Can income targeting be explained by dynamic sample selection?*

Pedro Bessone [†] João Pugliese [‡]

This version: July 2021

Abstract

In neoclassical models, transitory income shocks should not affect labor supply. This prediction has often been rejected empirically in favor of theories featuring reference-dependent preferences. We show that apparent negative daily income effects can be generated in a neoclassical model of labor supply by dynamic selection, where wage variation causes differential attrition throughout workers' shifts. Using data from an RCT with experimental variation in wages and fine measures of labor supply, we show that estimates of negative income effects are an artifact of dynamic selection in this setting, providing a neoclassical explanation to the findings of the income-targeting literature. JEL codes: C90, D90, J22

*We would like to thank Gautam Rao, Frank Schilbach, Heather Schofield, and Mattie Toma for allowing us to use the data from Bessone et al. (2021) (AEARCTR-0002494) in our paper. We would also like to thank Nicolas Ajzenman, Joshua Angrist, Abhijit Banerjee, Edward Davenport, Bruno Ferman, Lisa Ho, Deivy Houeix, Guilherme Lichand, Jeremy Majerovitz, Jacob Moscona, Charlie Rafkin, Gautam Rao, Tobias Salz, Karthik Sastry, Frank Schilbach, Dmitry Taubinsky, Sammuel Young, and participants of the Development Tea and Behavioral Lab at MIT and the Behavioral Lunch at MIT and Stanford for helpful comments. João is grateful to the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil (CAPES) for financial support.

[†]Uber; pedrobtepedino@gmail.com

[‡]Stanford University; jfpugli@stanford.edu

In neoclassical models of labor supply, workers consider their lifetime wealth when deciding how much labor to provide. This implies that transitory income shocks should not impact labor supply meaningfully. Starting with Camerer et al. (1997), a series of prominent papers have challenged this basic principle by showing a negative relation between daily income earned and labor supply in different settings. This empirical pattern is widely interpreted as evidence of income targeting, a theory in which workers set a daily earnings goal and choose their labor supply in order to reach it. Under this theory, a positive shock to daily earnings brings workers closer to their target, thereby reducing hours worked in the day.

In this paper, we propose and test the relevance of an alternative, neoclassical explanation for the negative relation between same-day earnings and labor supply estimated by hazard models.[1] We elucidate the relation between earnings and labor supply in a model of daily labor supply, in which workers face an optimal-stopping problem to decide when to end their shift. We allow for a general utility function that includes current wages, hours worked, and a term capturing unobserved (by the econometrician) heterogeneity. We allow the heterogeneity to vary across workers and shifts, capturing changes in workers' labor supply costs due to, for example, tiredness, malnutrition (Schofield, 2014), or financial distress (Kaur et al., 2019). By design, our model rules out income effects, which excludes the possibility of workers engaging in income-targeting strategies.

Our model generates two key insights. First, even without income effects, labor supply is negatively related to earnings under general conditions due to *dynamic selection* (Diamond and Hausman, 1984). It suffices for utility to be increasing in contemporaneous wages and weakly decreasing in hours worked, and there is some unobserved heterogeneity affecting labor supply. To understand this result, consider a worker at two different periods of a shift, in two states of the world: a low- and a high-energy state, where the worker has, respectively, a high and a low cost of effort. To estimate daily income effects, at $t = 2$ we compare the hazard rate (i.e., the probability of ending the shift) between workers who received, respectively, a high and a low wage at $t = 1$.

Consider a situation where the worker faces a low wage at $t = 1$. In that scenario, she may decide to keep working *only* in the high-energy state because the low financial incentives do not offset her high cost of effort in the low-energy state. In contrast, if she faces a high wage at $t = 1$, the high financial incentives may incentivize her to extend her shift even in the

[1] Researchers often estimate income targeting using hazard models (e.g., Farber, 2005). In this approach, income effects are identified by regressing a dummy of the decision to stop working for the day on cumulative daily earnings using intra-day data on labor supply and earnings. An alternative approach consists of exploring wage variation across days, regressing hours worked on daily wages (e.g., Fehr and Goette, 2007, Angrist et al., 2017). In this approach, a negative elasticity of labor supply with respect to wages is interpreted as evidence of income targeting. However, this approach cannot disprove expectation-based theories of income targeting Kőszegi and Rabin (2006).

low-energy state. In that case, at $t = 2$, all low-wage workers *who keep working* are in the high-energy state, while high-wage workers have a mix of high- and low-energy states. Thus, workers who receive a low wage at $t = 1$ and "survive" are positively selected. As such, they are less likely to stop working at $t = 2$, which implies a negative correlation between daily earnings and labor supply despite the nonexistence of income-targeting strategies.

Second, our model predicts that dynamic selection can generate a pattern where the relation between labor supply and earnings is stronger for more recent earnings, although this is not granted to occur. Thakral and Tô (2020) find such a pattern in the context of cab drivers and rationalize it with a theory of adaptive reference-point formation. In their theory, workers gradually adjust their reference point to account for variations in income accumulated earlier in their shift. Thus, earnings accumulated early in the shift help form the workers' income target to a larger degree than earnings accumulated later on. This reduces the impact of less recent earnings on labor supply in comparison to more recent ones. Our model can provide an alternative, neoclassical explanation to such a pattern.

We investigate the empirical relevance of dynamic selection using rich data from an RCT with low-income workers in Chennai, India (Bessone et al., 2021). Four hundred and fifty two participants worked for up to 28 days in a full-time data-entry job. They had considerable discretion over their labor supply, deciding their time of arrival and departure, when to take breaks, and how much effort to exert at their work, making the work in our setting similar to other jobs with flexible working hours. The data set contains precise measures of labor supply, captured by keystroke-by-keystroke data. Moreover, most of participants' earnings depended on an observable piece rate that fluctuated exogenously between a low and a high value throughout the day. This allows us to investigate income effects without the concern that wages might be endogenous.

We first follow the literature and estimate income effects using a standard hazard model, which does not address dynamic selection. We find a strong, negative relation between labor supply and cumulative earnings, consistent with both income targeting and dynamic selection. Workers randomly assigned to high piece rates earlier in the shift had a higher hazard rate, exerted less effort, and spent more time in work breaks. The results are precisely estimated and larger than estimates in other settings (e.g., Farber, 2015, Thakral and Tô, 2020). Similarly to Thakral and Tô (2020), we find that income effects are largely concentrated in more recent income variation, consistent with the theory of adaptive reference-point formation.

Next, we consider the importance of dynamic selection in driving the empirical patterns in our setting. We apply a nonparametric approach that eliminates dynamic selection. After accounting for dynamic selection, we find no evidence of negative income effects on hazard rates, effort, or work breaks. Importantly, we are well powered to reject the null of no daily income effects. Thus, in our setting, controlling for dynamic selection leads to the rejection of income-targeting and adaptive reference-point formation models.

We corroborate this finding with two additional pieces of evidence. First, the participants were randomly assigned to short shifts for a few days, in which they received a large incentive to stop working exactly at 5:00 pm. This effectively shut down the extensive margin of labor supply, a necessary condition for dynamic selection. Importantly, workers could still adjust labor supply via effort provision or work breaks, allowing us to test for income targeting in both dimensions. We find no evidence of income targeting in either. Second, in the Appendix we explore random variation in payments for a different study task. This task was mandatory, and the participants only learned about the payment level moments before the task. Thus, they could not end the shift in response to their payment level before or during the task, which could lead to dynamic selection. Further corroborating the dynamic selection hypothesis, we find that the payments do not affect labor supply for the rest of the day.

Our paper makes two contributions to the income-targeting literature (Camerer et al., 1997, Farber, 2005, 2008, 2015, Crawford and Meng, 2011, Chang and Gross, 2014, Andersen et al., 2018, Hammarlund, 2018, Dupas et al., 2020). First, we explore the existence of income targeting in a novel setting using randomized intra-day wage variation to address endogeneity concerns present in most previous work.[2] Moreover, our unusually rich data allow us to consider additional dimensions of labor supply such as effort and work breaks.

Second, we show the importance of dynamic selection when estimating income targeting using a novel empirical strategy. Our model shows that dynamic selection occurs under weak conditions, likely present in most settings. Dynamic selection may also generate patterns consistent with adaptive reference-point formation (Thakral and Tô, 2020), providing an alternative (albeit not mutually exclusive) explanation to time-varying income effects patterns. Importantly, our setting features many aspects conducive to income targeting, suggesting that dynamic selection could also be empirically relevant elsewhere.

[2] Angrist et al. (2017) and Fehr and Goette (2007) are notable exceptions, but they do not employ within-day variation in income, which is essential to identify daily income targeting under expectation-based theories of reference-point formation (Kőszegi and Rabin, 2006).

I. Model

A. Setup

We model a worker's single day of work ("shift"). Each shift is divided into periods $t \in \{1, 2, ...\}$. At the beginning of each period, the worker decides whether to work ($d_t = 1$) or to end the shift ($d_t = 0$). If they decide to work at t, they provide one unit of labor. Each period is associated with a wage $w_t \in \boldsymbol{W}$, a finite set. w_t may represent hourly wages, a piece rate, or any factor affecting the pecuniary incentive to work at a given moment in time, such as traffic conditions or weather for cab drivers. The worker observes the contemporaneous wage but not subsequent wages and believes wages follow an independent stochastic process.[3] We denote the *cumulative* earnings within a shift up to (and including) period $t-1$ by Y_{t-1}.

Worker preferences. If the worker stops working at the beginning of t, we normalize their one-period utility to zero. Otherwise, their one-period utility is given by $u(t, w_t, Y_{t-1}, \varepsilon_t)$, where ε_t is a random variable capturing unobserved heterogeneity. It represents factors affecting preferences for labor supply, such as temperature, sleep, noise, or fatigue, that may vary from shift to shift and possibly within the same shift. We focus on the simpler case where the unobserved heterogeneity is fixed within a shift, $\varepsilon_t = \varepsilon$ for all t, but still may vary *across* shifts for the same worker. In the end of this section, we discuss the robustness of our results to this assumption. Importantly, the unobserved heterogeneity is essential for the model to generate a negative relation between cumulative earnings and labor supply, as we discuss in detail below.

We make minimal structural assumptions about the one-period utility function. First, we assume $u_t < 0$, capturing an increasing cost of hours worked. Second, we assume that $u_w > 0$, which implies that higher *contemporaneous* wages increase the worker's propensity of supplying labor. In a model without income effects, this assumption is true since the substitution effect is always positive. In a model with daily income effects, this assumption could in principle be violated, but to the best of our knowledge, there is no credible evidence that the (contemporaneous) wage elasticity of labor supply is negative, even in papers detecting some form of income-targeting behavior (Fehr and Goette, 2007, Chen and Sheldon, 2015, Farber, 2015).[4] Third, we normalize $u_\varepsilon > 0$.

Daily income effects are captured by the derivative of the utility with respect to accumulated income at $t-1$, $u_{Y_{t-1}}$.[5] In neoclassical labor supply models, workers consider their

[3] We assume independence to emphasize that our results do not come from the autocorrelation of wages. Our results are valid if wages follow a Markov process.

[4] An exception is Camerer et al. (1997), but the exclusion restriction in their instrument is unlikely to hold (Farber, 2005).

[5] While contemporaneous changes in income also induce income effects, those are confounded with a

lifetime wealth when supplying labor, implying that $u_{Y_{t-1}} \approx 0$ for any reasonable *daily* income variation. In contrast, theories of daily income targeting posit that $u_{Y_{t-1}} < 0$, capturing the psychological phenomenon of agents deriving negative (positive) utility from earning below (above) their daily income target (Kőszegi and Rabin, 2006). Since our goal is to show that, even without income effects, greater cumulative earnings reduce labor supply, we impose that $u_{Y_{t-1}} = 0$. The per-period utility can then be written as $u(t, w_t, \varepsilon)$.

The worker stops working whenever the per-period utility of working at t plus the option value of working in the next period is smaller than the utility of ending the shift. We describe the full dynamic problem in Appendix A.

B. Testing for Daily Income Effects and Dynamic Selection Bias

Our labor supply model connects with the hazard models often used to detect income targeting. In papers employing hazard models, a dummy capturing the decision to quit working for the day is regressed on cumulative daily earnings using intra-day data on labor supply and earnings (e.g., Crawford and Meng, 2011, Chen and Sheldon, 2015, Thakral and Tô, 2020). This regression recovers the hazard rate $Pr(d_t = 0|d_{t-1} = 1, Y_{t-1})$, capturing the probability of ending the shift at t, conditional on having cumulative income Y_{t-1}. When researchers have access to wages (or other marginal financial incentives) at different moments of the day, they can instead estimate the hazard rate conditional on any past wage w_{t-k}. In either case, the income-targeting hypothesis implies that the hazard rate should be increasing on Y_{t-1} or w_{t-k}.

Proposition 1 is our main theoretical result.

Proposition 1. *Fix a time lag $k \geq 1$, and consider a pair of wages $w_L < w_H$. The bias resulting from dynamic sample selection at t conditional on wage variation at $t-k$ is*

$$B_t(k, w_L, w_H) = Pr(d_t = 0|d_{t-1} = 1, w_{t-k} = w_H) - Pr(d_t = 0|d_{t-1} = 1, w_{t-k} = w_L). \quad (1)$$

The following holds:

1. *For any k,*
$$B_t(k, w_L, w_H) \geq 0 \quad (2)$$
with strict inequality for at least $k = 1$.

2. *If $\boldsymbol{W} = \{w_L, w_H\}$, then $B_t(k, w_L, w_H)$ is nonincreasing in k. If $|\boldsymbol{W}| > 2$, $B_t(k, w_L, w_H)$ can be non-increasing, nondecreasing, or nonmonotonic in k.*

countervailing substitution effect. Thus, we focus on income effects from previously earned income during the shift like Farber (2005).

Item 1 shows there is a positive relation between wages accumulated early in the shift (given by the kth lag) and the exit likelihood at a subsequent period, even without daily income effects. The bias is caused by dynamic (sample) selection (Diamond and Hausman, 1984). The hazard rate is, by definition, conditioned on $d_{t-1} = 1$, that is, the sample of workers who have not dropped out before t. If a worker decides to keep working at $t-k$ after receiving a low wage, ε needs to be relatively high to compensate the low wage. Therefore, those workers tend to be positively selected in the distribution of ε, in comparison to workers who received a higher wage at $t-k$. Since the shock ε is persistent throughout the shift, the selection leads to the difference in hazard rates noted in Equation (2). As they qualitatively point to the same direction, the positive relation created by this bias is confounded with the existence of daily income effects even in the presence of randomized wages.

Item 2 of Proposition 1 discusses the timing patterns of the dynamic selection bias. When there are only two wages in the possible wages set, recent variations in wages cause a larger bias in the hazard rate than older ones. With more than two wages, the pattern can go either way.

We highlight the main intuitive force behind the pattern with two wages. Consider the same wage differential $w_H - w_L$ occurring at $t-1$ and $t-2$, and consider the dynamic selection bias at period t. Workers who survived in the sample until reaching period t must have decided to keep on working at both $t-1$ and $t-2$. Since at each period the cost of effort is increasing, the threshold $\bar{\varepsilon}$ necessary to continue working at $t-1$ is greater than the one at $t-2$. Thus, the positive selection in the distribution of the unobserved heterogeneity ε conditional on the wage difference at $t-1$ is greater than of those facing the same wage difference at $t-2$. In summary, the harder it is to keep working, the greater ε needs to be for a worker to keep on working and the greater the bias caused by dynamic sample selection.

However, if there are more than two wages in the possible wages set, we cannot establish the pattern's monotonicity. Although the main intuition is still true, the possible wage paths after the differential $w_H - w_L$ occurs in period $t-k$ may imply that all bias more recent than $t-k$ is equal to zero, while the bias from period $t-k$ is positive, thus implying a nondecreasing pattern. One can construct many examples since we impose minimal restrictions on the per-period utility function and no restriction on $\mathbf{W}$ beyond finiteness. One example is when the cross-derivative of $u(.)$ with respect to (t, w) is negative (i.e., decreasing differences). In this case, workers care more about wages accumulatd in earlier periods. Thus, the bias induced by the difference in wages at $t-1$ may be weaker than that in $t-2$. We use this example in Appendix *A* to prove the existence of nonmonotonic patterns.

Item 2 is important in light of the recent literature on reference-point formation. Thakral and Tô (2020) find a pattern where more recent income variation is associated with stronger point estimates on the hazard rate, corresponding to a nonincreasing pattern in our language. They rationalize this pattern with a model of adaptive reference-dependent preferences, in which workers gradually adjust their reference point to account for recent variations in income, reducing the weight of older variations in earnings on the hazard rate. Our model provides a possible alternative explanation to this phenomenon.

C. Addressing Dynamic Selection

Proposition 2 outlines a sub-sampling strategy to eliminate the dynamic selection bias.

Proposition 2. *Let* $\underline{w} = \min \boldsymbol{W}$. *Let* $w_{t-1} = \underline{w}$ *and* $\boldsymbol{w}^{t-2}$ *be an arbitrary wage history up to* $t-2$. *The wage history* $\boldsymbol{w}^{t-2}$ *does not predict the hazard rate at* t; *that is,*

$$Pr(d_t = 0 | d_{t-1} = 1, w_{t-1} = \underline{w}, \boldsymbol{w}^{t-2}) = Pr(d_t = 0 | d_{t-1} = 1, w_{t-1} = \underline{w}).$$

Proposition 2 states that any wage prior to session $t-1$ has no predictive effect on the stopping decision once we condition on $w_{t-1} = \underline{w}$, solving the dynamic selection issue. If past wages affect the stopping decision after conditioning on $w_{t-1} = \underline{w}$, there is evidence in favor of income targeting. The cost of our approach is that income effects from earnings in the immediately previous period are not identified.

Proposition 2 has the following intuition: given the cut-off strategy to quit, each decision to keep working only informs us that worker utility $u(t, w_t, \varepsilon)$ is above a given threshold—or equivalently that ε is above a threshold $\bar{\varepsilon}(t, w_t)$. Since $u_t < 0$ and $u_w > 0$, the lowest utility received by a worker is under wage $\underline{w}$ (lowest wage possible) and at period $t-1$ (last period, when the marginal cost of effort is highest). This implies that $\bar{\varepsilon}(t-1, w_{t-1}) \geq \bar{\varepsilon}(t', w_{t'})$ for $t' \leq t-1$. Thus, if the worker has decided not to quit despite having the lowest wage and the highest marginal cost of effort, the decisions and wages before $t-1$ do not add additional information about ε after conditioning on $w_{t-1} = \underline{w}$.

One shortcoming of our analysis is that we assume the unobserved heterogeneity is fixed within a shift, ($\varepsilon_t = \varepsilon$). Our results are partially robust to relaxing this assumption. Dynamic selection still biases income effects as long as there is some persistence in ε_t over the shift. This is a reasonable assumption since many factors such as heat, sleepiness, and fatigue are likely positively autocorrelated throughout the shift. The strategy to address dynamic selection outlined in Proposition 2 relies on how much the decision to keep working at $t-1$ is predictive of the shock at t. If the unobserved heterogeneity has a low persistence, then our strategy will not remove the bias. Note also that our strategy cannot increase the bias from dynamic selection, so in the worst case scenario, the estimates from our strategy will be as biased as conventional estimates.

II. Setting and Empirical Framework

A. Setting

We use data from an RCT with 452 low-income workers who were hired on a rolling basis for a four-week data-entry job, based in Chennai, India (Bessone et al., 2021).[6] Study participants' main activity consisted of transcribing alphanumeric data designed to mimic a real-world data-entry job. Participants spent, on average, two-thirds of their time on the data-entry work and the rest completing surveys and experimental tasks. The data they digitized were artificially generated, homogenizing the task across participants and allowing the measurement of their accuracy. The software also allows us to objectively measure time spent typing and in breaks, earnings, and output.[7]

Our setting is suitable to investigate income effects for at least three reasons. First, participants were free to choose when to arrive at the office and when to stop working. Workdays were classified as a "regular" or a "short" day. About half of the 28 workdays were regular days (12–14 days), when participants were allowed to work from 9:30 am to 8:00 pm, providing flexibility in labor supply choices in the extensive (how many hours to work) and intensive (how much effort to exert) margins. On short days (six to seven days in total), work was restricted from 11:00 am to 5:00 pm, and participants received a monetary incentive to comply with these working hours. As most participants complied, the extensive margin of labor supply is effectively shut down on short days. We exclude a few days from the analysis: days one and two were mostly for training, and a few other days included time-consuming experimental tasks, which could restrict labor supply.

Second, there is exogenous, frequent variation in hourly earnings. The daily payments for the data-entry work comprised 73% of participants' average daily earnings (Rs. 343) and were determined by two components: a piece-rate payment per unit of output (66% of data-entry earnings) and a constant hourly rate (other earnings include payments for surveys and for performance in experimental tasks). Each working day (shift hereafter) was divided into *sessions* that lasted for roughly 30 minutes of typing time.[8] Piece rates were randomized in the session level, taking either a high or a low value with equal probability (Rs. 0.02 or Rs. 0.005 per correct character). Participants were penalized by Rs. 0.10 per mistake. All

[6] Summary statistics are available in Table C2. Details about screening criteria can be found in Bessone et al. (2021)

[7] Output was pre-registered as the number of correct character entries minus eight times the number of mistakes (Bessone et al., 2021). Because the accuracy rate is above 99%, it is almost perfectly correlated with the number of correct entries.

[8] After 30 minutes of typing time in a session, the session would switch upon the worker submitting a data field. See more details in Figure C2.

payments accrued in a given day were carried out at the end of the same day.

Third, piece rates in the current session were easily observable, allowing workers to target income. In most shifts, the piece rate was always visible in the corner of their screen (Figure C2) and were associated with distinct colors. The participants' screen also blinked by the end of a session, indicating that the piece rate might have changed.[9] The participants produced 19% more under high piece rates, showing that they noticed the piece rates. A drawback is that participants could not track their exact cumulative earnings. However, earnings are determined by total output, typing time, and the current piece rate, which are all observable in each period. We show in Table C3 that piece rates account for most of the variance in total earnings.

B. Empirical Framework

We estimate the relation between labor supply and cumulative income with the regression

$$y_{idt} = \sum_{j=0}^{4} \beta_j \text{High}_{idt-j} + \gamma X_{idt} + \nu_{idt}, \tag{3}$$

where y_{idt} is one of the following labor supply measures for participant i at date d and session t: output, minutes spent in work breaks, and a dummy indicating the decision to stop at t. The variables High_{idt-j} are dummies indicating whether session $t-j$ had a high piece rate. X_{idt} is a vector of covariates including participant, day in study, session, and date fixed effects. Our empirical specification closely follows those used to estimate daily income effects in other income-targeting settings (Farber, 2005, 2008, 2015, Crawford and Meng, 2011, Chen and Sheldon, 2015, Hammarlund, 2018, Thakral and Tô, 2020). While they focus solely on the the decision to stop working, we can explore additional margins of labor supply.

The main coefficients of interest are β_j (for $j > 0$). They capture the impact of previous wage variation on labor supply. Assuming that wages are exogenous, that is, $\mathbb{E}[\mathbf{High}_{id} \cdot \nu_{idt}] = 0$, the neoclassical model predicts that $\beta_j = 0$ for any $j > 0$. Under income targeting, positive income shocks should decrease subsequent labor supply, implying that $\beta_j < 0$ for $j > 0$ when the outcome is output and $\beta_j > 0$ for breaks and the stopping decision. We also include the contemporaneous piece rate, with the associated coefficient β_0 capturing the price effect of high piece rates on labor supply. Because β_0 conflates substitution and potential income effects, $\beta_0 > 0$ does not contradict the income-targeting hypothesis.

[9] In some of the shifts, the piece rates were less salient, only being available for a few seconds after they changed (Figure C2, panel b). The results are similar in those days (Table C6).

Dynamic selection implies that $\mathbb{E}[\mathbf{High}_{id} \cdot \nu_{idt}] \neq 0$ even if wages are randomly assigned. In that case, estimates of income effects given by β_j will be downward biased. Controlling for observables X_{idt} including individual, date, day in study, and session fixed effects may help with dynamic selection, although as discussed in Appendix Section B and elsewhere (Kyriazidou, 1997), it may not entirely solve the dynamic selection bias. We still add controls to our analysis to increase the precision of our estimates.

We address dynamic selection in three ways. First, we employ the solution proposed in Proposition 2. Second, we restrict the analysis to short days. Because the extensive margin of labor supply–and therefore dynamic selection—is shut down, we look for income targeting in the intensive margin. Third, we explore in the appendix an exogenous variation in earnings, which cannot trigger dynamic selection. Together, these strategies allow us to estimate income effects without dynamic selection.

Estimating daily income effects presents other well-known econometric challenges. First, papers in this literature often rely on nonexperimental income variation. If cumulative wages are affected by supply shocks, estimates of income effects would be negatively biased. Second, labor supply decisions depend on the continuation value of working, which in turn depends on future expectations of wages. If wages are autocorrelated, the unobserved continuation value potentially confounds the estimation of income effects. Our paper bypasses these issues by randomly assigning wages throughout the shift. Third, higher past wages may induce more fatigue if workers exert more effort when wages are higher. If fatigue affects labor supply, cumulative income would be negatively related to labor supply. Therefore, our results from the piece-rate variation, as well as the results in the literature, should be interpreted as an upper bound of the magnitude of the income effects.

III. Results

A. Income Effect and Dynamic Selection

Dynamic selection only biases estimates of income effects when (i) the hazard rate is decreasing in contemporaneous wages and (ii) there is relevant unobserved heterogeneity in labor supply preferences. Figure 1 shows that (i) holds in our sample. In the beginning of the shift, there is very little quitting, so there is no difference in the hazard rate conditional on high (red) or low piece rates (blue). However, as the shift progresses, there are significant differences in quitting probabilities by piece rates. At 5.4 hours into the shift, the difference is already significant at the 5% level. By the end of the shift, a 20 percentage point (pp) gap between the hazard rate conditional on high and low wages appears. While condition

(ii) is not testable, it is hard to imagine a setting in which we can account for all heterogeneity shaping the labor supply decision, even when we add worker fixed effects. In the likely event that the differential attrition by piece rate is related to the participants' unobserved heterogeneity as in our model, the conventional estimates of income effects would suffer from dynamic selection bias.

In our setting, the estimates of the conventional hazard model indicate the presence of income targeting. Table 1 (columns 1–3) shows that when workers have discretion to choose their working hours and dynamic selection is not addressed, there are patterns consistent with income targeting for our three measures of labor supply: output, which falls by 98 units (6.6%, $p < 0.01$); (ii) pauses, which increase by 0.88 minutes (47%, $p < 0.01$); and (iii) the hazard rate, which increases by 4.3 pp (30.7%, $p < 0.01$). These estimates, which we call the *aggregate effect*, correspond to the sum of the first four piece-rate lags. Rows 2–5 confirm that most lags are statistically different than zero and all income effects are still significant (albeit smaller) if we exclude the first lag (row 2).

However, these results are an artifact of dynamic selection. In columns 4–6, we estimate the empirical model restricting the sample to workers who received a low piece rate in the immediately previous session. As discussed in Proposition 2, this sub-sample is not affected by dynamic selection. So, if income effects persisted, we would have credible evidence of income targeting; they did not. In this sub-sample, we find that the aggregate income effects (row 2) are close to zero and insignificant, and they even flip signs for output and the hazard rate. These results are not driven by excluding the first lag from the overall effect, which is required by Proposition 2. As discussed above, the conventional estimates still point toward income targeting, even excluding the first lag.

Even though this strategy halves the sample size, we are well powered to reject the null. For output, for example, we are powered to detect effects as small as 33.7 units (2.2%). Moreover, conditioning on receiving a high piece rate in the previous session, instead of conditioning on receiving a low one, yields even stronger evidence of income targeting in spite of both sub-samples having roughly the same size (Table C4). This is consistent with our model since dynamic selection should still affect the sample that just received a high piece rate.

Two additional results corroborate the absence of income targeting when dynamic selection is not a concern. First, we see no evidence of income targeting in the sample of short days (columns 7–9). On these days, participants are restricted to stop working at 5:00 pm, which effectively shuts down the extensive margin of labor supply. Importantly, workers could still adjust labor supply via effort provision or work breaks, allowing us to test for income targeting in both dimensions.[10] Nevertheless, the overall effects are small and statistically insignificant. There is a 14 second increase in work breaks, but it is offset by a tantamount increase in lag 4. Second, in Appendix Section C, we show that variation in payments for a different study task do not affect any dimension of labor supply in the rest of the day. Importantly, participants only learned about the amount they would be paid moments before the task. Thus, they could not end the shift in response to their payment level before or during the task, which could lead to dynamic selection.

After addressing dynamic selection, we find no evidence of income targeting. But is our setting suitable for income targeting? First, why would workers engage in income targeting in the first place? One possible reason is that participants in our sample are present biased on average (Bessone et al., 2021). This presents a rationale for income targeting as a way to deal with self-control issues (Read et al., 1999, Koch and Nafziger, 2016). Second, participants could be unable to track income throughout the day. However, as detailed in Section A, the participants could infer their accumulated income via piece rates, which explains most of the variation in income for a person. Third, perhaps income targeting evolves over time when participants develop more experience with the setting. This is unlikely since we see no differences in results when looking only at the last week of work (Table C7). Moreover, Farber (2015) finds that more experienced drivers are *less* likely to engage in income targeting.

B. Implication to the Literature

Dynamic selection may bias estimates of income effects from hazard models under two conditions that are likely present in most settings. First, workers must be heterogeneous in factors that are nonobservable and vary from shift to shift. It is hard to imagine a worker not subject to unobserved factors such as sleep, health, eating, or fatigue affecting the marginal cost of labor across shifts. Unobserved heterogeneity is likely to be an issue as long as the econometrician cannot control entirely for it. Second, the effect of current wages on the hazard rate needs to be positive. This condition holds under the neoclassical model and could hold even in the presence of income targeting (Kőszegi and Rabin, 2006). This is the case in our setting, for ride-share drivers, who are less likely to stop working during price surges (Chen and Sheldon, 2015), and bike messengers (Fehr and Goette, 2007), for example.

Dynamic selection might be an issue even in settings without an explicit instantaneous wage. In such settings, workers may still form beliefs about the marginal financial incentive to keep working. For example, cab drivers may use cues such as traffic or special events in the city that affect demand for rides to form beliefs about their instantaneous financial incentive

[10] In expectation-based models of income targeting, workers should target income even on short days since the target should be adjusted down to account for less working hours (Kőszegi and Rabin, 2006).

to keep working. As evidence that they do, Farber (2015) estimates the instantaneous wage at each minute of the day for NYC cab drivers, which reveals a positive wage elasticity of labor. Similarly, Giné et al. (2016) find that fishermen in India are more likely to work when predicted demand is high, even though they cannot directly observe the financial incentive to work. Replacing known wage rates by expected ones should not change the conclusions of the model since dynamic selection occurs as longs as workers supply more labor when expected instantaneous incentives are high.

Although it is likely that most income-targeting estimates in the literature present some degree of dynamic selection, the magnitude of the bias is unclear. In our setting, marginal incentives vary substantially and are very salient, which should contribute to the bias being larger than elsewhere. Since we cannot apply our strategy to separate income effects from dynamic selection for NYC cab drivers or in other settings, we investigate how our results compare with prominent papers in the literature. We re-estimate columns 3 and 6 of Table 1 with a specification directly comparable to Farber (2015) and Thakral and Tô (2020). We show in Figure C2 that a 10% increase in cumulative earnings early in the shift is associated with an increase of roughly 9.5% in the stopping probability later in the day. In comparison, Farber (2015) find a 9% increase when focusing on drivers working in daytime shifts, and Thakral and Tô (2020) report a 3.3% increase in a sample that includes day and nighttime shifts. While in our setting these effects are entirely driven by dynamic selection, the size of dynamic selection bias might be smaller elsewhere.

Our results also speak to the literature on reference-point formation. In a seminal contribution, Thakral and Tô (2020) show that cab drivers' income effects are more pronounced for recent earnings than for earlier ones. They rationalize this nonfungibility of income shocks over time with a theory of adaptive reference-point formation, in which earnings accumulated early in the shift help form the worker's income target. Dynamic selection provides an alternative (albeit nonmutually exclusive) explanation for their results (item 2 of Proposition 1). Our setting presents the same nonfungibility patterns (Figure 2). Not accounting for dynamic selection (left-hand graphs), income effects are positive for recent piece-rate lags but decrease over time for all measures of labor supply. However, when we separate the sample into low and high piece rates in the previous sessions, it is clear that dynamic selection is responsible for the nonfungibility pattern. When we use the low piece-rate sample (right-hand graphs, orange series), the pattern disappears. The nonfungibility pattern comes exclusively from the high piece-rate series (blue), which does not address dynamic selection.

IV. Conclusion

Our theoretical and empirical results suggest that studies considering daily income effects should investigate the empirical relevance of dynamic selection. In our setting, disregarding this bias would change the interpretation of the results. We propose two different strategies to deal with this issue. First, we use a model-based strategy to eliminate dynamic selection that may apply elsewhere. Second, researchers could investigate the presence of income effects in settings without an extensive margin of labor supply. Instead of focusing on the hazard rate, they could study income effects in other margins of labor supply, such as effort provision.

The solutions we propose are not feasible in every setting. When instantaneous wages are not observable, such as for a cab driver, the methods do not apply directly. But they might be used in the context of ride-sharing drivers, in which the researchers may observe or possibly even manipulate the trip multiplier for some drivers. Alternatively, future research could deal with dynamic selection using dynamic discrete choice structural models that incorporate dynamic selection explicitly (Abbring and Heckman, 2007).

References

Abbring, J. H. and J. J. Heckman (2007). Econometric evaluation of social programs, part iii: Distributional treatment effects, dynamic treatment effects, dynamic discrete choice, and general equilibrium policy evaluation. *Handbook of econometrics 6*, 5145–5303.

Andersen, S., A. Brandon, U. Gneezy, and J. A. List (2018). Toward an Understanding of Reference-Dependent Labor Supply: Theory and Evidence From a Field Experiment. *NBER Working Paper 60637*(No.20695), 1–38.

Angrist, J. D., S. Caldwell, and J. V. Hall (2017). Uber vs. taxi: A drivers eye view. Technical report, National Bureau of Economic Research.

Bessone, P., G. Rao, F. Schilbach, H. Schofield, and M. Toma (2021). Sleepless in Chennai: The Consequences of Sleep Deprivation Among the Urban Poor. *Quarterly Journal of Economics 136*(3), 18871941.

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

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

Chen, M. K. and M. Sheldon (2015). Dynamic Pricing in a Labor Market: Surge Pricing and the Supply of Uber Driver-Partners. *Working Paper*, 1–18.

Crawford, V. P. and J. Meng (2011). New York city cab drivers' labor supply revisited: Reference-dependent preferences with rational-expectations targets for hours and income. *American Economic Review 101*(5), 1912–1932.

Diamond, P. and J. Hausman (1984). Individual retirement and savings behavior. *Journal of Public Economics 23*, 81–114.

Dupas, P., J. Robinson, and S. Saavedra (2020). The daily grind: Cash needs and labor supply. *Journal of Economic Behavior and Organization 117*.

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

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

Farber, H. S. (2015). Why You Can't Find a Taxi in the Rain and Other Labor Supply Lessons From Cab Drivers. *The Quarterly Journal of Economics 130*(4), 1975–2026.

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

Giné, X., M. Martinez-Bravo, and M. Vidal-Fernández (2016). *Are labor supply decisions consistent with neoclassical preferences? Evidence from Indian boat owners.* The World Bank.

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

Henderson, P. W. and R. A. Peterson (1992). Mental accounting and categorization. *Organizational Behavior and Human Decision Processes 51*(1), 92–117.

Kaur, S., S. Mullainathan, S. Oh, and F. Schilbach (2019). Does financial strain lower productivity? Technical report, Working Paper.

Koch, A. K. and J. Nafziger (2016). Goals and bracketing under mental accounting. *Journal of Economic Theory 162*, 305–351.

Kőszegi, B. and M. Rabin (2006). A Model of Reference-Dependent Preferences. *The Quarterly Journal of Economics CXII*(August), 1133–1165.

Kyriazidou, E. (1997). Estimation of a panel data sample selection model. *Econometrica: Journal of the Econometric Society*, 1335–1364.

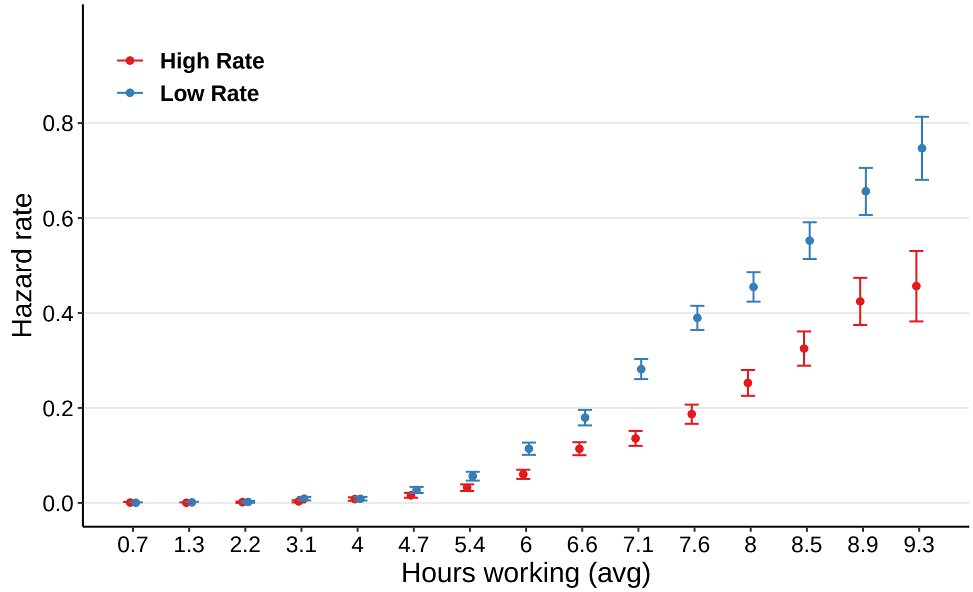
Lancaster, T. (2000). The incidental parameter problem since 1948. *Journal of econometrics 95*(2), 391–413.

Read, D., G. Loewenstein, and M. Rabin (1999). Choice Bracketing. *Journal of Risk and Uncertainty 19*(1-3), 171–197.

Schofield, H. (2014). The economic costs of low caloric intake: Evidence from india. *Unpublished Manuscript*.

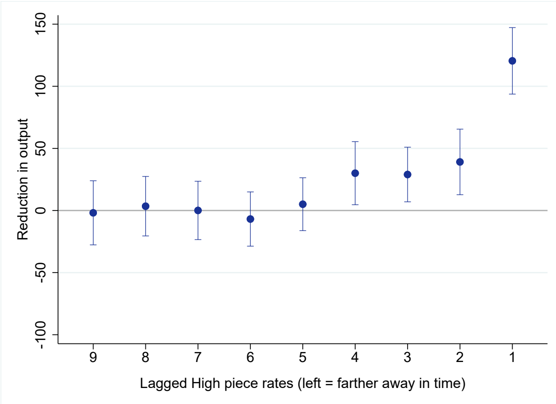
Thakral, N. and L. T. Tô (2020). Daily Labor Supply and Adaptive Reference Points. *American Economic Review Forthcoming.*

Figure 1: Hazard rate conditional on high versus low piece rate

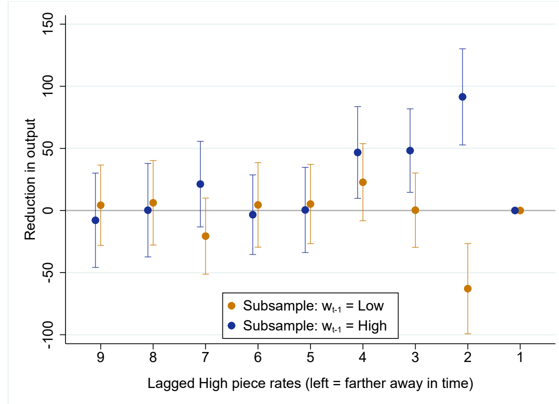


Notes: This figure presents the hazard rate (average probability of quitting) at different moments of the day (sessions t). We show the hazard rate conditional on high (red) and low (blue) piece rates. The whisker bars represent 95% confidence intervals. We show in the x-axis the average number of hours into the shift by the end of each session.

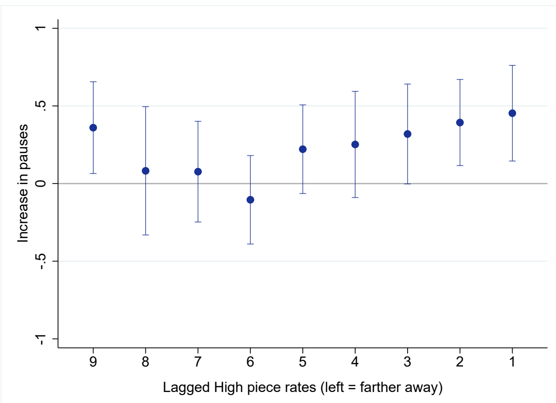
Figure 2: Income effects: Role of timing



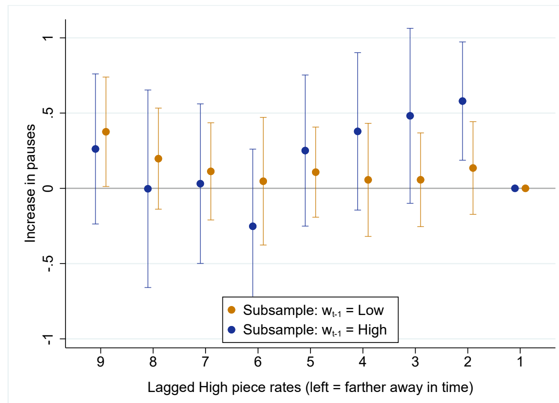
(a) Effects of lagged piece rates on output - full sample



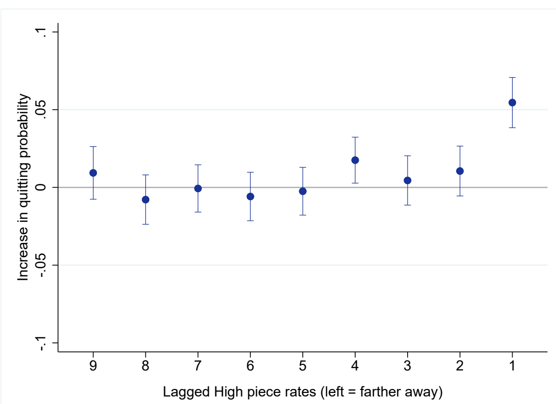
(b) Effects of lagged piece rates on output - sub-samples



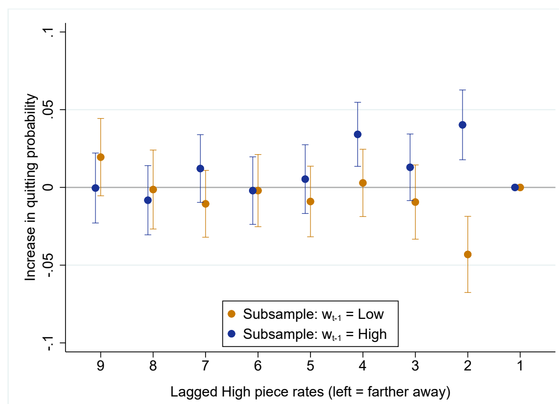
(c) Effects of lagged piece rates on pauses - full sample



(d) Effects of lagged piece rates on pauses - sub-samples



(e) Effects of lagged piece rates on probability of exit - full sample



(f) Effects of lagged piece rates on probability of exit - sub-samples

Notes: This figure presents the effects of nine piece rate lags on output, pauses, and the hazard rate. The specification follows Equation (3), but adds more lags. Panels on the left consider the whole sample, while panels on the right consider the sub-samples that received a high/low piece rate on the previous section. The bars represent 95% confidence intervals.

Table 1: Effect of piece-rate variation on labor supply

	Regular days			Regular, conditional on w_{t-1} = low			Short days		
	(1) Output	(2) Work Breaks	(3) Quit	(4) Output	(5) Work Breaks	(6) Quit	(7) Output	(8) Work Breaks	(9) Quit
Total effect: lags 1-4	-98.45	0.88	0.04				19.34	0.04	-0.01
	(12.31)	(0.19)	(0.01)				(19.19)	(0.19)	(0.01)
	[0.00]	[0.00]	[0.00]				[0.31]	[0.82]	[0.38]
Total effect: lags 2-4	-42.16	0.51	0.01	2.63	0.03	-0.01	10.18	-0.18	-0.00
	(9.49)	(0.14)	(0.01)	(12.05)	(0.11)	(0.01)	(16.29)	(0.15)	(0.01)
	[0.00]	[0.00]	[0.05]	[0.83]	[0.77]	[0.14]	[0.53]	[0.23]	[0.99]
High lag 1	-56.29	0.367	0.031				9.17	0.223	-0.011
	(6.17)	(0.083)	(0.004)				(9.54)	(0.076)	(0.006)
	[0.00]	[0.00]	[0.00]				[0.34]	[0.00]	[0.08]
High lag 2	-18.22	0.219	0.008	15.47	-0.068	-0.005	-1.17	-0.007	0.005
	(5.12)	(0.065)	(0.003)	(6.41)	(0.075)	(0.005)	(9.08)	(0.079)	(0.006)
	[0.00]	[0.00]	[0.01]	[0.02]	[0.36]	[0.28]	[0.90]	[0.93]	[0.44]
High lag 3	-11.45	0.127	0.000	-6.98	0.021	-0.001	-3.23	0.049	0.002
	(4.62)	(0.068)	(0.003)	(6.28)	(0.070)	(0.004)	(8.17)	(0.082)	(0.006)
	[0.01]	[0.06]	[0.94]	[0.27]	[0.76]	[0.79]	[0.69]	[0.55]	[0.71]
High lag 4	-12.48	0.167	0.003	-5.85	0.081	-0.006	14.58	-0.223	-0.007
	(5.68)	(0.069)	(0.003)	(6.16)	(0.066)	(0.004)	(9.02)	(0.070)	(0.006)
	[0.03]	[0.02]	[0.31]	[0.34]	[0.23]	[0.15]	[0.11]	[0.00]	[0.24]
High	284.71	-1.060	-0.083	261.39	-0.842	-0.077	156.60	-0.648	-0.004
	(12.13)	(0.100)	(0.004)	(12.23)	(0.090)	(0.005)	(12.11)	(0.089)	(0.007)
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.57]
Excluded Group Mean	1470.75	1.87	0.14	1512.75	1.70	0.11	1460.20	1.64	0.16
Excluded Group SD	941.48	6.24	0.34	925.67	5.00	0.31	922.50	4.57	0.37
Observations	33716	33716	33716	16181	16181	16181	9628	9628	9628
Participants	452	452	452	452	452	452	452	452	452

Notes: This table presents estimates of Equation (3) for three labor supply outcomes: output, work breaks, and a dummy indicating the decision to quit for the day. Columns 1-6 restrict the sample to regular days, when participants had full discretion to choose when to quit (see Section A). In columns 1-3 we use the entire sample, while in columns 4-6 we restrict the sample to sessions for which the previous session had a Low piece rate, addressing the dynamic selection (see Proposition 2). Columns 7 to 9 restrict the sample to short days, when participants received monetary incentives to quit at exactly 5 PM, effectively shutting down the extensive margin of labor supply, which also addresses dynamic selection. "Excluded group" stands for the sample where all piece-rate indicators included in the regression are equal to zero. Row 1 computes the sum of the four lags, while row 2 considers the effect of lags 2-4 to ensure comparability between columns 4-6 and the others. Errors clustered at the worker-level are displayed in parenthesis, with the associated p-values in brackets.

Online Appendix

A. Proofs

Under the assumptions we made, the optimal stopping rule for the agent follows a simple monotonic strategy, which we formalize in Lemmas 1 and 2.

Lemma 1. *For every wage w and time period t, there exists a unique value $\overline{\varepsilon}(t,w)$ such that the agent keeps working at t (i.e., $d^*(t,w,\varepsilon)=1$) iff $\varepsilon \geq \overline{\varepsilon}(t,w)$. Moreover, $\overline{\varepsilon}(t,w)$ is increasing on t and decreasing on w.*

Proof:

Let $V^e(t+1,\varepsilon) = \mathbb{E}\left[V(t+1,w_{t+1},\varepsilon)\right]$ and $U(t,w,\varepsilon) \equiv u(t,w_t,\varepsilon) + V^e(t+1,\varepsilon)$. The agent chooses $d^*(t,w,\varepsilon)=1$ iff $U(t,w,\varepsilon) \geq 0$.

By assumption, $u_\varepsilon(t,w,\varepsilon) > 0$ for any (t,w). This implies that $V^e(t+1,\varepsilon)$ is also (weakly) increasing on ε, since any agent facing two different values of ε can always choose exactly the same stopping-decision path and receive a higher inter-temporal utility under the higher value of ε. Thus, $U(t,w,\varepsilon)$ is increasing on ε. This implies that there is at most one point $\overline{\varepsilon}(t,w)$ such that $U(t,w,\overline{\varepsilon}(t,w))=0$. We know such a point exist because of technical assumption 1, implying that $U(t,w,\varepsilon) \geq 0$ iff $\varepsilon \geq \overline{\varepsilon}(t,w)$.

We now prove the comparative static results. Take two wages $w' > w$. We have that

$$U(t,w,\overline{\varepsilon}(t,w)) \equiv u(t,w,\overline{\varepsilon}(t,w)) + V^e(t+1,\overline{\varepsilon}(t,w)) = 0 \tag{4}$$
$$U(t,w',\overline{\varepsilon}(t,w')) \equiv u(t,w',\overline{\varepsilon}(t,w')) + V^e(t+1,\overline{\varepsilon}(t,w')) = 0 \tag{5}$$

Because $u_w > 0$, $u(t,w',\varepsilon) > u(t,w,\varepsilon)$ for any t,ε. Since, $U(t,w,\overline{\varepsilon}(t,w)) = U(t,w',\overline{\varepsilon}(t,w'))$ and $u_\varepsilon > 0$, this means that $\overline{\varepsilon}(t,w') < \overline{\varepsilon}(t,w)$.

Take now two time periods such that $t' > t$. Again,

$$U(t,w,\overline{\varepsilon}(t,w)) \equiv u(t,w,\overline{\varepsilon}(t,w)) + V^e(t+1,\overline{\varepsilon}(t,w)) = 0 \tag{6}$$
$$U(t',w,\overline{\varepsilon}(t',w)) \equiv u(t',w,\overline{\varepsilon}(t',w)) + V^e(t'+1,\overline{\varepsilon}(t',w)) = 0 \tag{7}$$

We have that both $u_t(t,w,\varepsilon) < 0$ and $V^e_t(t,\varepsilon) \leq 0$ for any t,w,ε, implying that $U(t',w,\varepsilon) < U(t,w,\varepsilon)$ for any ε. Since, (i) $U(t,w,\overline{\varepsilon}(t,w)) = U(t',w,\overline{\varepsilon}(t',w))$ and (ii) $U_\varepsilon \geq 0$, it must be the case that $\overline{\varepsilon}(t,w) < \overline{\varepsilon}(t',w)$. □

Lemma 2. *Take a wage history $\boldsymbol{w}^t = (w_1,...,w_t)$ and define*

$$\hat{\varepsilon}(\boldsymbol{w}^t) \equiv \max_{t' \in \{1,2,\ldots,t\}} \{\overline{\varepsilon}(t',w_{t'})\} \tag{8}$$

where $\overline{\varepsilon}(t',w_{t'})$ is defined as in Lemma 1.

The survival rate at $t+1$ is given by

$$\mathbb{E}\left[d_{t+1}|\boldsymbol{w}^t, d_t = 1\right] = \mathbb{E}\left[d_{t+1}|\varepsilon \geq \hat{\varepsilon}(\boldsymbol{w}^t)\right] \tag{9}$$

Proof

For each $t' \leq t$ we observe $w_{t'}$ and $d_{t'} = 1$. This is equivalent to observing $U(t', w_{t'}, \varepsilon) > 0 \iff \varepsilon > \overline{\varepsilon}(t', w_{t'})$. Aggregating the wage and stopping decision at each $t' \leq t$ yields $\varepsilon \geq \hat{\varepsilon}(\boldsymbol{w}^t)$ for all t. □

Proof of Proposition 1.

We prove the results for $\mathbb{E}\left[d_t|\boldsymbol{w}^{t-1}, d_{t-1} = 1\right] = 1 - Pr\left[d_t = 0|\boldsymbol{w}^{t-1}, d_{t-1} = 1\right]$. From now on we may omit the conditional $d_{t-1} = 1$ to ease the notation, but all the expectations below are taken conditioning on it.

Item 1, existence of dynamic selection bias: from Lemma 2, $\mathbb{E}\left[d_t|\boldsymbol{w}^t, d_{t-1} = 1\right] = \mathbb{E}\left[d_t|\varepsilon \geq \hat{\varepsilon}(\boldsymbol{w}^t)\right]$. Note first that

$$\mathbb{E}\left[d_t|\varepsilon \geq y\right] \geq \mathbb{E}\left[d_t|\varepsilon \geq x\right] \iff y \geq x \tag{10}$$

since the random variable $\varepsilon|\varepsilon > y$ first-order stochastically dominate (FOSD) $\varepsilon|\varepsilon > x$ for $y > x$, and $d_t = \mathbf{1}\left\{U(t, w, \varepsilon) \geq 0\right\}$ is a non-decreasing function of ε. Denote a wage history up to $t-1$ not including the period $t-k$ as $\boldsymbol{w}^t_{-(t-k)}$. Then, we can decompose the expectation of d_t in terms of all possible wage histories:

$$\mathbb{E}\left[d_t|w_{t-k} = w\right] = \sum_{\boldsymbol{w}^t_{-(t-k)}} \mathbb{E}\left[d_t|w_{t-k} = w, \boldsymbol{w}^t_{-(t-k)}\right] \cdot Pr\left(\boldsymbol{w}^t_{-(t-k)}\right) \tag{11}$$

We can then write $B_t(k, w_h, w_l) = \mathbb{E}\left[d_t|w_{t-k} = w_l\right] - \mathbb{E}\left[d_t|w_{t-k} = w_h\right]$ as

$$B_t(k, w_h, w_l) = \sum_{\boldsymbol{w}^t_{-(t-k)}} Pr\left(\boldsymbol{w}^t_{-(t-k)}\right) \cdot \left(\mathbb{E}\left[d_t|w_{t-k} = w_l, \boldsymbol{w}^t_{-(t-k)}\right] - \mathbb{E}\left[d_t|w_{t-k} = w_h, \boldsymbol{w}^t_{-(t-k)}\right]\right) \tag{12}$$

Note that

$$\begin{aligned} &\mathbb{E}\left[d_t|w_{t-k} = w_l, \boldsymbol{w}^t_{-(t-k)}\right] - \mathbb{E}\left[d_t|w_{t-k} = w_h, \boldsymbol{w}^t_{-(t-k)}\right] = \\ &= \mathbb{E}\left[d_t|\varepsilon > \max\left\{\overline{\varepsilon}(t-k, w_l), \hat{\varepsilon}\left(\boldsymbol{w}^t_{-(t-k)}\right)\right\}\right] - \mathbb{E}\left[d_t|\varepsilon > \max\left\{\overline{\varepsilon}(t-k, w_h), \hat{\varepsilon}\left(\boldsymbol{w}^t_{-(t-k)}\right)\right\}\right] \geq 0 \end{aligned}$$

This expression is always non-negative because

$$\max\left\{\overline{\varepsilon}(t-k, w_l), \hat{\varepsilon}\left(\boldsymbol{w}^t_{-(t-k)}\right)\right\} \geq \max\left\{\overline{\varepsilon}(t-k, w_h), \hat{\varepsilon}\left(\boldsymbol{w}^t_{-(t-k)}\right)\right\},$$

since Lemma 1 shows that $\overline{\varepsilon}(t-k, w)$ is decreasing on w, and because of inequality 10. Therefore $\mathbb{E}\left[d_t|w_{t-k} = w_l\right] \geq \mathbb{E}\left[d_t|w_{t-k} = w_h\right]$. When $k = 1$, this inequality is strict. To see that, we just need to show that one of the terms in 12 is strictly positive. To do that, we fix a wage history $\boldsymbol{w}^{t-2}$ such that $w_{t'} = w_h$ for each t'. Then by Lemma 1

$$\overline{\varepsilon}(t-1, w_l) > \overline{\varepsilon}(t-1, w_h) > \hat{\varepsilon}(\boldsymbol{w}^{t-2})$$

Implying that

$$\mathbb{E}\left[d_t|w_{t-1} = w_l, \boldsymbol{w}^{t-2}\right] = \mathbb{E}\left[d_t|\varepsilon > \overline{\varepsilon}(t-1, w_l)\right] > \mathbb{E}\left[d_t|\varepsilon > \overline{\varepsilon}(t-1, w_h)\right] = \mathbb{E}\left[d_t|w_{t-1} = w_h, \boldsymbol{w}^{t-2}\right]$$

The strict inequality is guaranteed because we know that for ε at the neighborhood of $\overline{\varepsilon}(t-1, w_h)$, $d_t = 0$, since $u_t < 0$. Thus, $\mathbb{E}\left[d_t|w_{t-1} = w_l\right] > \mathbb{E}\left[d_t|w_{t-1} = w_h\right]$.

Item 2, time pattern of the dynamic selection bias with two wages: Since we only consider two wages, we denote

$$B_t(k) \equiv B_t(k, w_L, w_H) = \mathbb{E}\left[d_t | w_{t-k} = w_L\right] - \mathbb{E}\left[d_t | w_{t-k} = w_H\right]$$

Consider the following partition of the wage history at $t' = t-1, ..., t-k+1$:

$$E = \{(w_{t-1}, w_{t-2}, ..., w_{t-k+1}) : w_{t'} = w_H \text{ for all } t' = t-1, ..., t-k+1\}$$
$$E^c = \{(w_{t-1}, w_{t-2}, ..., w_{t-k+1}) : w_{t'} = w_L \text{ for some } t' = t-1, ..., t-k+1\}$$

Note that the event E is a set with only one wage history where all wages are w_H from periods $t-1$ to period $t-k+1$ and E^c is the complement of E, i.e., any wage history in the same periods that involves at least one wage w_L. Using the law of iterated expectations, we can express $B_t(k)$ as

$$\begin{aligned} B_t(k) = Pr\left(E\right) \cdot \left(\mathbb{E}\left[d_t | w_{t-k} = w_L, E\right] - \mathbb{E}\left[d_t | w_{t-k} = w_H, E\right]\right) \\ + Pr\left(E^c\right) \cdot \left(\mathbb{E}\left[d_t | w_{t-k} = w_L, E^c\right] - \mathbb{E}\left[d_t | w_{t-k} = w_H, E^c\right]\right) \end{aligned}$$

We first show that $\mathbb{E}\left[B_t(k)|E^c\right] = \mathbb{E}\left[d_t|w_{t-k} = w_L, E^c\right] - \mathbb{E}\left[d_t|w_{t-k} = w_H, E^c\right] = 0$. Note that we can re-write

$$\begin{aligned} \mathbb{E}\left[B_t(k)|E^c\right] &= \mathbb{E}\left[d_t | w_{t-k} = w_L, E^c\right] - \mathbb{E}\left[d_t | w_{t-k} = w_H, E^c\right] \\ &= \sum_{\boldsymbol{w}^t_{-(t-k)} \in E^c} Pr\left(\boldsymbol{w}^t_{-(t-k)} | E^c\right) \cdot \left(\mathbb{E}\left[d_t | w_{t-k} = w_l, \boldsymbol{w}^t_{-(t-k)}\right] - \mathbb{E}\left[d_t | w_{t-k} = w_h, \boldsymbol{w}^t_{-(t-k)}\right]\right) \end{aligned}$$

By definition, $\boldsymbol{w}^t_{-(t-k)} \in E^c$ implies that $w_{t'} = w_L$ for some $t' > t-k$. Then $\bar{\varepsilon}(t', w_L) > \bar{\varepsilon}(t-k, w_L) > \bar{\varepsilon}(t-k, w_H)$, so $\mathbb{E}\left[d_t | w_{t-k} = w_L, \boldsymbol{w}^t_{-(t-k)}\right] = \mathbb{E}\left[d_t | w_{t-k} = w_H, \boldsymbol{w}^t_{-(t-k)}\right]$ (Lemma 2). Then, we conclude that for any $\boldsymbol{w}^t_{-(t-k)} \in E^c$,

$$\mathbb{E}\left[d_t | w_{t-k} = w_l, \boldsymbol{w}^t_{-(t-k)}\right] - \mathbb{E}\left[d_t | w_{t-k} = w_h, \boldsymbol{w}^t_{-(t-k)}\right] = 0$$

implying that $\mathbb{E}\left[B_t(k)|E^c\right] = 0$. Therefore, we have that

$$B_t(k) = Pr\left(E\right) \cdot \left(\mathbb{E}\left[d_t | w_{t-k} = w_L, E\right] - \mathbb{E}\left[d_t | w_{t-k} = w_H, E\right]\right)$$

Because we assume that w_t is iid, then $Pr(w_t = w_H) = p_H$ for every t and $Pr(E) = p_H^{k-1}$[11]. We then have that

$$\begin{aligned} B_t(k) = p_H^{k-1} \cdot \left(\mathbb{E}\left[d_t | w_{t-k} = w_L, w_{t-k+1} = w_H, ..., w_{t-1} = w_H\right]\right. \\ \left. - \mathbb{E}\left[d_t | w_{t-k} = w_H, w_{t-k+1} = w_H, ..., w_{t-1} = w_H\right]\right) \\ = p_H^{k-1} \cdot \left(\mathbb{E}\left[d_t | \varepsilon > \max\{\bar{\varepsilon}(t-k, w_L), \bar{\varepsilon}(t-1, w_H)\}\right] - \mathbb{E}\left[d_t | \bar{\varepsilon}(t-1, w_H)\right]\right) \end{aligned}$$

Note that if $\bar{\varepsilon}(t-k, w_L) \leq \bar{\varepsilon}(t-1, w_H)$, then the two expectations cancel out and $B_t(k) = 0$. However, if

[11] The exponent is $k-1$ because the wage history goes from $t-k+1$ to $t-1$

$\bar{\varepsilon}(t-k, w_L) > \bar{\varepsilon}(t-1, w_H)$, then

$$\mathbb{E}\left[d_t|\varepsilon > \max\{\bar{\varepsilon}(t-k, w_L), \bar{\varepsilon}(t-1, w_H)\}\right] - \mathbb{E}\left[d_t|\bar{\varepsilon}(t-1, w_H)\right] > 0$$

So, we can express

$$B_t(k) = \mathbb{1}_{[\bar{\varepsilon}(t-k,w_L)>\bar{\varepsilon}(t-1,w_H)]} \cdot p_H^{k-1} \cdot \left(\mathbb{E}\left[d_t|\varepsilon > \bar{\varepsilon}(t-k, w_L)\right] - \mathbb{E}\left[d_t|\varepsilon > \bar{\varepsilon}(t-1, w_H)\right]\right)$$

Analogously, for $k+1$, the bias will be

$$B_t(k+1) = \mathbb{1}_{[\bar{\varepsilon}(t-k-1,w_L)>\bar{\varepsilon}(t-1,w_H)]} \cdot p_H^{k} \cdot \left(\mathbb{E}\left[d_t|\varepsilon > \bar{\varepsilon}(t-k-1, w_L)\right] - \mathbb{E}\left[d_t|\varepsilon > \bar{\varepsilon}(t-1, w_H)\right]\right)$$

Note that

$$\begin{aligned}
&\mathbb{1}_{[\bar{\varepsilon}(t-k,w_L)>\bar{\varepsilon}(t,w_H)]} \geq \mathbb{1}_{[\bar{\varepsilon}(t-k-1,w_L)>\bar{\varepsilon}(t,w_H)]} \\
&Pr(w_{t'} = w_H)^{k-1} > Pr(w_{t'} = w_H)^{k} \\
&\mathbb{E}\left[d_t|\varepsilon > \bar{\varepsilon}(t-k, w_L)\right] > \mathbb{E}\left[d_t|\varepsilon > \bar{\varepsilon}(t-k-1, w_L)\right]
\end{aligned}$$

The first and third inequalities come from the fact that $\bar{\varepsilon}(t, w_L)$ is increasing in t (Lemma 1). These three inequalities imply that $B_t(k, w_L, w_H) \geq B_t(k+1, w_L, w_H)$.

Item 2, time pattern of the dynamic selection bias with three or more wages: We now show that, with three or more wages, it is possible for the bias at $t-k-1$ to be larger than the bias at $t-k$. We proceed by example and make a series of assumptions to impose the necessary structure for the result.

We iterate the value function as follows: at T, the last possible period of a shift, we set $V_T = 0$ as agents are forced to stop. We then proceed by solving the value function backwards, considering the future value as a expected value over the possible stochastic wages. To take expectations over the possible future wages, we simulate $1,000,000$ wage paths.

Before placing a specific structure on the utility function, consider the general format for the bias at times $t-k$ and $t-k-1$ (all expectations are conditional on $d_{t-1} = 1$:

$$\begin{aligned}
B_t(k, w_H, w_L) &= E[d_t|w_{t-k} = w_L] - E[d_t|w_{t-k} = w_H] \\
B_t(k+1, w_H, w_L) &= E[d_t|w_{t-k-1} = w_L] - E[d_t|w_{t-k-1} = w_H]
\end{aligned}$$

Taking the difference between these bias, we have:

$$\Delta B(w_H, w_L) := \left(E[d_t|w_{t-k} = w_L] - E[d_t|w_{t-k} = w_H]\right) - \left(E[d_t|w_{t-k-1} = w_L] - E[d_t|w_{t-k-1} = w_H]\right) \quad (13)$$

As we have shown, with two wages, $\Delta B(w_H, w_L) \geq 0$. With three or more wages, there are possible future paths of wages with wages above w_H and w_L, and these paths can act as to eliminate the bias stemming from $t-k$, while not necessarily eliminating those from $t-k-1$. However, the likelihood of such an event depends on specific formats for the utility function. Intuitively, wages closer to t must be less relevant for the labor supply decisions, as to eliminate the bias arising from wage differences at periods closer to t.

Based on this observation, we assume the following format for the utility function, common to all workers:

$$u(w, t, \varepsilon) = 4w - 0.2(t-1)^{1.01} - 0.49(w(t-1))^{0.75} + \varepsilon_t$$

There are two key ingredients in this utility function. First, the cost of effort function $(0.2(t-1)^{1.01})$ is only mildly convex, which translates into a smaller decay of bias. If the cost function is more strongly convex, the immediate previous time period has a much larger cost than those in the past, and thus selection is going to come almost entirely from $t-1$. Second, the interaction term $-w(t-1)$ guarantees that, as the shift advances, wages are less relevant to utility. This is necessary given our previous observation that, for $\Delta B(w_H, w_L)$ to be negative, the recent lag must be small - thus, wages must be more important to the stopping decision farther away in time. Third, the interaction term is concave, which follows a similar reasoning to the mild convexity of the cost of effort.

We assume that there are three wage levels, $w = \{0.2, 0.5, 0.8\}$, named w_L, w_M, w_H, respectively, and that there are at most $T = 20$ periods in a given shift. Then, assume that ε varies only at the shift-level (i.e., $\varepsilon_t = \varepsilon$), and that $\varepsilon \sim U[0, 2]$.

Table A1 presents the bias $B(k, w_j, w_i)$, for period $t = 12$, for all three possible wage combinations and $k \in [1, 4]$.

Table A1: Simulated dynamic selection bias, at $t = 12$, for four lags

Lag (k)	$B(k, w_H, w_M)$	$B(k, w_H, w_L)$	$B(k, w_M, w_L)$
1	0.111	0.146	0.035
2	0.023	0.095	0.072
3	0.003	0.0127	0.009
4	-0.002	-0.001	0.001

Thus, we can see that the bias from wages w_M to w_L at lag $k = 2$ is larger than the bias from $k = 1$. This proves the possibility of greater bias for latter lags.

□

Proof of Proposition 2

If $w_{t-1} = \underline{w}$, then by lemma 1 $\overline{\varepsilon}(t-1, \underline{w}) = \hat{\varepsilon}(\boldsymbol{w}^{t-1})$, since $\overline{\varepsilon}(t, w)$ is increasing in t and decreasing in w. By lemma 2, this implies that

$$\mathbb{E}\left[d_t | d_{t-1} = 1, w_{t-1} = \underline{w}, \mathbf{w}^{t-2}\right] = \mathbb{E}\left[d_t | \varepsilon > \overline{\varepsilon}(t-1, \underline{w})\right]$$

for any $\mathbf{w}^{t-2}$. □

B. Additional Results

A. Additional Details Model

Dynamic Model. At t, the agent chooses $\boldsymbol{d^t} = (d_t, d_{t+1}, ...)$ to maximize their intertemporal utility

$$U(\boldsymbol{d^t}, t, \varepsilon) = d_t u(t, w_t, \varepsilon) + \sum_{t'=t+1}^{\infty} \mathbb{E}\left[d_{t'}(w_{t'}, \varepsilon)u(t', w_{t'}, \varepsilon)\right]$$

where $d_t = 0$ is absorbing (upon quitting, the current shift is ended). The agent's problem is equivalent to the Bellman equation

$$V(t, w_t, \varepsilon) = \max_{d_t \in \{0,1\}} \{d_t u(t, w_t, \varepsilon) + d_t \mathbb{E}\left[V(t+1, w_{t+1}, \varepsilon)\right]\} \tag{14}$$

implying that they stops working at t iff

$$u(t, w_t, \varepsilon) + \mathbb{E}\left[V(t+1, w_{t+1}, \varepsilon)\right] < 0 \tag{15}$$

Technical Assumption. To avoid dealing with corner solutions, we assume that for any t and w, $\exists\, \tilde{\varepsilon}$ such that $u(t, w, \tilde{\varepsilon}) = 0$. This is equivalent to assuming probabilities of stopping strictly between 0 and 1, making the algebra less cumbersome without affecting results qualitatively.

B. Controlling Dynamic Selection by Fixed Effects

Instead of using the approach proposed in Proposition 2, we could estimate the hazard model controlling for fixed effects. If the unobserved heterogeneity varies across shifts for a given worker, then we would need to add fixed effects in the shift-worker level. Such high-dimensional control would lead to the incidental parameter problem (Lancaster, 2000). This problem could be avoided if the hazard rate is linear function of the error term, but this gives rise to another issue. Since a linear model is unlikely to represent well a binary outcome variable, model mispecification could be severe, and fixed effects would not deal with the unobserved heterogeneity.

We assess how shift-level fixed effects might help our results in two ways. First, we estimate the results in Table 1 using shift-worker FEs, and they do not solve the issue of dynamic selection in our setting. Table C8 shows that the difference in total income effect between the standard estimates (columns 1-3) are very similar to the estimates controlling for participant-day fixed effects (columns 4-6) for output, work breaks, and stopping decision.

Second, we conduct a simulation exercise that shows that a linear fixed effect model does not help with dynamic selection while the solution outlines at Proposition 2 does. We simulate a sample of 4520 workers (we use a sample 10 times higher than our own to minimize simulation noise) deciding when to stop working based on the model outlined in Section I. For each worker-day pair, we draw a heterogeneity value ϵ from a $[0, 1]$-uniform distribution and a wage path. We assume that the work day is divided in 20 sessions, so each wage path has 20 elements. We assume the instantaneous utility function is separable:

$$u(t, w, \epsilon) = \phi w - \beta(t-1) + \epsilon$$

and we set $\phi = 5$ and $\beta = 0.1$. Just like in our setting we assume there are only 2 different wages and set the low wage to 0.5 and the high wage to 2, also replicating the ratio between high and low wages in our setting.

With this model, we simulate a data-set in the incentive-session level that shows for each worker-day-session, what was the piece rate and whether the worker decided to keep working ($y_{ids} = 1$) or if the worker stopped working ($y_{ids} = 0$).

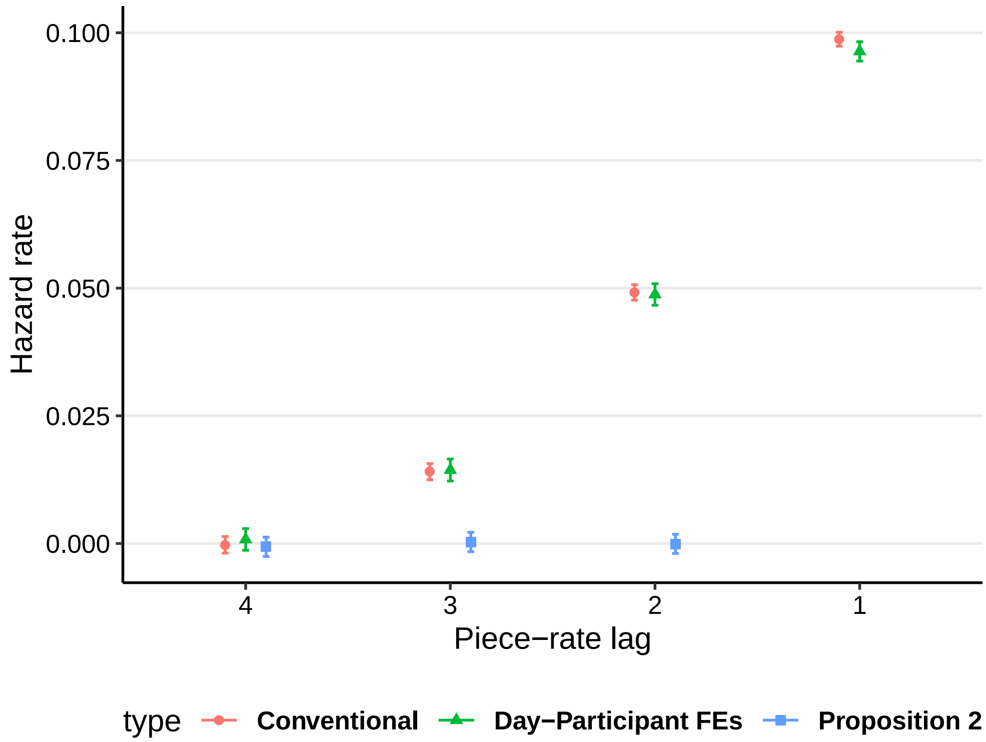
We estimate three variants of equation 3 using the simulate data set. First, we estimate a model with no controls and full sample. This is supposed to be the benchmark result, which mimics the estimation procedure commonly used in the literature.[12] Second, we estimate the same model, still with the full sample, but adding participant-session fixed effects. This model controls for fixed-effects in the same level as the unobserved heterogeneity varies, which perhaps indicate it could suffice to deal with dynamic selection. Third, we estimate the equation without controls, but using only the sample of sessions in which the first lag of the piece rate is the low piece rate. This is our proposed solution to the dynamic selection issue, outlined in Proposition 2.

The simulation exercise (Figure B1) has three takeaways. First, the benchmark estimates presents a positive relationship between past wages and the probability of stopping the shift. Moreover, this relationship is stronger for more recent lags of piece rates. This is precisely the prediction of the income targeting model with adaptive reference-point formation Thakral and Tô (2020). Note however that in this simulation, this is coming exclusively from dynamic selection. Second, we show that including participant-day fixed effects do not change the income effects estimates qualitatively. Thus, adding fixed effects in this example do not help with dynamic selection. Third, the solution outlined in Proposition 2 completely solves the dynamic selection issue. In the sub-sample not affected by dynamic selection, all lags of piece rate are very close to zero and not significant.[13]

[12] Note that in the literature, researchers often add a series of controls, including participant fixed effects. We do not here for simplicity, since in our simulated data we rule out correlations between participant-level covariates and the outcome variable by design.

[13] Note that with our proposed solution, we cannot estimate the first piece-rate lag. That is why we do not add it to the graph.

Figure B1: Simulation of participant-day fixed effects versus Proposition 2 solution to dynamic selection



Notes: This figure shows the estimates from three income effect estimators using data simulated from our model in Section I. Each series is a variation of equation 3 using the simulate data set: The red (circle) series use estimates from a model with no controls and full sample. This is the benchmark result, which mimics the estimation procedure commonly used in the literature. The green series (triangle), stills uses the full sample, but controls for participant-session fixed effects. The blue (square) series, shows estimates of the equation without controls, but using only the sample of sessions in which the first lag of the piece rate is the low piece rate. This is our proposed solution to the dynamic selection issue, outlined in Proposition 2. Note that we cannot estimate the first piece-rate lag in the last series. The bars represent 1% confidence intervals.

C. Monetary Windfall

We corroborate the results from the previous subsection using an exogenous shock to income that cannot generate dynamic selection.

The participants performed two cognitive tasks each day, receiving daily payments for their performance. Together, performing both tasks took around 20 minutes In each day, each of these tasks had a 50% probability of having their payments doubled, which was observed by the participants. Right before starting the cognitive task (see Section A), participants were told whether their payment would be doubled. Since participants could not leave during the task, they could not adjust their decision to stop working as a response to the payment before or during the task, disallowing dynamic selection. However, they could adjust labor supply afterward. This represents a positive income shock of about 6% to their cumulative daily income *per payment doubled*, bringing them closer to their daily income target.

We estimate the regression

$$\mathrm{y}_{id} = \beta \ \mathrm{DoublePay}_{id} + \gamma X_{id} + \varepsilon_{id} \tag{16}$$

where i represents a participant and d a day. y_{id} captures one of the following margins of labor supply: effort, work breaks, or the time in hours a participant decides to quit working. The exogenous variation in income is captured by the dummy $\mathrm{DoublePay}_{id}$, which indicates whether the participant received a High payment on either of the two cognitive tasks. X_{idt} is a vector of covariates capturing participant, date, day in study, and window fixed effects.

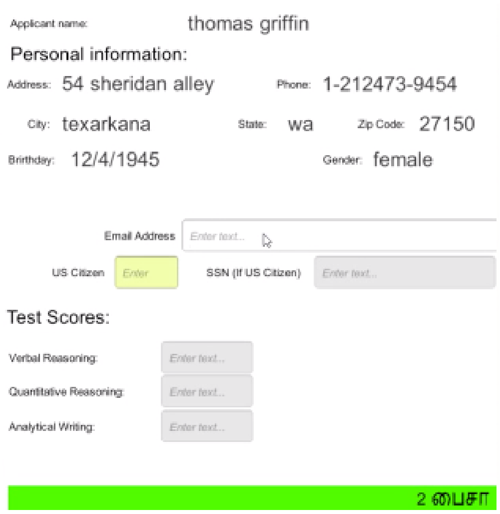
Table C5 shows no evidence of income targeting. The effect on output (col. 1), work breaks (col. 2), and the time participants stopped working (col. 3) are small and insignificant, in spite of high detection power.[14] The sign for work breaks is actually flipped, suggesting fewer breaks after a doubled payment.

These results complement the findings of Dupas et al. (2020) and Andersen et al. (2018). They also find no evidence of income effects using random variation in one-time payments. They use a monetary windfall as an income shifter. In contrast, the income variation from cognitive tasks in our setting is part of the participants' day-to-day work. This is important because workers might use a separate mental account for unearned, unexpected monetary windfalls (Henderson and Peterson, 1992). In that case, the monetary windfall may not count for achieving the income target. Our paper addresses this concern and corroborates their findings of no daily income effects.

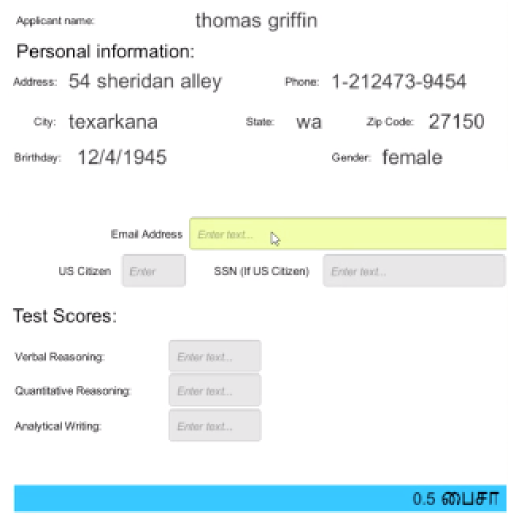
[14] We could reject changes in output of 3%, in work breaks of 42 seconds, and in stopping time of 3.6 minutes with 95% confidence.

C. Supplementary Tables and Figures

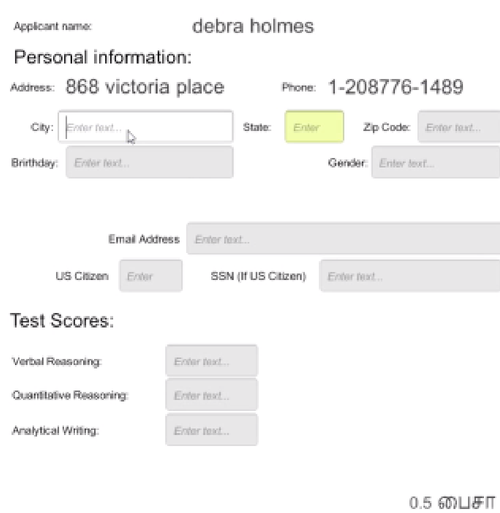
Figure C2: Screens and piece-rates under salient and non-salient conditions



(a) Salient high piece-rate



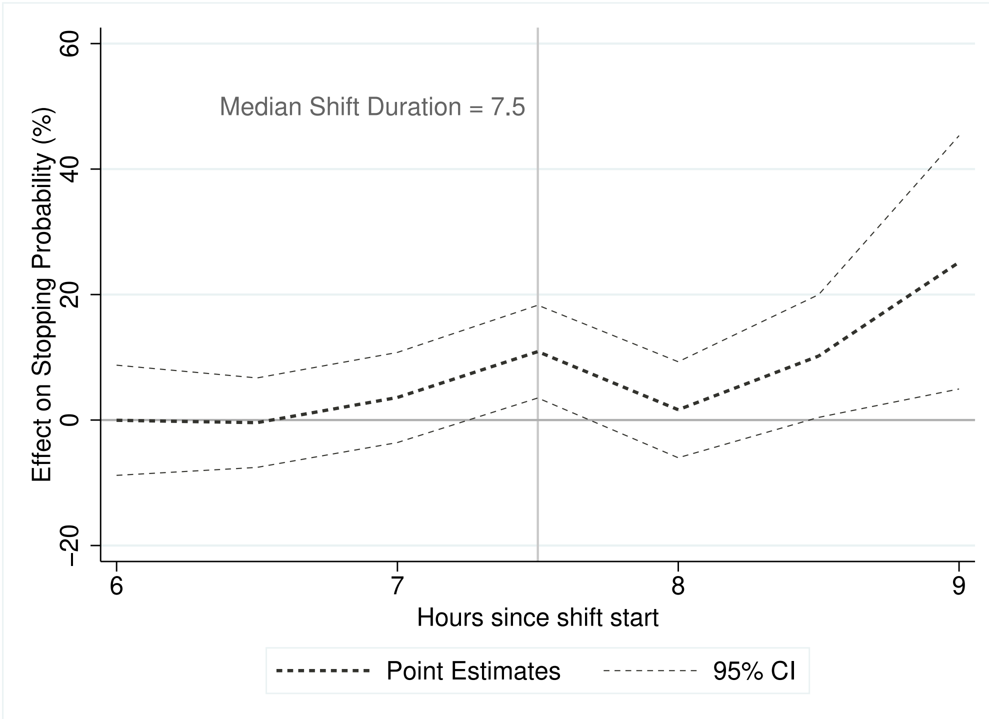
(b) Salient low piece-rate



(c) Non-salient piece-rate

Notes: This figure provides an example of the participants' computer screen, where they entered the data in their data-entry work. The participants submit the data in these fields sequentially and cannot edit a field after submission nor skip a field. After 30 minutes typing in a session, a new session begins when the participant submits the current field of data they are entering. When the session changes, the piece rates may change as well. In (a), the piece-rate for the current session is High (Rs. 2 per 100 characters), while (b) reproduces the screen for Low piece-rates (Rs. 0.5 per 100 characters). Panel (c) shows a low piece-rate during the non-salient condition. In the non-salient condition, there is no difference in color between the Low and High piece-rates.

Figure C2: Estimates of the elasticity of stopping decision in response to a 10% increase in accumulated earnings



Notes: This figure presents estimates of the following regression (Thakral and Tô, 2020):

$$d_{idp} = \sum_{j} \mathbb{1}\{h_{idp} \in H_j\}(\alpha_j h_{idp} + \beta_j I_{idp} + \gamma_j X_{idp}) + \varepsilon_{idp} \quad (17)$$

where i is an individual, d a date, and p is a form page. The indicator function represents the partition of the data by accumulated hours working. We use a partition of 30-minute windows. h_{idp} represents cumulative hours in the office (including typing and break times), and I_{idp} represents cumulative income, which is instrumented by the share of High sessions among all preceding sessions, not including the current session. X_{idp} are controls, which include: worker, date, day-in-study, and session fixed effects; a dummy for whether the current session has high piece rates; and hour-of-day indicators interacted with worker indicators. Standard errors are clustered at the worker-level.

In the figure, estimates β_j are multiplied by 10% of the cumulative earnings and divided by the average hazard rate in the window H_j. They represent the reactivity of the stopping decision to an increase of 10% in cumulative earnings.

Table C2: Summary statistics at participant, shift and session level

Panel A: Participant-level

	Mean	Std. Deviation	Median	p5	p25	p75	p95
Age	34.95	7.22	33.00	27.00	29.00	40.00	50.00
Years of Education	10.18	2.88	10.00	5.00	9.00	12.00	15.00
Number of Children	1.35	1.06	2.00	0.00	0.00	2.00	3.00
Female	0.66	0.47	-	-	-	-	-
Used Computer Before	0.28	0.45	-	-	-	-	-
Observations	452						

Panel B: Shift-level (long days)

	Mean	Std. Deviation	Median	p5	p25	p75	p95
Output	17186.19	11777.57	14501.00	3983.00	9141.00	21930.00	40618.00
Typing Time (Hr.)	5.22	1.38	5.30	3.01	4.45	6.15	7.29
Productivity (Output/Min)	54.03	30.53	47.65	17.60	31.98	68.53	117.36
Voluntary Pauses (Min.)	15.56	22.74	8.83	0.00	3.40	18.82	53.67
Scheduled Pauses (Min.)	110.07	25.58	115.23	77.03	95.00	125.00	154.80
Typing Earnings	339.49	179.36	302.26	133.24	221.95	417.46	693.01
Number of Incentive Sessions	10.91	2.47	11.00	7.00	9.00	13.00	15.00
Time in Office	8.04	1.17	8.07	6.18	7.40	8.77	9.83
Observations	4945						

Panel C: Session-level (regular days)

	Mean	Std. Deviation	Median	p5	p25	p75	p95
Output	1588.02	977.87	1404.00	343.00	895.00	2055.00	3558.00
Productivity (Output/Min)	55.20	32.24	48.46	16.50	32.00	70.35	120.81
Typing Time (Min.)	29.13	7.26	29.92	13.75	29.58	30.77	34.53
Voluntary Pauses (Min.)	1.47	5.97	0.00	0.00	0.00	0.07	7.80
Scheduled Pauses (Min.)	10.20	17.10	0.00	0.00	0.00	15.00	40.00
Typing Earnings	31.42	20.80	24.37	9.28	16.81	42.30	73.09
Performance Earnings	20.81	20.29	13.32	1.26	5.68	31.22	62.12
Attendance Earnings	10.61	2.50	10.84	5.00	10.76	11.20	12.55
Observations	54286						

Notes: This table presents summary statistics for three different units of observations: participants, shifts, and incentive sessions. The panels in the shift and incentive session levels only include regular days (see section A). p5, p25, p75, and p95 are the 5th, 25th, 75th, and 95th percentile of the variables, respectively.

Table C3: Earnings variation explained by piece-rates

Panel A: Session-level earnings

	(1) Earnings	(2) Earnings	(3) Earnings	(4) Earnings
Mean squared error	434.1	227.3	314.6	108.3
R^2	0.00	0.48	0.28	0.75
High piece-rate dummy		X		X
Worker fixed effects			X	X
Observations	53194	53194	53194	53194

Panel B: Cumulative earnings (Y_{t-1})

	(1) Cum. Earnings	(2) Cum. Earnings	(3) Cum. Earnings	(4) Cum. Earnings
Mean squared error	23702.8	7896.2	17759.9	3711.8
R^2	0.00	0.67	0.26	0.84
Cumulative number of High piece-rates		X		X
Worker fixed effects			X	X
Observations	53194	53194	53194	53194

Notes: This table presents the amount of earnings variation explained by piece-rate variation. In panel A, the variable of interest is earnings at the session-level, while panel B focuses on cumulative earnings until session $t-1$. Overall, the piece-rates alone explain 48% of the earnings in a session and 67% of cumulative earnings (column 1). It also has significantly more explanatory power than worker-specific intercepts (comparing columns 2 and 3). Together, worker-specific intercepts and piece-rates explain 75% of session-level earnings and 84% of cumulative earnings (column 4). These results indicate that piece-rates can provide reasonable inferences about earnings in a shift, while also being easily trackable (see Figure C2 and Table C6 for evidence that piece-rates were easily observable and workers did, in fact, observe them). This exercise is partly conservative since one component of earnings is a fixed hourly-rate, which is constant throughout the whole study and linear in the number of hours worked, making it easy to track by workers.

Table C4: Effects of piece-rate variation on labor supply

	Regular days			Regular, conditional on w_{t-1} = low			Regular, conditional on w_{t-1} = high		
	(1) Output	(2) Work Breaks	(3) Quit	(4) Output	(5) Work Breaks	(6) Quit	(7) Output	(8) Work Breaks	(9) Quit
Total effect: lags 1-4	-98.45	0.88	0.04						
	(12.31)	(0.19)	(0.01)						
	[0.00]	[0.00]	[0.00]						
Total effect: lags 2-4	-42.16	0.51	0.01	2.63	0.03	-0.01	-77.75	0.90	0.03
	(9.49)	(0.14)	(0.01)	(12.05)	(0.11)	(0.01)	(14.13)	(0.24)	(0.01)
	[0.00]	[0.00]	[0.05]	[0.83]	[0.77]	[0.14]	[0.00]	[0.00]	[0.00]
High Lag 1	-56.292	0.367	0.031						
	(6.172)	(0.083)	(0.004)						
	[0.00]	[0.00]	[0.00]						
High Lag 2	-18.225	0.219	0.008	15.469	-0.068	-0.005	-42.293	0.476	0.018
	(5.120)	(0.065)	(0.003)	(6.405)	(0.075)	(0.005)	(7.907)	(0.115)	(0.005)
	[0.00]	[0.00]	[0.01]	[0.02]	[0.36]	[0.28]	[0.00]	[0.00]	[0.00]
High Lag 3	-11.453	0.127	0.000	-6.982	0.021	-0.001	-16.482	0.191	0.001
	(4.616)	(0.068)	(0.003)	(6.282)	(0.070)	(0.004)	(7.558)	(0.122)	(0.005)
	[0.01]	[0.06]	[0.94]	[0.27]	[0.76]	[0.79]	[0.03]	[0.12]	[0.77]
High Lag 4	-12.482	0.167	0.003	-5.852	0.081	-0.006	-18.974	0.231	0.011
	(5.607)	(0.069)	(0.003)	(6.157)	(0.066)	(0.004)	(8.340)	(0.115)	(0.005)
	[0.03]	[0.02]	[0.31]	[0.34]	[0.23]	[0.15]	[0.02]	[0.04]	[0.03]
High	284.706	-1.060	-0.083	261.394	-0.842	-0.077	309.734	-1.267	-0.089
	(12.128)	(0.100)	(0.004)	(12.226)	(0.090)	(0.005)	(14.230)	(0.154)	(0.005)
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Excluded Group Mean	1470.75	1.87	0.14	1512.75	1.70	0.11	1416.59	2.10	0.18
Excluded Group SD	941.48	6.24	0.34	925.67	5.00	0.31	958.83	7.53	0.38
Observations	33716	33716	33716	16181	16181	16181	17532	17532	17532
Participants	452	452	452	452	452	452	452	452	452

Notes: This table presents estimates of Equation (3) for three labor supply outcomes: output, work breaks, and a dummy indicating the decision to quit for the day. All estimates use the sample of regular days, when participants had full discretion to choose when to quit, replicating the same columns in Table 1. Columns 1-3 use the entire sample. In columns 4-6, we restrict the sample to sessions for which the previous session had a Low piece rate, addressing dynamic selection (see Proposition 2). In columns 7-9, we restrict the sample to sessions for which the previous session had a High piece rate. Row 1 computes the sum of the four lags, while row 2 considers the effect of lags 2-4 to ensure comparability between columns 1-3 and the others. "Excluded group" stands for the sample where all piece-rate indicators included in the regression are equal to zero. Standard errors clustered at the participant level are displayed in parenthesis, with the associated p-values in brackets.

Table C5: Effect of earnings variation from cognitive tasks on labor supply

	(1) Output (post-task)	(2) Voluntary pauses (post-task)	(3) Time stopped typing (hr.)
DoublePay	-8.246	-0.399	-0.037
	(109.334)	(0.354)	(0.043)
	[0.94]	[0.26]	[0.39]
Control Group Mean	6657.76	6.25	17.78
Control Group SD	5596.49	15.84	1.32
Observations	7416	7416	4912
Participants	452	452	452

Notes: This table presents estimates from Equation (16) for three labor supply outcomes: output, voluntary pauses, and the hour of day a participant decides to quit typing. "DoublePay" is a dummy capturing whether the payment for any of the cognitive tasks was doubled. We only consider outcomes after the cognitive task has taken place. The sample includes observations for both short and regular days, except for column 3, which includes only regular days, when participants had full discretion to choose when to quit (see Section A). Working days in which participants did not complete the cognitive task are excluded from the sample (33 observations). Standard errors clustered at the worker level are displayed in parenthesis, with the associated p-values in brackets.

Table C6: Effect of piece-rate variation on labor supply - differences by piece-rate salience

	(1) Output	(2) Work Breaks	(3) Quit
High Lag 1 × Salience	-16.384	-0.113	0.004
	(12.615)	(0.152)	(0.008)
	[0.19]	[0.46]	[0.60]
High Lag 2 × Salience	-15.370	0.085	0.007
	(11.788)	(0.162)	(0.007)
	[0.19]	[0.60]	[0.32]
High Lag 3 × Salience	24.617	0.097	-0.010
	(10.738)	(0.162)	(0.007)
	[0.02]	[0.55]	[0.18]
High Lag 4 × Salience	11.635	-0.156	-0.001
	(10.548)	(0.165)	(0.007)
	[0.27]	[0.34]	[0.88]
High	292.030	-1.046	-0.083
	(14.569)	(0.145)	(0.006)
	[0.00]	[0.00]	[0.00]
Salience	-6.579	0.055	0.001
	(15.747)	(0.173)	(0.009)
	[0.68]	[0.75]	[0.86]
High × Salience	18.232	-0.154	-0.008
	(12.431)	(0.160)	(0.007)
	[0.14]	[0.34]	[0.23]
High Lag 1	-50.249	0.438	0.030
	(9.638)	(0.134)	(0.006)
	[0.00]	[0.00]	[0.00]
High Lag 2	-11.693	0.164	0.003
	(8.635)	(0.135)	(0.006)
	[0.18]	[0.23]	[0.56]
High Lag 3	-25.916	0.019	0.004
	(7.844)	(0.125)	(0.005)
	[0.00]	[0.88]	[0.48]
High Lag 4	-23.285	0.230	0.006
	(8.229)	(0.133)	(0.005)
	[0.00]	[0.08]	[0.29]
Excluded Group Mean	1562.35	1.89	0.14
Excluded Group SD	953.44	6.63	0.34
Observations	27654	27654	27654
Participants	449	449	449

Notes: This table reproduces the regressions from Equation (3) interacting the piece-rate indicators with variables capturing whether a shift has salient or non-salient piece-rates (for a visual comparison between salient and non-salient piece-rates, see Figure C2). The coefficients of interest are the difference-in-differences interaction coefficients displayed in rows 1-4. Taken together, the interactions shows that there is little difference between reactions to lagged piece rates under salient and non-salient conditions. If piece rates are negatively related to labor supply and salience improves their observation, we should expect greater magnitudes in coefficients (in this case, more negative coefficients). Not only the signs are overall mixed, but the estimates are also statistically insignificant. The difference in number of observations when compared to Table 1 of the main body is due to: i) the randomization of the salient/non-salient conditions only starts at day 6 of the experiment; ii) days 6, 7 and 8 are special days and thus excluded from the sample (see section A). The difference in number of participants when compared to Table 1 is due to three workers being present only until the 8^{th} day of the study. Errors clustered at the worker-level are displayed in parenthesis, with the associated *p*-values in brackets.

Table C7: Effect of piece-rate variation on labor supply - first week versus last week

	(1) Output	(2) Work Breaks	(3) Quit	(4) Output	(5) Work Breaks	(6) Quit
Total effect: lags 1-4	-74.27	0.60	0.03	-90.66	0.62	0.03
	(18.77)	(0.29)	(0.01)	(31.20)	(0.26)	(0.02)
	[0.00]	[0.04]	[0.02]	[0.00]	[0.02]	[0.17]
High Lag 1	-42.79	0.198	0.023	-54.83	0.298	0.024
	(7.89)	(0.101)	(0.006)	(12.96)	(0.128)	(0.007)
	[0.00]	[0.05]	[0.00]	[0.00]	[0.02]	[0.00]
High Lag 2	-14.57	0.209	0.005	-19.29	0.146	0.004
	(7.82)	(0.098)	(0.005)	(11.31)	(0.090)	(0.007)
	[0.06]	[0.03]	[0.36]	[0.09]	[0.10]	[0.61]
High Lag 3	-8.79	0.153	-0.001	-6.10	0.220	-0.001
	(7.51)	(0.107)	(0.006)	(11.13)	(0.119)	(0.008)
	[0.24]	[0.15]	[0.81]	[0.58]	[0.06]	[0.92]
High Lag 4	-8.11	0.038	0.002	-10.44	-0.045	0.001
	(7.40)	(0.100)	(0.006)	(12.06)	(0.108)	(0.007)
	[0.27]	[0.71]	[0.75]	[0.39]	[0.68]	[0.89]
High	232.32	-0.823	-0.077	344.754	-1.128	-0.090
	(12.81)	(0.119)	(06)	(17.337)	(0.130)	(0.007)
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Excluded Group Mean	1295.24	1.76	0.13	1616.37	1.95	0.14
Excluded Group SD	890.78	5.69	0.34	937.00	5.85	0.35
Observations	12887	12887	12887	10092	10092	10092
Participants	452	452	452	426	426	426

Notes: This table presents estimates of Equation 3 for three labor supply outcomes: output, work breaks, and a dummy indicating the decision to quit for the day. Columns 1 to 3 restrict the sample to the first seven regular days, when participants had full discretion to choose when to quit (see section A). This covers days 3 through 12 (days 6, 7 and 8 are "special days"). Columns 4 to 6 uses only the last seven regular days, covering days 20 through 27, except for day 26 (days 26 and 28 are "special days"). Row 1 presents the sum of the four lags. "Excluded group" stands for the sample where all piece-rate indicators included in the regression are equal to zero. Standard errors clustered at the participant-level are displayed in parenthesis, with the associated p-values in brackets.

Table C8: Effect of piece-rate variation on labor supply - controlling for worker-date fixed effects

	(1) Output	(2) Work Breaks	(3) Quit	(4) Output	(5) Work Breaks	(6) Quit
Total effect: lags 1-4	-98.45	0.88	0.04	-73.43	0.85	0.02
	(12.31)	(0.19)	(0.01)	(15.75)	(0.22)	(0.01)
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.08]
High Lag 1	-56.29	0.367	0.031	-45.65	0.372	0.022
	(6.17)	(0.083)	(0.004)	(6.81)	(0.086)	(0.004)
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
High Lag 2	-18.22	0.219	0.008	-12.79	0.217	0.003
	(5.12)	(0.065)	(0.003)	(5.75)	(0.078)	(0.004)
	[0.00]	[0.00]	[0.01]	[0.03]	[0.01]	[0.50]
High Lag 3	-11.45	0.127	0.000	-5.37	0.117	-0.005
	(4.62)	(0.068)	(0.003)	(5.36)	(0.069)	(0.004)
	[0.01]	[0.06]	[0.94]	[0.32]	[0.09]	[0.20]
High Lag 4	-12.48	0.167	0.003	-9.61	0.146	-0.001
	(5.61)	(0.069)	(0.003)	(5.95)	(0.073)	(0.004)
	[0.03]	[0.02]	[0.31]	[0.11]	[0.05]	[0.84]
High	284.71	-1.060	-0.083	287.98	-1.032	-0.086
	(12.13)	(0.100)	(0.004)	(12.53)	(0.099)	(0.004)
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Worker-Date FE				X	X	X
Excluded Group Mean	1470.75	1.87	0.14	1470.75	1.87	0.14
Excluded Group SD	941.48	6.24	0.34	941.48	6.24	0.34
Observations	33716	33716	33716	33676	33676	33676
Participants	452	452	452	452	452	452

Notes: This table presents estimates of Equation 3 for three labor supply outcomes: output, work breaks, and a dummy indicating the decision to quit for the day. Columns 1 to 3 restrict the sample to regular days, when participants had full discretion to choose when to quit (see section A). Columns 4 to 6 also uses only regular days, but includes worker-date indicators as controls. Row 1 presents the sum of the four lags. "Excluded group" stands for the sample where all piece-rate indicators included in the regression are equal to zero. Standard errors clustered at the participant level are displayed in parenthesis, with the associated p-values in brackets.