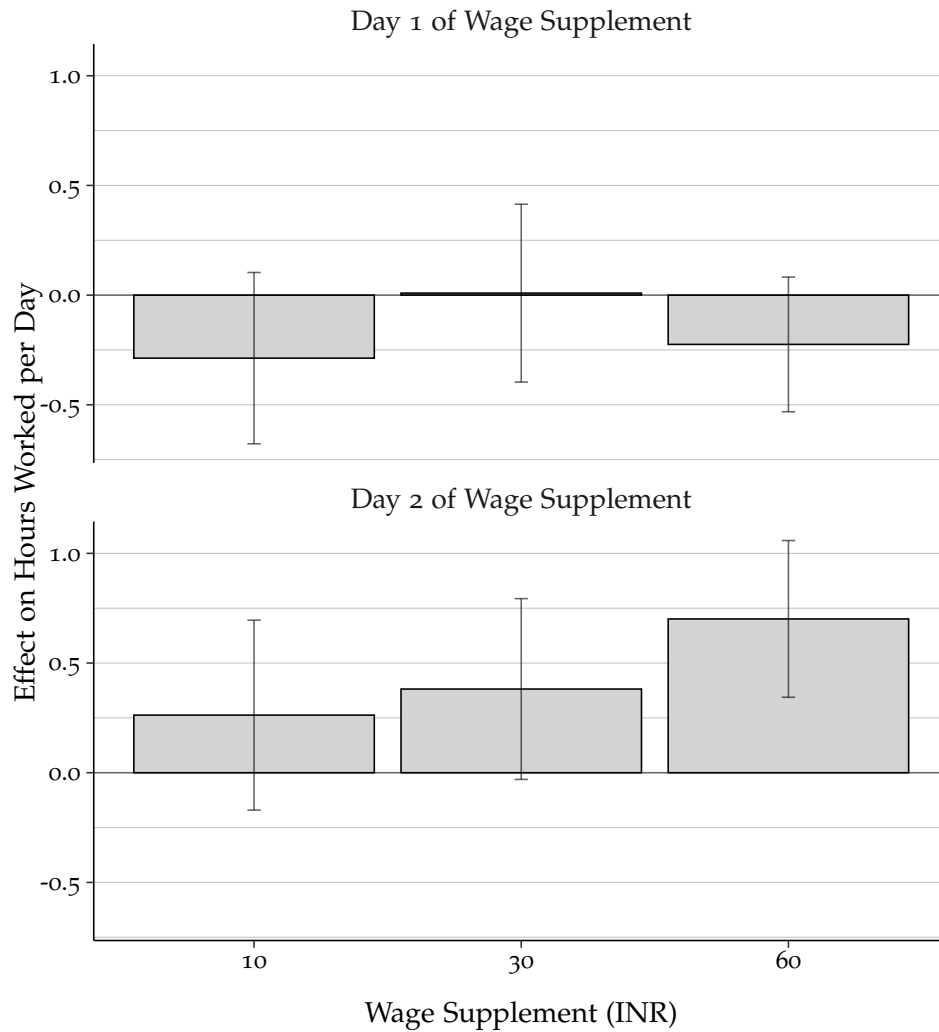
Thakral, Neil and Linh T. Tô, “Daily Labor Supply and Adaptive Reference Points,” *American Economic Review*, 111(8):2417–2443, 2021.

Table 1: Pre-Incentive Period Balance in Market Survey Experiment

| | (1) | (2) | (3) | (4) | (5) |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| | Worked | Hours Worked | Start Time | Break Hours | Quit Time |
| Treatment Diff. | | | | | |
| INR 10 | 0.08 (0.05) | 0.42 (0.51) | 0.24 (0.20) | -0.05 (0.05) | -0.04 (0.13) |
| INR 30 | -0.02 (0.02) | -0.39 (0.21) | 0.30 (0.17) | -0.08 (0.05) | 0.00 (0.08) |
| INR 60 | 0.00 (0.00) | 0.29 (0.24) | -0.29 (0.24) | 0.00 (0.00) | 0.00 (0.00) |
| Control | 0.97 (0.02) | 8.28 (0.18) | 9.66 (0.09) | 0.05 (0.02) | 18.28 (0.05) |
| Vendors | 250 | 250 | 245 | 245 | 245 |

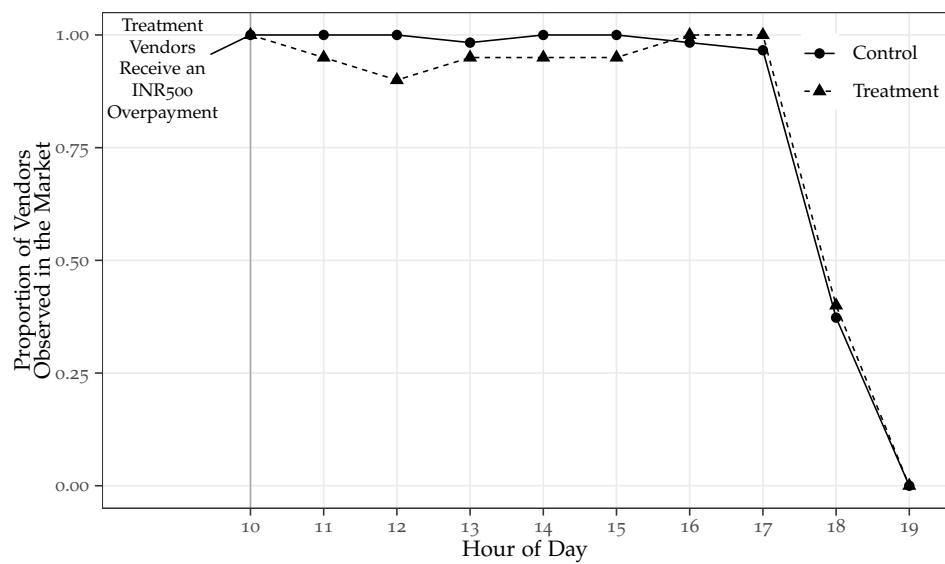
Note: This table reports the coefficients from a regression of labor supply outcomes on the incentive offered to treatment group vendors and stratum fixed effects. The control group average is reported by taking the weighted average of control vendor outcomes across each stratum. The regression is fit with outcomes that are observed during the first day of the Pre-Incentive Period, during which treatment group vendors have yet to be notified of the incentive they will be offered during the Incentive Period. Clustered standard errors that are robust to heteroskedasticity and vendor autocorrelation are reported in parentheses under each coefficient.

Figure 1: Effect of Incentives in Market Survey Experiment



Note: This figure reports the effect of each wage supplement and day of the wage supplement on hours worked. Effects are obtained by regressing daily hours worked on an indicator for each day and level of the wage supplement in the Market Survey Experiment, with a fixed effect for each strata also included. This regression is run on data observed in the Pre-Incentive and Incentive Periods. Brackets indicate 95 percent confidence intervals, which are constructed with clustered standard errors that are robust to heteroskedasticity and vendor autocorrelation.

Figure 2: Share of Vendors in Market After Overpayment Administered in Betel Nut Experiment



Note: This figure plots the proportion of vendors in the treatment and control group who were observed in the market after the overpayment was administered.

Table 2: Incentive Period Spillovers on Control Group Vendors in Market Survey Experiment

| | (1) Hours Worked | (2) Start Time | (3) Break Hours | (4) Quit Time |
|---------------|------------------------|----------------------|-----------------------|---------------------|
| Share Treated | -0.07 (0.49) | -0.53 (0.42) | 0.45 (0.25) | -0.15 (0.28) |
| Constant | 8.31 (0.20) | 9.87 (0.13) | -0.08 (0.06) | 18.33 (0.08) |
| Vendors | 112 | 109 | 109 | 109 |
| Observations | 224 | 218 | 218 | 218 |

Note: This table reports the coefficients from a regression of control vendor labor supply outcomes on the share of vendors receiving the treatment in a stratum and stratum fixed effects. The reported constant is calculated by taking the weighted average across each stratum of the predicted outcome when the share of vendors receiving the treatment is 0. The regression is fit with outcomes that are observed during the first two days of the Incentive Period, during which treatment group vendors receive a wage supplement for each hour they complete the Market Survey. See Online Appendix A for more on the timing of the Incentive Period during the two sessions of the Market Survey Experiment. Clustered standard errors that are robust to heteroskedasticity and vendor autocorrelation are reported in parentheses under each coefficient.

Table 3: Pre-Income Shock Balance in Betel Nut Experiment

| | (1) Start Time | (2) Initial Offer | (3) Final Price |
|---------------------|--------------------------|--------------------------|--------------------------|
| Control | 9.29 (0.09) | 17.87 (0.91) | 17.16 (0.91) |
| Treatment | 9.05 (0.15) | 15.95 (0.99) | 15.13 (0.85) |
| Control = Treatment | $p\text{-value} = 0.182$ | $p\text{-value} = 0.093$ | $p\text{-value} = 0.067$ |
| Vendors | 79 | 52 | 52 |

Note: This table reports the average outcomes for treatment and control group vendors prior to the income shock's administration. The outcomes considered are the time vendors arrived to the market on the day of the income shock and the initial price offer and final price elicited by confederates before the income shock. Average initial and final prices control for the type of betel nut. The p -value reports a test of equality between treatment and control group averages. This inference is conducted with clustered standard errors that are robust to heteroskedasticity and vendor autocorrelation.

Table 4: Effect of Incentives on Hours Worked in Market Survey Experiment

| | (1) | (2) | (3) | (4) |
|--------------------------|-------------------|-------------------|------------------|-------------------|
| | Hours | Hours | Hours | Hours |
| | Worked | Worked | Worked | Worked |
| Wage Supplement on Day 1 | -0.003 (0.003) | -0.002 (0.003) | | |
| Wage Supplement on Day 2 | 0.013 (0.003) | 0.015 (0.003) | | |
| Wage Supplement | | | 0.005 (0.003) | 0.007 (0.003) |
| Constant | 8.443 (0.087) | 8.427 (0.026) | 8.443 (0.087) | 8.427 (0.026') |
| Fixed Effects | Strata | Vendor | Strata | Vendor |
| Vendors | 250 | 250 | 250 | 250 |
| Observations | 820 | 820 | 820 | 820 |

Note: This table reports the coefficients from a regression of hours worked per day on the magnitude of the wage supplement for the first and second day of the Market Survey Experiment. This regression is estimated on data from the Pre-Incentive and Incentive Periods. Column 1 reports the coefficients after adjusting for strata fixed effects. Column 2 reports the coefficients after adjusting for vendor fixed effects. The reported constant is calculated by taking the weighted average across each stratum or vendor of the predicted outcome when the wage supplement is 0. Clustered standard errors that are robust to heteroskedasticity and vendor autocorrelation are reported in parentheses under each coefficient.

Table 5: Effect of Income on Decision to Quit Working in Betel Nut Experiment

| | (1) | (2) | (3) | (4) |
|---------------|-----------------------|----------------------|-----------------------|----------------------|
| | Quit (pp) | Quit (pp) | Quit (pp) | Quit (pp) |
| Treatment | -0.04092 (0.11973) | 0.01957 (0.12182) | | |
| Overpayment | | | -0.00008 (0.00024) | 0.00004 (0.00024) |
| Constant | 9.65630 (0.06271) | 9.64094 (0.06099) | 9.65630 (0.06271) | 9.64094 (0.06099) |
| Fixed Effects | | Start Time | | Start Time |
| Vendors | 79 | 79 | 79 | 79 |
| Observations | 819 | 819 | 819 | 819 |

Note: This table reports the coefficients from a regression of the binary decision to quit working for the day on the overpayment in the Betel Nut Experiment. Columns 1 and 2 report the effect of assignment to receive the overpayment and columns 3 and 4 report the effect of the magnitude of the overpayment. Columns 1 and 3 feature no controls, while columns 2 and 4 include a fixed effect for the time of day when the vendor started working. The reported constant in columns 2 and 4 are calculated by taking the weighted average across each start time. Clustered standard errors that are robust to heteroskedasticity and vendor autocorrelation are reported in parentheses under each coefficient.

Table 6: Effect of Income on Betel Nut Haggling in Betel Nut Experiment

| | (1) Initial Price | (2) Initial Price | (3) Final Price | (4) Final Price |
|--------------------------------|-------------------------|-------------------------|-----------------------|-----------------------|
| Treatment | -1.44 (1.10) | | -1.93 (1.04) | |
| Day of Overpayment | 0.75 (1.08) | -0.17 (1.12) | 0.49 (1.01) | 0.17 (1.09) |
| Treatment X Day of Overpayment | 1.67 (1.49) | 2.53 (1.50) | 1.83 (1.39) | 2.22 (1.44) |
| Constant | 18.29 (0.68) | 18.10 (0.52) | 17.85 (0.66) | 17.20 (0.50) |
| Fixed Effects | Betel Type | Betel Type, Vendors | Betel Type | Betel Type, Vendor |
| Vendors | 55 | 55 | 55 | 55 |
| Observations | 155 | 155 | 155 | 155 |

Note: This table reports effects on the initial and final prices of betel nut transactions. Three effects are reported: The effect of assignment to treatment, the effect of the day on which treatment was administered, and the effect of receiving the treatment on the day the treatment was administered. Columns 1 and 2 report these effects on initial price offered by a vendor and columns 3 and 4 report these effects on the final transaction price. Columns 1 through 4 control for the type of betel nut that was haggled over. Columns 2 and 4 also control for vendor. Estimated are obtained with data observed on a day prior to the overpayment and the day of the overpayment. Clustered standard errors that are robust to heteroskedasticity and vendor autocorrelation are reported in parentheses under each coefficient.

Online Appendix for “Towards an Understanding of Reference-Dependent Labor Supply: Theory and Evidence from a Field Experiment”

Online Appendix A

In this Appendix we describe the administration of the Market Survey Experiment in greater detail. This experiment was conducted over the course of two sessions. In the subsequent two sections, we describe the administration of each session.

1. Session 1

Session 1 of the Market Survey Experiment was conducted over the course of six days in Spring 2006. A total of 90 vendors were identified that met the following criteria: Vendors working for a single-employee firm with a permanent location in the market selling non-perishables. The 90 vendors were divided equally across six strata determined by the location and the goods sold in that area. Randomization was conducted within each stratum, with vendors assigned to one of three arms: Control, Treatment Wave 1, and Treatment Wave 2. Control vendors were untouched over the course of the experiment while vendors in Treatment Wave 1 and Wave 2 received an expected hourly wage supplement of INR 10 or INR 30 for two consecutive days. To get a sense of the magnitude of the wage supplements, an informal survey of the market suggested baseline hourly earnings of INR 10. Online Appendix Table 1 summarizes the count of vendors in each strata and arm of the experiment.

Online Appendix Figure 1 illustrates the timeline of vendors in the experiment. On Day 1 all vendors were untouched until the end of the day, when only Treatment Wave 1 vendors were approached by research staff and offered a fixed supplemental wage for each hour they

completed the Market Survey on Days 2 and 3. On Days 2 and 3, the research staff visited vendors in Treatment Wave 1 every hour to collect that hour's Market Survey. When vendors in Treatment Wave 1 were ready to leave the market on both Day 2 and 3 they were paid their promised wage for each hour a survey had been collected that day. On Days 4-6 vendors in Treatment Wave 1 were untouched. Vendors in Treatment Wave 2 experienced the same timeline as vendors in Treatment Wave 1 lagged by a day. Vendors in Control were untouched over the course of the experiment.

These days are linked to the periods of the experiment that are described in Section 3.1 in the following way. Vendors in Treatment Wave 1 were in the Pre-Incentive Period on Day 1, the Incentive Period on Days 2 and 3, and the Post-Incentive Period on Days 4 through 6. Vendors in Treatment Wave 2 were in the Pre-Incentive Period on Days 1 and 2, the Incentive Period on Days 3 and 4, and the Post-Incentive Period on Days 5 and 6. Control vendors were in the Pre-Incentive Period on Day 1, the Incentive Period on Days 2 through 4, and the Post-Incentive Period on Days 5 and 6.

A quirk of the Post-Incentive Period in Session 1 was that hours worked was not observed for two calendar days. These two days were "Market Days," wherein the market was simply too busy for us to reliably observe the labor market participation of vendors.

Online Appendix Figure 1 also presents the daily mean, between vendor standard deviation, and within vendor standard deviation of hours worked, probability of working, start time, and quit time for vendors in Control. Importantly, there is sizeable between and within variation in labor supply behavior, which supports the claim that vendors were free to choose their hours of work. Additionally, the average start time at 9:30am is well after the market opening at 8:00am and the average quit time of 6:20pm is before the market's closing time at

7:00pm, ensuring that vendors have two margins to increase their supply. Online Appendix Table 2 evaluates the balance of hours worked separately for each stratum and day of the Pre-Incentive Period.

2. Session 2

Session 2 of the Market Survey Experiment was conducted over the course of four days in Fall 2007. At this point, the market opened at 6 instead of 8 AM. Online Appendix Table 3 summarizes the count of vendors in each stratum and arm of the experiment. A total of 160 vendors were identified according to the same criteria used in Session 1. Vendors were divided equally across eight strata. Randomization was again conducted within each stratum, with vendors assigned to one of two arms: Control and Treatment. Control vendors were again untouched over the course of the experiment while vendors in Treatment received an expected hourly wage supplement of INR 10, INR 30, or INR 60 over two consecutive days. Unlike Session 1, not all strata received both between and within variation in expected wage. Furthermore, two strata were purely control and received no supplemental wage. They are labeled INR 0 in Online Appendix Table 2.

Online Appendix Figure 2 presents the timeline of Session 2 as well as descriptive statistics of the vendors in Control. The timeline shows that vendors assigned to Treatment experience the same timeline as vendors in Treatment Wave 1 but with only one day of post-treatment observation. A quirk of Session 2 was that hours worked were not observed for two calendar days in between the Pre-Incentive and Incentive Periods. These two days were “Market Days,” wherein the market was too busy for us to reliably observe the labor market participation of vendors. The descriptive statistics in Online Appendix Figure 2 show that vendors in Session

2 behave similarly to vendors in Session 1, with similar levels and variation in behavior. Online Appendix Table 4 evaluates the balance of hours worked separately for each stratum and day of the Pre-Incentive Period.

3. Market Days

As discussed in the two subsections above, Market Days occurred just as the Incentive Period ended in Session 1 and just before the Incentive Period began in Session 2. On these Market Days we did not observe hours worked. To determine whether Market Days influenced our findings, we separately estimate the effect of the incentives in each session. Online Appendix Table C7 reports these estimates. Therein we find that regardless of the timing of the Market Days, the same staggered response to the incentives is found. We interpret this consistency in the staggered response as evidence that the timing of Market Days did not confound our results.

Online Appendix B

In this appendix we formalize a reference-dependent model of labor supply where vendors determine their hours of work and effort in bargaining. Suppose customers approach a vendor according to a Poisson arrival process. Once there is a meeting, the vendor and customer engage in alternating price offers, with the transaction price increasing in the vendor's effort and the vendor paying a convex cost for effort exerted. Each day, vendors solve,

$$\begin{aligned} \max_{Y, \{e_\tau\}} & \int_0^Y \pi e^{-(\pi+\rho)\tau} \tilde{v}(e_\tau) d\tau + u(M) - v(Y) + G(M) - \lambda(1 - G(M)), \\ \text{s. t. } & M = X + \sum_\tau p_\tau, \\ & p_\tau = f(e_\tau), \end{aligned} \tag{B1}$$

where Y is hours worked, e_τ is effort exerted on transaction τ , π is the Poisson arrival rate of customer transactions, ρ is the discount rate, \tilde{v} maps effort into disutility, u is a function that

maps accumulated daily income into utility, M is accumulated daily income, v is a function that maps hours worked into disutility, G is a CDF of the continuum of reference levels of income as in Koszegi and Rabin (2006), λ captures the intensity of loss aversion, X is the overpayment, p_τ is the price of transaction τ , and f maps effort on a transaction into the price of the transaction. Solving and taking the total derivative with respect to X reveals that the price of transaction τ is decreasing in the overpayment. Concerns with dynamic selection bias can be seen by incorporating into the model in equation B1 an arrival process for western customers that is increasing in the prior receipt of an overpayment. Under such a dual customer formulation, the hours worked effect of the overpayment is indeterminate, but the effect on transaction prices with our confederates remains decreasing in the overpayment.

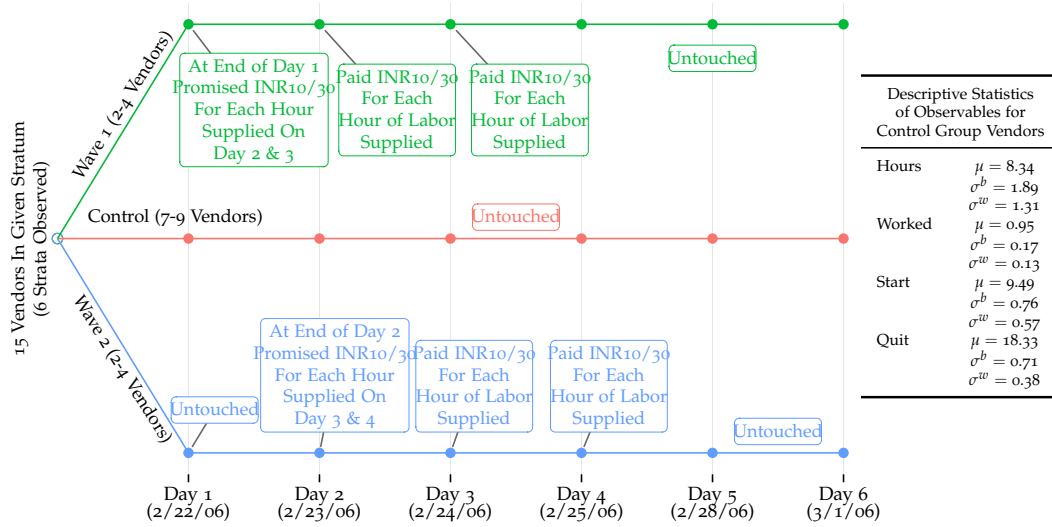
Online Appendix C

This appendix reports evidence on the robustness of our findings.

References

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

Online Appendix Figure A1: Market Survey Experiment—Session 1: Experimental Design



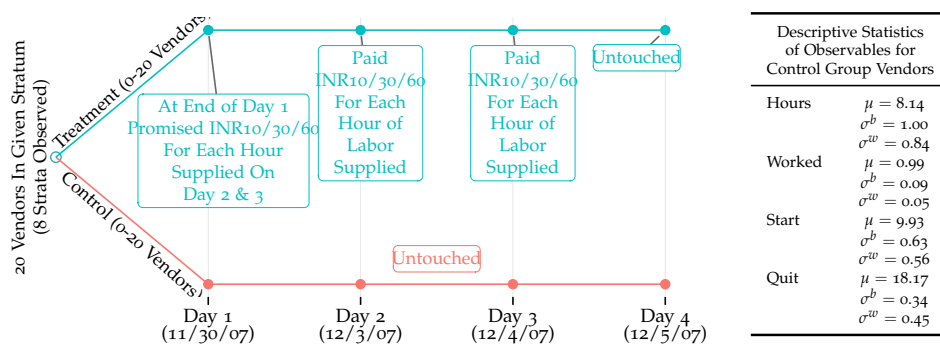
Note: Timeline of subjects and descriptive statistics are presented to summarize Session 1 of the Market Survey Experiment. A total of 90 vendors are divided into six equal sized strata. Expected supplemental wage is constant within stratum, with half of strata assigned to INR10 and other half to INR30. The right panel presents the labor supply behavior of vendors in Control observed over the six days of the experiment. Hours measures hours of labor supply per day, Worked reflects the proportion of vendors supplying a positive quantity of labor per day, Start measures the time a vendor arrives in the market in decimal military time, and Quit captures the time a vendor leaves the market in decimal military time. μ measures vendor average, σ^b measures the standard deviation between vendors, and σ^w measures the standard deviation within vendor. There is no data on 2/26/2006 and 2/27/2006 because those were “Market Days” where the market was so busy that it was infeasible to conduct the Market Survey Experiment.

Online Appendix Table A1: Market Survey Experiment—Session 1: Vendors per Stratum

| | Control | Treatment Wave 1 | Treatment Wave 2 |
|-----------|---------|------------------|------------------|
| INR 10 | | | |
| Stratum 1 | 8 | 3 | 4 |
| Stratum 2 | 9 | 2 | 4 |
| Stratum 3 | 8 | 3 | 4 |
| INR 30 | | | |
| Stratum 1 | 8 | 4 | 3 |
| Stratum 2 | 7 | 4 | 4 |
| Stratum 3 | 9 | 4 | 2 |

Note: Count of vendors per stratum and treatment arm.

Online Appendix Figure A2: Market Survey Experiment—Session 2: Experimental Design



Note: Timeline of subjects and descriptive statistics are presented to summarize Session 2 of the Market Survey Experiment. A total of 160 vendors are divided into eight equal sized strata. Expected supplemental wage is constant within stratum, with one-quarter assigned to INR 10, one-quarter to INR 30, one-quarter to INR 60, and one-quarter to pure control. The right panel presents the labor supply behavior of vendors in Control observed over the four days of the experiment. Hours measures hours of labor supply per day, Worked reflects the proportion of vendors supplying a positive quantity of labor per day, Start measures the time a vendor arrives in the market in decimal military time, and Quit captures the time a vendor leaves the market in decimal military time. μ measures vendor average, σ^b measures the standard deviation between vendors, and σ^{iw} measures the standard deviation within vendor. There is no data on 12/1/2007 and 12/2/2007 because those were “Market Days” where the market was so busy that it was infeasible to conduct the Market Survey Experiment.

Online Appendix Table A2: Market Survey Experiment—Session 1: Balance of Hours Worked in Pre-Incentive Period

| | Pre-Treatment Day 1 | | | Pre-Treatment Day 2 | |
|----------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
| | Control | Treatment Wave 1 | Treatment Wave 2 | Control | Treatment Wave 2 |
| INR 10: Stratum 1 | 7.38 (3.16) $n = 8$ | 8.33 (1.15) $n = 3$ | 8.50 (0.58) $n = 4$ | 8.88 (0.99) $n = 8$ | 8.50 (0.58) $n = 4$ |
| | | $p = 0.65$ | | $p = 0.43$ | |
| INR 10: Stratum 2 | 6.89 (2.85) $n = 9$ | 7.00 (0.00) $n = 2$ | 7.75 (0.50) $n = 4$ | 7.44 (3.21) $n = 9$ | 8.50 (1.00) $n = 4$ |
| | | $p = 0.03$ | | $p = 0.40$ | |
| INR 10: Stratum 3 | 9.13 (0.99) $n = 8$ | 8.67 (0.58) $n = 3$ | 8.75 (0.50) $n = 4$ | 9.50 (1.20) $n = 8$ | 9.75 (0.50) $n = 4$ |
| | | $p = 0.61$ | | $p = 0.62$ | |
| INR 30: Stratum 1 | 9.50 (0.53) $n = 8$ | 8.75 (0.96) $n = 4$ | 9.67 (0.58) $n = 3$ | 8.88 (3.60) $n = 8$ | 9.33 (0.58) $n = 3$ |
| | | $p = 0.28$ | | $p = 0.74$ | |
| INR 30: Stratum 2 | 9.00 (0.82) $n = 7$ | 7.50 (1.29) $n = 4$ | 7.75 (0.96) $n = 4$ | 8.57 (0.53) $n = 7$ | 7.75 (1.26) $n = 4$ |
| | | $p = 0.05$ | | $p = 0.23$ | |
| INR 30: Stratum 3 | 8.11 (0.60) $n = 9$ | 8.00 (0.00) $n = 4$ | 8.00 (0.00) $n = 2$ | 7.56 (0.73) $n = 9$ | 7.00 (0.00) $n = 2$ |
| | | $p = 0.61$ | | $p = 0.06$ | |

Note: Average hours worked by treatment arm and day of the pre-incentive period is reported with standard deviations in parentheses and sample sizes reported below. A p -value from test of equality is also reported.

Online Appendix Table A3: Market Survey Experiment—Session 2: Vendors per Strata

| | Control | Treatment |
|-----------|---------|-----------|
| INR 10 | | |
| Stratum 1 | 0 | 20 |
| Stratum 2 | 0 | 20 |
| INR 30 | | |
| Stratum 1 | 4 | 16 |
| Stratum 2 | 6 | 14 |
| INR 60 | | |
| Stratum 1 | 13 | 7 |
| Stratum 2 | 0 | 20 |
| INR 0 | | |
| Stratum 1 | 20 | 0 |
| Stratum 2 | 20 | 0 |

Note: Count of vendors per strata/treatment arm.

Online Appendix Table A4: Market Survey Experiment—Session 2: Balance of Hours Worked in Pre-Incentive Period

| | | Pre-Treatment Day | | |
|-----------|----------|-------------------|-----------|------------|
| | | Control | Treatment | |
| INR 10: | | | 9.70 | |
| Stratum 1 | | | (0.92) | |
| | | | $n = 20$ | |
| INR 10: | | | 10.20 | |
| Stratum 2 | | | (1.01) | |
| | | | $n = 20$ | |
| INR 30: | 7.75 | 7.44 | | $p = 0.70$ |
| Stratum 1 | (0.50) | (3.03) | | |
| | $n = 4$ | $n = 16$ | | |
| INR 30: | 8.00 | 8.14 | | $p = 0.63$ |
| Stratum 2 | (0.63) | (0.53) | | |
| | $n = 6$ | $n = 14$ | | |
| INR 60: | 8.00 | 8.29 | | $p = 0.25$ |
| Stratum 1 | (0.58) | (0.49) | | |
| | $n = 13$ | $n = 7$ | | |
| INR 60: | | 8.40 | | |
| Stratum 2 | | (0.50) | | |
| | | $n = 20$ | | |
| INR 0: | 7.65 | | | |
| Stratum 1 | (0.81) | | | |
| | $n = 20$ | | | |
| INR 0: | 7.30 | | | |
| Stratum 2 | (1.92) | | | |
| | $n = 20$ | | | |

Note: Average hours worked by treatment arm and day of the pre-incentive period is reported with standard deviations in parentheses and sample sizes reported below. A p -value from test of equality is also reported.

Online Appendix Table C1: Robustness of Pre-Incentive Period Balance in Market Survey Experiment

| | (1) | (2) | (3) | (4) | (5) |
|-----------------|----------------|-----------------|-----------------|-----------------|-----------------|
| | Worked | Hours Worked | Start Time | Break Hours | Quit Time |
| Treatment Diff. | | | | | |
| INR 10 | 0.06 (0.04) | 0.24 (0.41) | 0.20 (0.21) | 0.00 (0.03) | -0.03 (0.11) |
| INR 30 | 0.00 (0.02) | -0.33 (0.23) | 0.45 (0.15) | -0.05 (0.04) | 0.02 (0.08) |
| INR 60 | 0.00 (0.00) | 0.29 (0.24) | -0.29 (0.24) | 0.00 (0.00) | 0.00 (0.00) |
| Control | 0.97 (0.02) | 8.35 (0.18) | 9.62 (0.08) | 0.04 (0.01) | 18.30 (0.04) |
| Vendors | 250 | 250 | 246 | 246 | 246 |
| Observations | 340 | 340 | 333 | 333 | 333 |

Note: This table reports the coefficients from a regression of labor supply outcomes on the incentive offered to treatment group vendors and stratum fixed effects. The control group average is reported by taking the weighted average of control vendor outcomes across each stratum. The regression is fit with outcomes that are observed during the first and, when relevant, second day of the Pre-Incentive Period, during which treatment group vendors have yet to be notified of the incentive they will be offered during the Incentive Period. See Online Appendix A for more on the timing of the Pre-Incentive Period in the two sessions of the Market Survey Experiment. Clustered standard errors that are robust to heteroskedasticity and vendor autocorrelation are reported in parentheses under each coefficient.

Online Appendix Table C2: Robustness of Incentive Period Spillovers on Control Group Vendors in Market Survey Experiment

| | (1) Hours Worked | (2) Start Time | (3) Break Hours | (4) Quit Time |
|---------------|------------------------|----------------------|-----------------------|---------------------|
| Share Treated | -0.74 (0.65) | -0.18 (0.32) | 0.32 (0.23) | -0.02 (0.24) |
| Constant | 8.50 (0.18) | 9.73 (0.11) | -0.03 (0.06) | 18.30 (0.07) |
| Vendors | 112 | 110 | 110 | 110 |
| Observations | 273 | 265 | 265 | 265 |

Note: This table reports the coefficients from a regression of control vendor labor supply outcomes on the share of vendors receiving the treatment in a stratum and stratum fixed effects. The reported constant is calculated by taking the weighted average across each stratum of the predicted outcome when the share of vendors receiving the treatment is 0. The regression is fit with outcomes that are observed during the first two days and, when relevant, the third day of the Incentive Period, during which treatment group vendors receive a wage supplement for each hour they complete the Market Survey. See Online Appendix A for more on the timing of the Incentive Period during the two sessions of the Market Survey Experiment. Clustered standard errors that are robust to heteroskedasticity and vendor autocorrelation are reported in parentheses under each coefficient.

Online Appendix Table C3: Effect of Incentives on Alternative Outcomes in Market Survey Experiment

| | (1) | (2) | (3) | (4) | (5) Post Hours Worked |
|--------------------------|------------------|-------------------|-------------------|-------------------|--------------------------------|
| | Worked | Start Time | Break Hours | Quit Time | |
| Wage Supplement on Day 1 | 0.000 (0.000) | 0.003 (0.002) | -0.001 (0.000) | -0.003 (0.001) | |
| Wage Supplement on Day 2 | 0.000 (0.000) | -0.008 (0.003) | 0.000 (0.001) | 0.003 (0.001) | |
| Wage Supplement | | | | | -0.005 (0.009) |
| Constant | 0.979 (0.008) | 9.667 (0.045) | 0.032 (0.010) | 18.327 (0.023) | 8.513 (0.190) |
| Vendors | 250 | 247 | 247 | 247 | 250 |
| Observations | 820 | 804 | 804 | 804 | 360 |
| R^2 | 0.04 | 0.17 | 0.15 | 0.48 | 0.22 |

Note: This table reports the coefficients from a regression of an outcome of interest on the magnitude of the wage supplement in the Market Survey Experiment. The outcomes in Columns 1 through 4 are respectively the decision to work, start time, break hours, and quit time. For these outcomes, the effect of the wage supplement is considered for the first and second day of the Market Survey Experiment. These effects are estimated with data from the Pre-Incentive and Incentive Periods. Column 5 considers hours worked in the Post-Incentive Period. This regression is estimated on data from Post-Incentive Period. All five regressions control for strata fixed effects. The reported constant is calculated by taking the weighted average across each stratum of the predicted outcome when the wage supplement is 0. Clustered standard errors that are robust to heteroskedasticity and vendor autocorrelation are reported in parentheses under each coefficient.

Online Appendix Table C4: Effect of Incentives on Hazard Rate of Quitting in Market Survey Experiment

| | (1) | (2) |
|--------------------------|-------------------|-------------------|
| | Quit (pp) | Quit (pp) |
| Wage Supplement on Day 1 | 0.007 (0.003) | 0.007 (0.003) |
| Wage Supplement on Day 2 | -0.003 (0.003) | -0.003 (0.003) |
| Fixed Effects | Strata | Vendor |
| Vendors | 247 | 247 |
| Observations | 8,325 | 8,325 |

Note: This table reports the coefficients from a regression of the decision to quit working for the day on the wage supplement. The wage supplement is interacted with the day of the wage supplement. Controls are strata or vendor fixed effects. The regression is estimated on data observed in the Pre-Incentive and Incentive Periods. Clustered standard errors that are robust to heteroskedasticity and vendor autocorrelation are reported in parentheses under each coefficient.

Online Appendix Table C5: Effect of Wave on Initial Response to Incentives in Market Survey Experiment

| | (1) Hours Worked | (2) Hours Worked |
|-------------------------------------|------------------------------|------------------------------|
| Wage Supplement on Day 1 for Wave 1 | -0.009 (0.010) | 0.004 (0.009) |
| Wage Supplement on Day 1 for Wave 2 | 0.003 (0.013) | 0.006 (0.011) |
| Wave 1 = Wave 2 Fixed Effects | p -value = 0.392 Strata | p -value = 0.878 Vendor |
| Vendors | 90 | 90 |
| Observations | 299 | 299 |

Note: This table reports the coefficients from a regression of hours worked on the magnitude of the wage supplement in the first session of the Market Survey Experiment. The magnitude of the wage supplement is interacted with whether a treatment vendor was in the first or second wave of the wage supplement's rollout. Estimates are obtained with data from the Pre-Incentive Period and the first day of the Incentive Period. Treatment vendors in the first wave began receiving the wage supplement one day earlier than vendors in the second wave. For more on the administration of these waves, see Online Appendix A. Estimates are reported when strata and vendor fixed effects are used. The p -value testing the null hypothesis that the incentive effect for wave 1 equals the incentive effect for wave 2 is also reported. Clustered standard errors that are robust to heteroskedasticity and vendor autocorrelation are reported in parentheses under each estimate.

Online Appendix Table C6: Effect of Incentives on Hours Worked When Controlling for Lagged Hours Worked in Market Survey Experiment

| | (1) Hours Worked | (2) Hours Worked |
|--------------------------|-----------------------------|-----------------------------|
| Wage Supplement on Day 1 | -0.004 (0.004) | 0.010 (0.008) |
| Wage Supplement on Day 2 | 0.013 (0.003) | 0.025 (0.008) |
| Controls | Lagged Hours, Strata FEs | Lagged Hours, Vendor FEs |
| Vendors | 250 | 250 |
| Observations | 570 | 570 |

Note: This table reports the coefficients from a regression of hours worked per day on the magnitude of the wage supplement for the first and second day of the Market Survey Experiment. This regression is estimated on data from the Pre-Incentive and Incentive Periods. Column 1 reports the coefficients after adjusting for lagged hours worked and strata fixed effects. Column 2 reports the coefficients after adjusting for lagged hours worked and vendor fixed effects. Clustered standard errors that are robust to heteroskedasticity and vendor autocorrelation are reported in parentheses under each coefficient.

Online Appendix Table C7: Effect of Incentives on Hours Worked By Session of Market Survey Experiment

| | (1) Hours Worked Session 1 | (2) Hours Worked Session 1 | (3) Hours Worked Session 2 | (4) Hours Worked Session 2 |
|--------------------------|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| Wage Supplement on Day 1 | -0.003 (0.009) | 0.006 (0.007) | -0.003 (0.003) | -0.003 (0.003) |
| Wage Supplement on Day 2 | 0.014 (0.009) | 0.023 (0.007) | 0.012 (0.003) | 0.013 (0.004) |
| Fixed Effects | Strata | Vendor | Strata | Vendor |
| Vendors | 90 | 90 | 160 | 160 |
| Observations | 340 | 340 | 480 | 480 |

Note: This table reports the coefficients from a regression of hours worked per day on the magnitude of the wage supplement for the first and second day of each session of the Market Survey Experiment. This regression is estimated on data from the Pre-Incentive and Incentive Periods. Columns 1 and 2 report the coefficients from the first session of the Market Survey Experiment, while columns 3 and 4 report the coefficients from the second session. Columns 1 and 3 report the coefficients after adjusting for strata fixed effects. Columns 2 and 4 report the coefficients after adjusting for vendor fixed effects. Clustered standard errors that are robust to heteroskedasticity and vendor autocorrelation are reported in parentheses under each coefficient.

Online Appendix Table C8: Effect of Income on Betel Nut Haggling Day After Overpayment in Betel Nut Experiment

| | (1) Initial Price | (2) Initial Price | (3) Final Price | (4) Final Price |
|--------------------------------------|-------------------------|-------------------------|-----------------------|-----------------------|
| Treatment | -1.97 (1.13) | | -2.37 (1.06) | |
| After Day of Overpayment | 2.92 (0.80) | 2.49 (0.91) | 2.03 (0.71) | 1.99 (0.86) |
| Treatment X After Day of Overpayment | 1.10 (1.24) | 1.85 (1.33) | 2.01 (1.23) | 2.18 (1.32) |
| Constant | 18.39 (0.65) | 17.71 (0.43) | 17.94 (0.61) | 16.99 (0.42) |
| Fixed Effects | | Vendor | Vendor | Vendor |
| Vendors | 57 | 57 | 57 | 57 |
| Observations | 150 | 150 | 150 | 150 |

Note: This table reports effects on the initial and final prices of betel nut transactions. Three effects are reported: The effect of assignment to treatment, the effect of the day on which treatment was administered, and the effect of receiving the treatment on the day the treatment was administered. Columns 1 and 2 report these effects on initial price offered by a vendor and columns 3 and 4 report these effects on the final transaction price. Columns 1 through 4 control for the type of betel nut that was haggled over. Columns 2 and 4 also control for vendor. Estimated are obtained with data observed on a day prior to the overpayment and the day after the overpayment. Clustered standard errors that are robust to heteroskedasticity and vendor autocorrelation are reported in parentheses under each coefficient.