For Online Publication

Appendix A. Further Discussion

A.1. *Uncertainty about Workload*

In the main body of the paper, the agent knows task features $\mathcal{T} = (w, T)$. However, in many cases, the agent is uncertain about the task difficulty w due to lack of experience and thus undergoes learning by doing. In this appendix, I analyze a simple extension in which a potentially present-biased and naive agent needs to complete a long-term task before a fixed deadline T but she is uncertain about the task's difficulty w.

In this scenario, workload uncertainty may well be another plausible source for procrastination. A natural question to ask is: to what degree does an agent procrastinate due to her innate inclination to put off work; and to what degree is the procrastination simply caused by her uncertainty about the task workload? The separation of preference and belief has always been an intriguing problem in empirical studies and has accrued growing attention. To this end of identifying the source of procrastination, we need a model that incorporates present bias, naivete, and workload uncertainty.

In *Appendix A.1.1*, I first characterize the agent's work schedule and individual welfare in the binary case where the workload can be either low or high. I then examine the empirical content of the model in *Appendix A.1.2* and the value of short-term goals *Appendix A.1.3*. In *Appendix A.1.4*, I generalize the prior belief on the workload.

A.1.1. *Binary Case.* The workload can take two values, i.e., $w \in \{w_L, w_H\}$ with $w_H > w_L > 0$. The agent has the prior that $w = w_L$ with probability $\mu \in [0,1]$. Since the agent gets the reward once she completes the task, the workload uncertainty is naturally resolved after she finishes the workload w_L: if she attains the reward, she learns $w = w_L$; otherwise, she learns $w = w_H$ and continues to complete the work left $w_H - w_L$ within the remaining time. Thus, completing w_L reveals the workload information to the agent.

Let $\mathcal{T} = \langle w_L, w_H, \mu, T\rangle$ represent task features and let $\mathcal{B} = \langle \beta, \hat{\beta}\rangle$ represent behavioral frictions of the agent. *Proposition A.1* characterizes the agent's work schedule and effort costs to complete a task before the deadline.

PROPOSITION A.1 (Work Schedule and Effort Costs for an Uncertain Task)**.**
Let $B = (\beta/\hat{\beta})^{\frac{1}{\alpha-1}}(\alpha-1)/(\alpha-\hat{\beta})$, and $\lambda = w_L + (1-\mu)^{\frac{1}{\alpha}}(w_H - w_L)$. The time when an agent $(\beta, \hat{\beta})$ finishes the low workload w_L is:

$$\tau = \left[1 - (1 - w_L/\lambda)^{\frac{1}{B}}\right] T. \tag{A.1}$$

The unique work schedule for the agent is:

$$x_t(\mathcal{T}, \mathcal{B}) = \begin{cases} \lambda \left[1 - \left(1 - \frac{t}{T}\right)^B\right] & \text{if } 0 \le t < \tau, \\ w_L & \text{if } \tau \le t < T \text{ and } w = w_L, \\ w_L + (w_H - w_L)\left[1 - \left(1 - \frac{t-\tau}{T-\tau}\right)^B\right] & \text{if } \tau \le t < T \text{ and } w = w_H; \end{cases}$$

$$y_t(\mathcal{T}, \mathcal{B}) = \begin{cases} \frac{\lambda B}{T}\left(1 - \frac{t}{T}\right)^{B-1} & \text{if } 0 \le t < \tau, \\ 0 & \text{if } \tau \le t < T \text{ and } w = w_L, \\ \frac{(w_H - w_L)B}{T-\tau}\left(1 - \frac{t-\tau}{T-\tau}\right)^{B-1} & \text{if } \tau \le t < T \text{ and } w = w_H. \end{cases}$$

The cost function (or the ex-ante perceived cost) is

$$C(\mathcal{T}, \mathcal{B}) = \frac{\gamma B^{\alpha-1}\lambda^\alpha}{T^{\alpha-1}}. \tag{A.2}$$

The long-run cost associated with the work schedule is

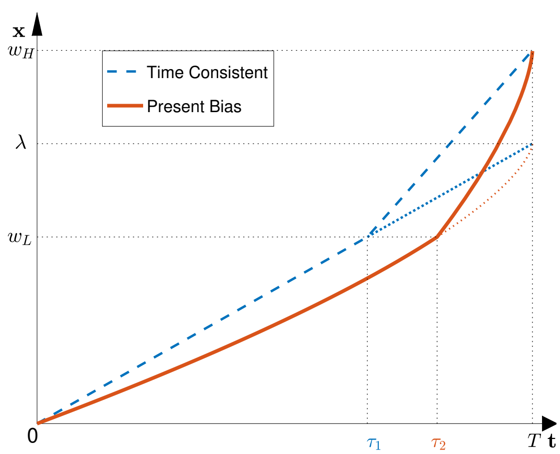
$$LC(\mathcal{T}, \mathcal{B}) = \frac{\gamma B^\alpha \lambda^\alpha}{[1 - \alpha(1-B)]T^{\alpha-1}} \tag{A.3}$$

if $\alpha(1-B) < 1$ and $LC(\mathcal{T}, \mathcal{B}) = \infty$ otherwise.

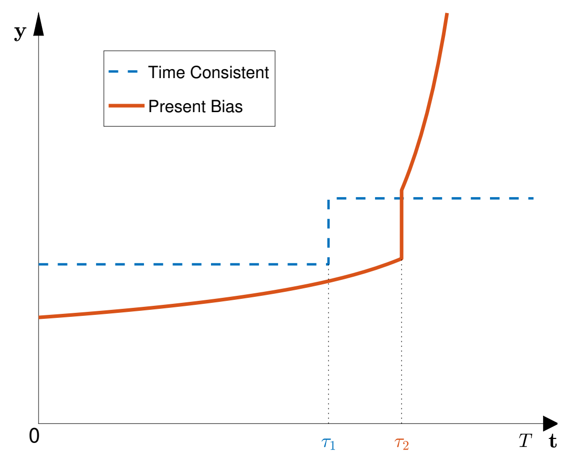
Here, λ can be interpreted as the agent's initial workload target under workload uncertainty. *Proposition 1* implies that the first-best benchmark is attained by dynamic choices if and only if the agent is time consistent ($B = 1$) and there is no workload uncertainty ($\lambda = w$). If the workload is uncertain ($\mu \in (0,1)$), task completion involves two phases: completing w_L by time $\tau < T$, followed by a divergence in the work schedule based on whether the reward is granted at time τ.

Recall that, once the agent completes w_L, the agent knows the task workload — based on whether or not she obtains the task completion reward. It follows that work schedule after τ can be characterized by *Proposition 1* in the main body of the paper. Therefore, to examine the impact of workload uncertainty, we only need to study: (i) how the agent should exert effort over time before she completes w_L, and (ii) how τ is determined.

Proposition A.1 suggests that, before she finishes w_L, the agent acts as if her workload is $\lambda = w_L + (1-\mu)^{\frac{1}{\alpha}}(w_H - w_L) \in (\mu w_L + (1-\mu)w_H, w_H)$.The time to learn the actual task workload, τ, is determined accordingly as the accumulated effort reaches w_L, which is given by (A.1). As depicted in *Figure A.1*, the agent's learning-by-doing process features an endogenous "moving-the-goalpost" pattern. Suppose the actual workload is high and the agent is uncertain about that. The agent would initially aim at an intermediate target λ until she completes the workload w_L. Without receiving the reward for task completion, she then raises her target to the actual workload w_H.[1] Uncertainty affects the agent's work schedule through this misspecified target, $\lambda \notin \{w_L, w_H\}$. It depends on task features (w_H, w_L, μ), and is independent of the behavioral frictions of present bias and naivete.



(a) Work Trajectory



(b) Effort Trajectory

FIGURE A.1. Work Schedule: Sophisticated Agent & Uncertain Workload

[1]This dynamic pattern manifests itself more clearly when the workload uncertainty is resolved gradually; check the online appendix available from the author's webpage for an illustration. Moreover, as shown clearly in *Figure A.1(b)*, an upsurge in effort occurs at time τ. The effort jump strictly increases in the prior μ, capturing the surprise at a high workload.

To examine the impact of workload uncertainty, I take a mean-preserving spread of the workload prior and let it represent an increased workload uncertainty. Formally, denote the prior mean of workload as $\bar{w} = \mu w_L + (1-\mu) w_H$, and let

$$w_H = (1+\varepsilon)\bar{w}, \tag{A.4}$$

$$w_L = [1 - (1-\mu)\varepsilon/\mu]\,\bar{w}. \tag{A.5}$$

Fixing $\mu \in (0,1)$, I use $\varepsilon \in [0, \mu/(1-\mu)]$ to measure the magnitude of the workload uncertainty. If $\varepsilon = 0$, then $w_H = w_L = \bar{w}$ and there is no workload uncertainty in the prior. As ε rises, the dispersion of the workload prior expands, capturing the idea that the agent is less certain about the task difficulty. Plugging (A.4) and (A.5) in the expression of λ, we can obtain that the initial workload target $\lambda = \left\{1 + \left[(1-\mu)^{\frac{1}{\alpha}} - (1-\mu)\right]\varepsilon/\mu\right\}\bar{w} \geq \bar{w}$. The equality holds if and only if there is no workload uncertainty (i.e., $\varepsilon = 0$), and the initial target is boosted up as the workload uncertainty ε increases. This implies that an increased workload uncertainty alleviates procrastination and accelerates the learning of the workload.

So how does the workload uncertainty affect the agent's welfare? As a benchmark, I first calculate the expected cost when uncertainty is resolved *before* the agent chooses her work schedule:

$$\begin{aligned}\Pi(\mathcal{T},\mathcal{B}) &= \mu \cdot LC(w_L, T, \mathcal{B}) + (1-\mu)\cdot LC(w_H, T, \mathcal{B}) \\ &= \frac{\gamma B^{\alpha}[\mu w_L^{\alpha} + (1-\mu) w_H^{\alpha}]}{[1-\alpha(1-B)]T^{\alpha-1}}.\end{aligned}$$

The gap between $LC(\mathcal{T},\mathcal{B})$ and $\Pi(\mathcal{T},\mathcal{B})$, denoted by $I(\mathcal{T},\mathcal{B})$, captures the value of learning workload before starting to work for a present-biased agent:

$$\begin{aligned}I(\mathcal{T},\mathcal{B}) &\equiv LC(\mathcal{T},\mathcal{B}) - \Pi(\mathcal{T},\mathcal{B}) \\ &= \frac{\gamma B^{\alpha}}{[1-\alpha(1-B)]T^{\alpha-1}}\left[\lambda^{\alpha} - \mu w_L^{\alpha} - (1-\mu) w_H^{\alpha}\right] \geq 0,\end{aligned} \tag{A.6}$$

where the equality holds if and only if there is no workload uncertainty (see *Figure A.2* for an illustration).

Here, the interactive impact of behavioral frictions and task features rearises: for a given task $\mathcal{T}$, the welfare loss inflicted by workload uncertainty (i.e., the value of information) grows with the agent's present bias and naivete. Intuitively, a present-biased agent, without commitment, cannot overcome the innate inclination to put off what she can do today until tomorrow, and thus fail to finish w_L as early and ideally as a time-consistent agent. (A.1) formally confirms this intuition since τ strictly decreases in B. Therefore, under workload uncertainty, apart from the effort misdistribution that affects agents of

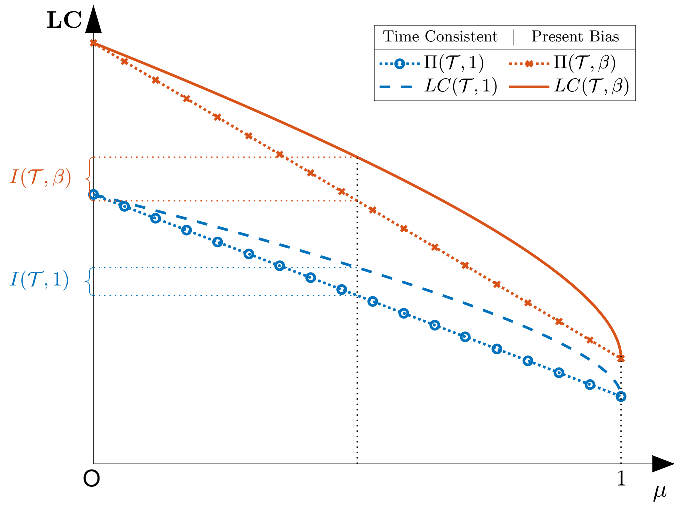


FIGURE A.2. The Welfare Effect of Uncertainty

all time preferences, behavioral frictions additionally induce a delay in discovering the actual workload.

COROLLARY A.1. *Consider an agent who is uncertain about the workload.*

(i) An increased workload uncertainty leads to a higher initial workload target, thus promoting more early effort and alleviating procrastination.

(ii) Workload uncertainty undermines the agent's welfare, and behavioral friction amplifies the welfare loss from workload uncertainty (i.e., the value of information).

A.1.2. *Joint Identification of Belief and Time Preference.* If the workload is high, then apart from the behavioral frictions of present bias and naivete, there is alternative force that contributes to procrastination; namely, a belief that the workload can be low. This section is devoted to separating time preference from workload belief in explaining an agent's procrastination.

By the closed-form work trajectory in *Proposition A.1*, time preference parameters $(\beta, \hat{\beta})$ are entirely summarized in B, whereas prior belief factors, namely the possible binary workloads (w_H and w_L) and the likelihood (μ) are entirely summarized in λ. More importantly, B and λ affect the work trajectory in a distinguishable way: time preferences control how much the work trajectory is tilted towards the deadline, whereas the prior belief on workload determines the initial target under the workload uncertainty. We can thus exploit this variation to jointly identify time preference and belief on workload.

Formally, fix any vector of (B, λ). Assume an agent's work trajectory $\{x_t : t \in [0, T]\}$ is observed. If the work trajectory is smooth and has no kink, then one can conclude immediately that the dynamic choices are made under workload certainty ($w_L = x_T, \mu = 1$), and B can be identified as in the *Section 3.2*.

If, otherwise, the work trajectory has one kink at τ, we can first recover B by

$$B = \varphi^{-1}(x_s/x_t),$$

where $0 < s < t \leq \tau$. Then recover λ by

$$\lambda = \frac{x_t}{1 - (1 - t/T)^B}.$$

COROLLARY A.2 (Identification)**.** *Suppose a work trajectory $\boldsymbol{x} = \{x_t : t \in [0, T]\}$ is observed. A measure of time preference $B = \left(\beta/\hat{\beta}\right)^{\frac{1}{\alpha-1}} (\alpha - 1)/(\alpha - \hat{\beta})$ and a measure of prior belief $\lambda = w_L + (1-\mu)^{\frac{1}{\alpha}}(w_H - w_L)$ can be jointly identified.*

A.1.3. *The Value of Short-Term Goals under Uncertain Workload.* Since $\tau(\mathcal{B}) > \tau(\hat{\beta}, \hat{\beta}) > \tau(1, 1)$ for $\beta < \hat{\beta} \leq 1$, a present-biased agent recovers the workload later than a time-consistent agent; and a naive agent recovers the workload later than she initially anticipates. Now I allow the agent to commit to short-term goals, and in particular, the time to finish w_L and recover the workload. We then ask: what is the optimal committed time, $\hat{\tau}^*$? Furthermore, will this additional commitment power improve individual welfare?

Let $\hat{\tau} \in (0, T)$ be the committed time to finish w_L. We first divide all the other short-term goals except $(w_L, \hat{\tau})$ into two phases. Phase I includes short-term goals before the workload is learnt (i.e., $\hat{\tau}_i < \hat{\tau}$), and Phase II includes those after (i.e., $\hat{\tau}_i > \hat{\tau}$). short-term goals in Phase II are executed only when the agent does not get the reward at time $\hat{\tau}$. Still, I focus on the effective goals that cannot be implemented without intermediate deadlines. Since within Phase I, the agent chooses her optimal short-term goals as if the task requirement is certain as $(w_L, \hat{\tau})$, the short-term goals in Phase I yield no value to the agent according to *Section 4*, and the minimal perceived cost is $C(w_L, \hat{\tau}, \mathcal{B})$ regardless of the number of short-term goals in Phase I. Likewise, the short-term goals in Phase II yield no value to the agent. With probability μ, no work needs to be done in Phase II and thus the overall cost in this phase is 0; with probability $1 - \mu$, the agent chooses her optimal short-term goals taking the task requirement as $(w_H - w_L, T - \hat{\tau})$, and the minimal perceived cost is $C(w_H - w_L, T - \hat{\tau}, \mathcal{B})$. Therefore, the only hope for the agent to reduce her overall cost by an short-term goal is by setting the optimal $\hat{\tau}$.

Now, I characterize the optimal committed time to finish w_L, $\hat{\tau}^*$. Since

$$\hat{C}^* = \min_{\hat{\tau} \in [0,T]} C(w_L, \hat{\tau}) + (1-\mu)C(w_H - w_L, T - \hat{\tau}),$$

where $C(\cdot)$ is given by *Proposition 1*, the first-order condition with regard to $\hat{\tau}$ yields

$$-(\alpha-1)\frac{w_L^\alpha}{\hat{\tau}^{*\alpha}}+(1-\mu)(\alpha-1)\frac{(w_H-w_L)^\alpha}{(T-\hat{\tau}^*)^\alpha}=0.$$

Thus, $\hat{\tau}^* = T/\left[(1-\mu)^{\frac{1}{\alpha}}(w_H/w_L-1)+1\right] = w_L T/\lambda$. Similarly, the long-run cost $LC(\cdot)$ features the same the optimal committed time to finish w_L, $\hat{\tau}^*$. Observe that $\hat{\tau}^*$ is totally determined by task features (w_L, w_H, μ, T) and is invariant with the agent's time preference $(\beta, \hat{\beta})$. Therefore, as in the certain workload case, the optimal short-term goals are aligned with the time-consistent agent's work trajectory.

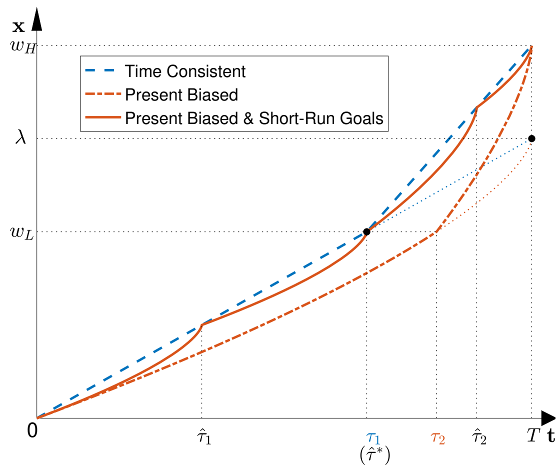


FIGURE A.3. Work Trajectory under Uncertain Workload with Optimal short-term Goals

As illustrated in *Figure A.3*, the agent procrastinates less with the optimal short-term goals. The welfare analysis is analogous to the case when the workload is certain. Again, short-term goals cannot reduce either the perceived cost or the long-run cost, since

$$\hat{C}^* = \gamma[w_L + (1-\mu)^\alpha(w_H-w_L)](B\lambda)^{\alpha-1}/T^\alpha = C(\mathcal{T},\mathcal{B}),$$
$$\hat{LC}^* = B\cdot\hat{C}^*/(\alpha B+1-\alpha) = B\cdot C(\mathcal{T},\mathcal{B})/(\alpha B+1-\alpha) = LC(\mathcal{T},\mathcal{B}).$$

A.1.4. *General Prior Belief on Workload.* In previous sections, I analyze the binary case when the agent is uncertain about workload. Now let's think about what happens if the workload follows any discrete distribution. Assume the workload draws from k positive values, i.e., $w \in \{w_1, w_2, \cdots, w_k\}$ with $0 < w_1 < w_2 < \cdots < w_k$. The agent's prior is that $w = w_i$ with probability $\mu_i \in [0,1]$, and $\sum_{i=1}^k \mu_i = 1$. Compared to the binary case, now learning is done gradually instead of once and for all: the belief on workload is updated every time when the agent completes $\Delta w_i = w_i - w_{i-1}$ (with the convention that $w_0 = 0$).

Let λ_i be the workload target in phase $i \in \{1, 2, \cdots, k\}$ when the accumulated effort exceeds w_{i-1} but is less than w_i, and the task has not been completed yet. Denote $\xi_i \equiv$

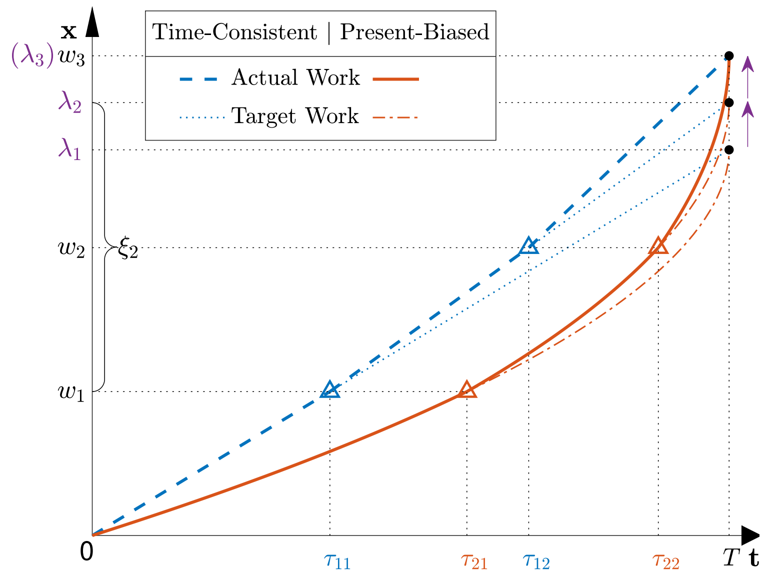


FIGURE A.4. Moving the Goalposts

$\lambda_i - w_{i-1}$ as the additional work to do in phase i. By Bayesian updating, the posterior that workload is w_j for any $1 \leq i \leq j \leq k$ is

$$\mu_j' = \frac{\mu_j}{1 - \sum_{m=1}^{i-1} \mu_m}. \tag{A.7}$$

Let τ_i be the time that a sophisticated agent with present bias β finishes w_i when $w \geq w_i$, and denote $\Delta\tau_i = \tau_i - \tau_{i-1}$, with the convention that $\tau_0 = 0$ and $\tau_k = T$. By backward induction, for any phase $i \in \{1, 2, \cdots, k-1\}$, we have $\xi_i = \Delta w_i + (1 - \mu_i')^{\frac{1}{\alpha}} \xi_{i+1}$ where μ_i' is given by (A.7) and $\xi_k = \Delta w_k$. In particular,

(i) In the last phase k, the workload is certain to be w_k, so $\mu_k' = 1$ and $\xi_k = \Delta w_k$.
(ii) In the second to last phase $k-1$, we are back to the binary case, so $\mu_{k-1}' = \frac{\mu_{k-1}}{\mu_{k-1} + \mu_k}$ and $\xi_{k-1} = \Delta w_{k-1} + (1 - \mu_{k-1}')^{\frac{1}{\alpha}} \xi_k$.
(iii) Continuing in this way, in the initial phase 1, $\mu_1' = \mu_1$ and we have $\xi_1 = w_1 + (1 - \mu_1)^{\frac{1}{\alpha}} \xi_2$ by induction.

Solving the series of $\{\xi_n\}$, we get $\xi_i = [\sum_{m=i}^{k} (\sum_{j=m}^{k} \mu_j)^{\frac{1}{\alpha}} \Delta w_m] / [(\sum_{j=i}^{k} \mu_j)^{\frac{1}{\alpha}}] \geq \Delta w_i$. Therefore, the work trajectory in any phase $i \in \{1, 2, \cdots, k\}$ is

$$\tilde{x}_t = w_{i-1} + \xi_i \left[1 - \left(1 - \frac{t - \tau_{i-1}}{T - \tau_{i-1}}\right)^B\right],$$

and $\tau_i = \tau_{i-1} + \left[1 - \left(1 - \frac{w_i - w_{i-1}}{\xi_i}\right)^{\frac{1}{B}}\right] (T - \tau_{i-1})$.

Note that since $\xi_i = \Delta w_i + \sqrt{1-\mu_i'}\xi_{i+1}$ for any $i = 1, 2, \cdots, k-1$, we obtain

$$\lambda_i = w_{i-1} + \xi_i = w_i + \sqrt{1-\mu_i'}\xi_{i+1} < w_i + \xi_{i+1} = \lambda_{i+1}.$$

This implies the agent "moves the goalposts" when working on a long-term project with uncertain workload, as illustrated in *Figure A.4*. Whenever her accumulated effort reaches some possible workload but is still not enough for the task completion, she aims higher until she completes the task.

Also since $\Delta\tau_i = \left[1 - \left(1 - \frac{w_i}{\lambda_i}\right)^{\frac{1}{B}}\right](T - \tau_{i-1})$ strictly decreases in B for any $i \in \{1, \cdots, k\}$ and $\tau_0 = 0$, τ_i is smaller for an agent with less present bias and naivete. Therefore, as in *Appendix A.1.1*, a more present-biased and naive agent postpones resolving the uncertainty.

A.2. *Exponential Discount Rate $\delta > 0$*

In this section, I look at the case when $\delta > 0$. Now we need to account for the change in continuation cost due to discount. Therefore, the HJB equation (14) becomes

$$c(\mathbf{e}^S(x,t)) - \delta V^S(x,t) + V_t^S(x,t) + V_x^S(x,t)\mathbf{e}^S(x,t) = 0.$$

Following the procedure in*Section 3.3.1*, I first characterize work schedule and individual welfare when $\delta > 0$ in the proposition below.

PROPOSITION A.2. *Let $\delta > 0$. The unique work schedule for a sophisticated agent with the time preference (β, β, δ) and a certain task (w, T) is: for all $t \in [0, T]$,*

$$x_t^S = w - w\left(\frac{e^{\frac{\delta}{\alpha-1}T} - e^{\frac{\delta}{\alpha-1}t}}{e^{\frac{\delta}{\alpha-1}T} - 1}\right)^{\frac{\alpha-1}{\alpha-\beta}},$$

$$y_t^S = \frac{\delta w}{(\alpha-\beta)(e^{\frac{\delta}{\alpha-1}(T-t)} - 1)}\left(\frac{e^{\frac{\delta}{\alpha-1}T} - e^{\frac{\delta}{\alpha-1}t}}{e^{\frac{\delta}{\alpha-1}T} - 1}\right)^{\frac{\alpha-1}{\alpha-\beta}}.$$

The cost (or perceived cost) function is

$$C(w, T, \beta, \delta) = \gamma w^\alpha \left[\frac{\delta}{(\alpha-\beta)(e^{\frac{\delta}{\alpha-1}T} - 1)}\right]^{\alpha-1}.$$

The continuation cost (or long-run cost) function is

$$LC(w, T, \beta, \delta) = \frac{\gamma w^\alpha}{\beta}\left[\frac{\delta}{(\alpha-\beta)(e^{\frac{\delta}{\alpha-1}T} - 1)}\right]^{\alpha-1}.$$

Note that *Section 3.3.1* is a limiting case of *Proposition A.2* when $\delta > 0$ approaches to 0.

Next, I characterize the optimal short-term goals when $\delta > 0$. Again, I focus on finite effective short-term goals:

$$\boldsymbol{G}^k = \{(\hat{w}_i, \hat{\tau}_i) : 1 \leq i \leq k, \hat{w}_k = w, \hat{\tau}_k = T, \hat{w}_i > \hat{w}_{i-1} + x^S_{\hat{\tau}_i - \hat{\tau}_{i-1}}(w - \hat{w}_{i-1}, T - \hat{\tau}_{i-1}, \beta)\}.$$

To set the optimal short-term goals, the agent/her advisor minimizes

$$\sum_{i=1}^{k} \frac{(\hat{w}_i - \hat{w}_{i-1})^\alpha}{\left[e^{\frac{\delta(\hat{\tau}_i - \hat{\tau}_{i-1})}{\alpha-1}} - 1\right]^{\alpha-1}}.$$

Fix $\{\hat{\tau}_i : 1 \leq i \leq k\}$ and optimize the short-term committed workload $\{\hat{w}_i : 1 \leq i \leq k\}$. By F.O.C. with regard to $\hat{w}_i$ we have for all $1 \leq i \leq k-1$,

$$\frac{\alpha(\hat{w}_i - \hat{w}_{i-1})^{\alpha-1}}{\left[e^{\frac{\delta(\hat{\tau}_i - \hat{\tau}_{i-1})}{\alpha-1}} - 1\right]^{\alpha-1}} = \frac{\alpha(\hat{w}_{i+1} - \hat{w}_i)^{\alpha-1}}{\left[e^{\frac{\delta(\hat{\tau}_{i+1} - \hat{\tau}_i)}{\alpha-1}} - 1\right]^{\alpha-1}} \equiv \alpha\eta^{\alpha-1}.$$

Therefore $\Delta\hat{w}_i = (e^{\frac{\delta\Delta\hat{\tau}_i}{\alpha-1}} - 1)\eta$. Then I minimize $\sum_{i=1}^k \frac{(\Delta\hat{w}_i)^\alpha}{\left[e^{\frac{\delta\Delta\hat{\tau}_i}{\alpha-1}} - 1\right]^{\alpha-1}} = \eta^\alpha \sum_{i=1}^k \left[e^{\frac{\delta\Delta\hat{\tau}_i}{\alpha-1}} - 1\right]$ with regard to the intermediate deadlines $\{\hat{\tau}_i : 1 \leq i \leq k\}$. By F.O.C.,

$$\frac{\delta}{\alpha-1}e^{\frac{\delta\Delta\hat{\tau}_i}{\alpha-1}} - \frac{\delta}{\alpha-1}e^{\frac{\delta\Delta\hat{\tau}_{i+1}}{\alpha-1}} = 0 \Rightarrow \ \Delta\hat{\tau}_i = \Delta\hat{\tau}_{i+1}.$$

Therefore, we have $\Delta\hat{\tau}_i = \frac{T}{k}$, and $\Delta\hat{w}_i = \frac{w}{k}$ for any $1 \leq i \leq k$. In other words, the optimal short-term goals are equally spaced and spread the total workload over time evenly.

I then examine the welfare impact of short-term goals. The minimal perceived cost and long-run cost under optimal short-term goals are given by

$$\hat{C}^* = \gamma w^\alpha \left[\frac{\delta}{k(\alpha-\beta)(e^{\frac{\delta T}{k(\alpha-1)}} - 1)}\right]^{\alpha-1} > \gamma w^\alpha \left[\frac{\delta}{(\alpha-\beta)(e^{\frac{\delta T}{\alpha-1}} - 1)}\right]^{\alpha-1} = C(w, T, \beta, \delta),$$

$$\hat{LC}^* = \frac{\hat{C}^*}{\beta} > \frac{C(w, T, \beta, \delta)}{\beta} = LC(w, T, \beta, \delta).$$

In sum, for any agent (including a time-consistent agent), setting optimal short-term goals makes them procrastinate less. Nevertheless, the cost of being myopic to urgent goals (*Tunnel-Vision Effect*) always exceeds the benefit from an even work schedule (*Keep-on-Track Effect*). Therefore, the net value of short-term goals with $\delta > 0$ is always negative.

APPENDIX B. OMITTED PROOFS FOR AUXILIARY RESULTS

B.1. *Proof of Inequality (19)*

The following lemma is useful in establishing *Inequality (19)* in the main body of the paper of the paper.

LEMMA B.1. *For any $A > 0$, $B > 0$, $\alpha > 1$, and $k \in \mathbb{N}_+$,*

$$\frac{x_1^\alpha}{y_1^{\alpha-1}} + \frac{x_2^\alpha}{y_2^{\alpha-1}} + \cdots + \frac{x_k^\alpha}{y_k^{\alpha-1}} \geq \frac{A^\alpha}{B^{\alpha-1}},$$

where $x_i > 0$ and $y_i > 0$ for $i = 1, \cdots, k$, $\sum_{i=1}^k x_i = A$, and $\sum_{i=1}^k y_i = B$.

PROOF. Clearly *Lemma B.1* holds for $k = 1$. For $k \geq 2$, fix any $y_1, y_2, \cdots, y_k > 0$ such that $\sum_{i=1}^k y_i = B$, and define $f(x_1, x_2, \cdots, x_k) \equiv \frac{x_1^\alpha}{y_1^{\alpha-1}} + \frac{x_2^\alpha}{y_2^{\alpha-1}} + \cdots + \frac{x_k^\alpha}{y_k^{\alpha-1}}$. I will show that $\min_{x_1, x_2, \cdots, x_k} f(x_1, x_2, \cdots, x_k) = \frac{A^\alpha}{B^{\alpha-1}}$ such that $\sum_{i=1}^k x_i = A$. Take the Lagrangian function of the constrained optimization

$$L(x_1, x_2, \cdots, x_k; \lambda) = \frac{x_1^\alpha}{y_1^{\alpha-1}} + \frac{x_2^\alpha}{y_2^{\alpha-1}} + \cdots + \frac{x_k^\alpha}{y_k^{\alpha-1}} + \lambda(A - \sum_{i=1}^k x_i),$$

where $\lambda > 0$ is the Lagrangian multiplier. By the first-order condition with regard to any x_i, we have

$$\alpha \left(\frac{x_i^*}{y_i}\right)^{\alpha-1} - \lambda = 0.$$

Hence, $\frac{x_1^*}{y_1} = \frac{x_2^*}{y_2} = \cdots = \frac{x_k^*}{y_k} = \frac{\sum_{i=1}^k x_i^*}{\sum_{i=1}^k y_i} = \frac{A}{B}$. Note also that $L(x_1, x_2, \cdots, x_k; \lambda)$ is a strictly convex function in $(x_1, \cdots, x_k)$, which implies that the global minimum of $L(\cdot)$ is attained when the first-order conditions hold. Therefore, the unique constrained minimizer is $(x_1^*, x_2^*, \cdots, x_k^*)$, and thus $LHS \geq (A/B)^\alpha (y_1 + y_2 + \cdots + y_k) = RHS$. ■

Take $x_i = (1 - \tau_{i-1}/T)^B - (1 - \tau_i/T)^B$, $y_i = \tau_i - \tau_{i-1}$, $A = 1$ and $B = T$ in *Lemma B.1*, and then we can obtain (19).

B.2. *Proof of Proposition A.1*

Since the agent gets to know her task workload as soon as she finishes w_L, the agent's objective function of the dynamic optimization before she finishes w_L is formulated as follows:

$$\min_{\tau \in [0,T]} \left\{ \min_{\{y: \int_0^\tau y_s \mathrm{d}s = w_L\}} \int_0^\tau D_0(t)c(y_t)\mathrm{d}t + (1-\mu)\beta V^S(w_L, \tau; w_H, T, \beta) \right\},$$

where $V^S(\cdot)$ is the continuation cost given by *(15)* in the main body of the paper. This objective function is the sum of two terms. The first term is the perceived effort cost to finish w_L. The second term is the expected cost when w_L is finished: with probability μ, the agent completes the task and incurs no more cost; with probability $1-\mu$, she still has $w_H - w_L$ to do. Following the four-step procedure in *Section 3.3.1*, I then obtain the work schedule and effort costs for a sophisticated agent under workload uncertainty as follows.

LEMMA B.2. *Let $\lambda = w_L + (1-\mu)^{\frac{1}{\alpha}}(w_H - w_L), \tau = \left[1-(1-w_L/\lambda)^{\frac{\alpha-\beta}{\alpha-1}}\right] T$. For $t \in [0,\tau)$, the unique work schedule for a sophisticated agent with present bias $\beta \in (0,1]$ and prior $\mu \in [0,1]$ is*

$$x_t(\mathcal{T},\beta) = \lambda[1-(1-t/T)^{\frac{\alpha-1}{\alpha-\beta}}],$$
$$y_t(\mathcal{T},\beta) = [(\alpha-1)\lambda/(\alpha-\beta)T]\,(1-t/T)^{-\frac{1-\beta}{\alpha-\beta}}.$$

For $t \in [\tau, T)$, if $w = w_H$, then the unique work schedule is given by Proposition 1 with the workload $w_H - w_L$, the time available $T-\tau$, and $\hat{\beta} = \beta$; otherwise, the agent stops working.

The cost function (or ex-ante perceived cost) is

$$C(\mathcal{T},\beta) = \gamma\lambda^{\alpha}/T^{\alpha-1}[(\alpha-1)/(\alpha-\beta)]^{\alpha-1}. \tag{B.1}$$

The continuation cost (or long-run cost) is $LC(\mathcal{T},\beta) = C(\mathcal{T},\beta)/\beta$

PROOF OF LEMMA B.2. The dynamic programming problem is formulated as follows. Fix the initial state $(x,t) \in [0,w_L) \times [0,T]$. Let $\langle \tilde{\boldsymbol{x}}^S(x,t), \tilde{\boldsymbol{y}}^S(x,t)\rangle$, $\tilde{W}^S(x,t)$ and $\tilde{V}^S(x,t)$ be the work schedule, cost function, and continuation cost function, respectively, for a sophisticated agent starting from the state (x,t), and let $\tau(x,t)$ be the corresponding optimal time to finish w_L.

We can follow the procedure described in *Section 3.3.1* to solve this dynamic programming. The only twists here consist in the long-run cost function and boundary conditions:

$$\tilde{V}^S(x,t) = \int_t^{\tau(x,t)} c(\tilde{y}^S_s(x,t))\mathrm{d}s + (1-\mu)V^S(w_L,\tau(x,t);w_H,T,\beta),$$
$$\tilde{x}^S_{\tau(x,t)}(x,t) = w_L,$$
$$\lim_{x'\uparrow w_L,\, t'\uparrow\tau(x,t)} \tilde{V}^S(x',t') = (1-\mu)V^S(w_L,\tau(x,t);w_H,T,\beta), \tag{B.2}$$

for any $(x,t) \in [0,w_L) \times [0,T]$, where $V^S(\cdot)$ is the continuation cost given by equation (15) in the main text.

Suppose $\tilde{V}^S(x,t) = f(t)g(x)$ where $f(\cdot), g(\cdot)$ are two continuously differentiable functions, and $f'(t) > 0, g'(x) < 0$. Again, the optimal effort equation (11) gives us

$$\tilde{\mathbf{e}}^S(x,t) = \left[-\frac{\beta}{\alpha\gamma}f(t)g'(x)\right]^{\frac{1}{\alpha-1}}.$$

Plugging in the HJB equation (12) and rearranging terms, we have

$$\frac{f'(t)}{[f(t)]^{\frac{\alpha}{\alpha-1}}} = \left(\frac{1}{\gamma}\right)^{\frac{1}{\alpha-1}} \left[\left(\frac{\beta}{\alpha}\right)^{\frac{1}{\alpha-1}} - \left(\frac{\beta}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}\right] \frac{[-g'(x)]^{\frac{\alpha}{\alpha-1}}}{g(x)}.$$

Following the proof steps for *Proposition 1* and solving the first-order ordinary differential equations, we obtain:

$$\begin{aligned}
f(t) &= \left(\frac{\alpha-1}{-Ht+J}\right)^{\alpha-1}, \\
g(x) &= (Ax+B)^{\alpha}, \\
\tilde{V}^S(x,t) &= f(t)g(x) = \left(\frac{\alpha-1}{-Ht+J}\right)^{\alpha-1}(Ax+B)^{\alpha}, \qquad \text{(B.3)} \\
\tilde{\mathbf{e}}^S(x,t) &= \left(-\frac{\beta}{\alpha\gamma}f(t)g'(x)\right)^{\frac{1}{\alpha-1}} = \frac{\alpha-1}{-Ht+J}(Ax+B)\left(-\frac{\beta A}{\gamma}\right)^{\frac{1}{\alpha-1}},
\end{aligned}$$

where $A = -\gamma^{\frac{1}{\alpha}}/\beta\left[H\beta/(\alpha-\beta)\right]^{\frac{\alpha-1}{\alpha}}$, and $B, H, J \in \mathbb{R}$ are three constants.

Then using the boundary condition (B.2), we have for any $(x,t) \in [0, w_L) \times [0, T]$,

$$\tilde{V}^S(w_L, \tau(x,t)) = \lim_{x'\uparrow w_L, t'\uparrow\tau(x,t)} \tilde{V}^S(x',t') = (1-\mu)V^S(w_L, \tau(x,t); w_H, T, \beta). \qquad \text{(B.4)}$$

Combining (B.3) and (B.4), we have,

$$\begin{aligned}
\left[\frac{\alpha-1}{-H\tau(x,t)+J}\right]^{\alpha-1}(Aw_L+B)^{\alpha} &= \frac{\gamma}{\beta}\left[\frac{\alpha-1}{(\alpha-\beta)(\frac{J}{H}-\tau(x,t))}\right]^{\alpha-1}\left(-\frac{B}{A}-w_L\right)^{\alpha} \\
&= (1-\mu)\frac{\gamma}{\beta}\left[\frac{\alpha-1}{(\alpha-\beta)(T-\tau(x,t))}\right]^{\alpha-1}(w_H-w_L)^{\alpha}.
\end{aligned}$$

Thus, $J = HT$ and $B = -\lambda A$, where $\lambda = w_L + (1-\mu)^{\frac{1}{\alpha}}(w_H - w_L)$.

Since

$$\dot{\tilde{x}}^S_s(x,t) = \tilde{y}^S_s(x,t) = \tilde{\mathbf{e}}^S(s, \tilde{x}^S_s(x,t)) = \frac{\alpha-1}{\alpha-\beta}\frac{\lambda-\tilde{x}_s(x,t)}{T-s},$$

for any $s \in [t, T]$, we have

$$\tilde{x}^S_s(x,t) = E\left[H(T-s)/A\right]^{\frac{\alpha-1}{\alpha-\beta}} + \lambda,$$

where E is a function of the initial state variables (x,t) and is invariant with s. Then the boundary conditions $\tilde{x}^S_\tau(x,t) = w_L$ and $\tilde{x}^S_t(x,t) = x$ imply that

$$E\left[H(T-\tau(x,t))/A\right]^{\frac{\alpha-1}{\alpha-\beta}} = w_L - \lambda,$$

$$E\left[H(T-t)/A\right]^{\frac{\alpha-1}{\alpha-\beta}} = x - \lambda.$$

Hence, we obtain

$$\left[\frac{T-\tau(x,t)}{T-t}\right]^{\frac{\alpha-1}{\alpha-\beta}} = \frac{\lambda - w_L}{\lambda - x} \Rightarrow \tau(x,t) = (T-t) - (T-t)\left(\frac{\lambda - w_L}{\lambda - x}\right)^{\frac{\alpha-\beta}{\alpha-1}},$$

$$E\,(H/A)^{\frac{\alpha-1}{\alpha-\beta}} = (x-\lambda)(T-t)^{\frac{\alpha-\beta}{\alpha-1}} \Rightarrow \tilde{x}_s(x,t) = (x-\lambda)\left[(T-s)/(T-t)\right]^{\frac{\alpha-1}{\alpha-\beta}} + \lambda.$$

Finally, we can derive the cost function, the continuation cost, the policy function, and the corresponding optimal state and control trajectories in any state $(x,t) \in [0, w_L) \times [0, T]$ as follows:

$$\begin{aligned}
\tilde{W}^S(x,t) &= \beta \tilde{V}^S(x,t), \\
\tilde{V}^S(x,t) &= \frac{\gamma}{\beta}\left(\frac{\alpha-1}{\alpha-\beta}\right)^{\alpha-1} \frac{(\lambda-x)^\alpha}{(T-t)^{\alpha-1}}, \\
\tilde{x}^S_s(x,t) &= \lambda - (\lambda - x)\left(\frac{T-s}{T-t}\right)^{\frac{\alpha-1}{\alpha-\beta}} \text{ for any } s \in [t,T], \\
\tilde{y}^S_s(x,t) &= \dot{x}_s(x,t) = \frac{\alpha-1}{\alpha-\beta}(\lambda-x)\frac{(T-s)^{\frac{\alpha-1}{\alpha-\beta}-1}}{(T-t)^{\frac{\alpha-1}{\alpha-\beta}}} \text{ for any } s \in [t,T], \\
\tau(x,t) &= (T-t) - (T-t)\left(\frac{\lambda - w_L}{\lambda - x}\right)^{\frac{\alpha-\beta}{\alpha-1}},
\end{aligned}$$

where $\lambda = w_L + (1-\mu)^{\frac{1}{\alpha}}(w_H - w_L)$.

Taking the initial state (x,t) as $(0,0)$, we obtain *Lemma B.2*.

■

Now, we can characterize the dynamic work schedule and welfare for a potentially naive agent under workload uncertainty. Fix the initial state $(x,t) \in [0, w_L) \times [0, T]$. Let $\langle \tilde{\boldsymbol{x}}(x,t), \tilde{\boldsymbol{y}}(x,t)\rangle$, $\tilde{W}(x,t)$ and $\tilde{V}(x,t)$ be the work schedule, cost function, and continuation cost function starting from the state (x,t), respectively.

I first relate the cost function for a naive agent $(\beta, \hat{\beta})$, $\tilde{W}(\cdot)$, to the continuation cost for a sophisticated agent $\hat{\beta}$, $\tilde{V}^S(\cdot)$:

$$\tilde{W}(x,t;\mathcal{T},\mathcal{B}) = \beta\tilde{V}^S(x,t;\mathcal{T},\hat{\beta}) = \gamma\beta\left[(\alpha-1)/(\alpha-\hat{\beta})\right]^{\alpha-1}(\lambda-x)^\alpha / \left[\hat{\beta}(T-t)^{\alpha-1}\right]$$

for $x < w_L$, where $\lambda = w_L + (1-\mu)^{\frac{1}{\alpha}}(w_H - w_L)$. Thus, the cost function at the start is

$$C(\mathcal{T},\mathcal{B}) = \tilde{W}(0,0;\mathcal{T},\mathcal{B}) = \frac{\gamma\beta}{\hat{\beta}}\left(\frac{\alpha-1}{\alpha-\hat{\beta}}\right)^{\alpha-1}\frac{\lambda^\alpha}{T^{\alpha-1}}.$$

We can then obtain the optimal current effort by the F.O.C. as

$$\tilde{\mathbf{e}}(x,t;\mathcal{T},\mathcal{B}) = \left[-\frac{\beta}{\alpha\gamma}\tilde{V}^S_x(x,t;\mathcal{T},\hat{\beta})\right]^{\frac{1}{\alpha-1}} = B\frac{\lambda-x}{T-t},$$

where $B = (\beta/\hat{\beta})^{\frac{1}{\alpha-1}}(\alpha-1)/(\alpha-\hat{\beta})$. Therefore,

$$\dot{\tilde{x}}_s(x,t) = \tilde{\mathbf{e}}(\tilde{x}_s(x,t), s; \mathcal{T}, \mathcal{B}) = B\,(\lambda - \tilde{x}_s(x,t))/\,(T-s).$$

Solving this first-order differential equation with the boundary condition $\tilde{x}_t(x,t) = x$, we obtain the work trajectory starting from the state $(x,t) \in [0, w_L) \times [0, T]$ as follows:

$$\tilde{x}_s(x,t) = \lambda - (\lambda - x)\left(1 - \frac{s-t}{T-t}\right)^B.$$

Finally taking the initial state $(x,t) = (0,0)$, we get:

(i) A potentially naive agent's work and effort trajectories before she finishes w_L are

$$x_t = \tilde{x}_t(0,0) = \lambda - \lambda\,(1 - t/T)^B,$$
$$y_t = \dot{x}_t = B\,(1-t/T)^{B-1}\,\lambda/T.$$

In particular, when the agent knows the workload w, her effort trajectory is

$$y_t(w, T, \mathcal{B}) = B\,(1-t/T)^{B-1}\,w/T. \tag{B.5}$$

(ii) The time when the naive agent finishes the low workload is given by

$$x_\tau = \lambda - \lambda\,(1-\tau/T)^B = w_L$$
$$\Rightarrow \quad \tau = \left[1 - (1 - w_L/\lambda)^{\frac{1}{B}}\right] T.$$

(iii) The expected long-run cost associated with this work schedule,

$$\begin{aligned} LC &= \int_0^\tau c\,(y_t)\,\mathrm{d}t + (1-\mu)\int_0^{T-\tau} c\,(y_t(w_H - w_L, T-\tau, \mathcal{B}))\,\mathrm{d}t \\ &= \frac{\gamma T}{\alpha(B-1)+1}\left(\frac{B\lambda}{T}\right)^\alpha \left[1 - \left(1 - \frac{\tau}{T}\right)^{\alpha(B-1)+1}\right] + \frac{\gamma(1-\mu)(T-\tau)}{\alpha(B-1)+1}\left[\frac{B(w_H - w_L)}{T-\tau}\right]^\alpha \\ &= \frac{\gamma B^\alpha \lambda^\alpha}{(\alpha B + 1 - \alpha)T^{\alpha-1}}, \end{aligned}$$

where $y_t(\cdot)$ is given by (B.5).

B.3. *Proof of Proposition A.2*

Recall that (i) $W^S = \beta V^S$; (ii) $\mathbf{e}^{\mathbf{S}} = (-\frac{\beta}{\alpha\gamma}V_x^S)^{\frac{1}{\alpha-1}}$; (iii) $\gamma y^\alpha - \delta V^S + V_t^S + V_x^S \mathbf{e}^{\mathbf{S}} = 0$. Suppose for all $(x,t) \in [0,w] \times [0,T]$, $V^S(x,t) = f(t)g(x)$ where $f(\cdot), g(\cdot)$ are two continuously differentiable functions with $f'(t) > 0$ and $g'(x) < 0$. Then by (ii), we have

$$\mathbf{e}^{\mathbf{S}}(x,t) = \left[-\frac{\beta}{\alpha\gamma} f(t) g'(x)\right]^{\frac{1}{\alpha-1}}.$$

Combining this with (iii), we have

$$\gamma \mathbf{e}^{\mathbf{S}}(x,t)^{\alpha} - \delta V^S + f'(t)g(x) + f(t)g'(x)\mathbf{e}^{\mathbf{S}}(x,t) = 0$$

$$\Rightarrow \ H \equiv \frac{f'(t) - \delta f(t)}{[f(t)]^{\frac{\alpha}{\alpha-1}}} = \left(\frac{1}{\gamma}\right)^{\frac{1}{\alpha-1}} \left(\frac{\alpha}{\beta} - 1\right) \frac{\left[-\frac{\beta}{\alpha} g'(x)\right]^{\frac{\alpha}{\alpha-1}}}{g(x)},$$

where $H \in \mathbb{R}$ is a constant. I solve the first-order differential equations and obtain

$$f(t) = \left[\frac{\delta/H}{e^{\frac{\delta}{\alpha-1}(J-t)} - 1}\right]^{\alpha-1},$$

$$g(x) = (Ax+B)^{\alpha},$$

where $A = -\gamma^{\frac{1}{\alpha}}/\beta\,[H\beta/(\alpha-\beta)]^{\frac{\alpha-1}{\alpha}}$ and $B, J \in \mathbb{R}$ are constants. Since the boundary condition $V(w,t) = 0$ holds for any $t \in [0,T]$, we have $B = -Aw$. Therefore,

$$V^S(x,t) = f(t)g(x) = \frac{\gamma}{\beta}\left[\frac{\delta/(\alpha-\beta)}{e^{\frac{\delta}{\alpha-1}(J-t)} - 1}\right]^{\alpha-1} (w-x)^{\alpha},$$

$$\mathbf{e}^{\mathbf{S}}(x,t) = \frac{\delta/(\alpha-\beta)}{e^{\frac{\delta}{\alpha-1}(J-t)} - 1}(w-x).$$

Fix the initial state $(x,t) \in [0,w) \times [0,T)$. Since for any $s \in [t,T]$, $\dot{x}_s(x,t) = y_s(x,t) = \mathbf{e}^{\mathbf{S}}(s, x_s(x,t)) = \frac{\delta/(\alpha-\beta)}{e^{\frac{\delta}{\alpha-1}(J-s)}-1}(w - x_s(x,t))$, we have

$$x_s(x,t) = Ee^{\frac{-\delta}{\alpha-\beta}(J-s)}[e^{\frac{\delta}{\alpha-1}(J-s)} - 1]^{\frac{\alpha-1}{\alpha-\beta}} + w,$$

where E is a function of (x,t) and is unrelated to s. Then by $x_T(x,t) = w$ and $x_t(x,t) = x$,

$$x_T(x,t) = Ee^{\frac{\delta}{\alpha-\beta}(T-J)}[e^{\frac{\delta}{\alpha-1}(J-T)} - 1]^{\frac{\alpha-1}{\alpha-\beta}} + w = w,$$

$$x_t(x,t) = Ee^{\frac{-\delta}{\alpha-\beta}(J-t)}[e^{\frac{\delta}{\alpha-1}(J-t)} - 1]^{\frac{\alpha-1}{\alpha-\beta}} + w = x.$$

These two equations imply that $J = T$, $E = (x-w)e^{\frac{\delta}{\alpha-\beta}(T-t)}[e^{\frac{\delta}{\alpha-1}(T-t)} - 1]^{-\frac{\alpha-1}{\alpha-\beta}}$. Therefore, we have

$$W^S(x,t) = \beta V^S(x,t),$$

$$V^S(x,t) = \frac{\gamma}{\beta}\left[\frac{\delta/(\alpha-\beta)}{e^{\frac{\delta}{\alpha-1}(T-t)} - 1}\right]^{\alpha-1} (w-x)^{\alpha},$$

$$x_s(x,t) = w - (w-x)e^{\frac{\delta}{\alpha-\beta}(s-t)} \left[\frac{e^{\frac{\delta}{\alpha-1}(T-s)} - 1}{e^{\frac{\delta}{\alpha-1}(T-t)} - 1}\right]^{\frac{\alpha-1}{\alpha-\beta}},$$

$$y_s(x,t) = \mathbf{e}^{\mathbf{S}}(s, x_s(x,t)) = \frac{\delta/(\alpha-\beta)}{e^{\frac{\delta}{\alpha-1}(T-s)}-1}(w-x)e^{\frac{\delta}{\alpha-\beta}(s-t)}\left[\frac{e^{\frac{\delta}{\alpha-1}(T-s)}-1}{e^{\frac{\delta}{\alpha-1}(T-t)}-1}\right]^{\frac{\alpha-1}{\alpha-\beta}}.$$

Take the initial state $(x,t) = (0,0)$. We can then obtain that

$$C = \gamma w^{\alpha}\left[\frac{\delta/(\alpha-\beta)}{e^{\frac{\delta}{\alpha-1}T}-1}\right]^{\alpha-1},$$

$$LC = \frac{\gamma w^{\alpha}}{\beta}\left[\frac{\delta/(\alpha-\beta)}{e^{\frac{\delta}{\alpha-1}T}-1}\right]^{\alpha-1},$$

$$x_t^S = w - w\left(\frac{e^{\frac{\delta}{\alpha-1}T}-e^{\frac{\delta}{\alpha-1}t}}{e^{\frac{\delta}{\alpha-1}T}-1}\right)^{\frac{\alpha-1}{\alpha-\beta}},$$

$$y_t^S = \frac{\delta w/(\alpha-\beta)}{e^{\frac{\delta}{\alpha-1}(T-t)}-1}\left(\frac{e^{\frac{\delta}{\alpha-1}T}-e^{\frac{\delta}{\alpha-1}t}}{e^{\frac{\delta}{\alpha-1}T}-1}\right)^{\frac{\alpha-1}{\alpha-\beta}}.$$

B.4. *Short-Term Goals Enables New Implementable Work Schedules*

Consider the optimal short-term goals

$$\boldsymbol{G}^k = \{(\hat{w}_i, \hat{\tau}_i) | \hat{w}_i = \frac{w}{T}\hat{\tau}_i, 1 \le i \le k, \hat{w}_k = w, \hat{\tau}_k = T\}.$$

Let $(\boldsymbol{x}^G, \boldsymbol{y}^G)$ be the corresponding work schedule with these goals. $k = 1$ is exactly the case when no short-term goal is available for the agent. In this appendix, I show that, for $k \ge 2$, $(\boldsymbol{x}^G, \boldsymbol{y}^G)$ is not implementable for a present-biased agent if she has no access to any precommited short-term goals. I use this as evidence of the fact that the commitment device can add work schedules implementable to a present-biased agent.

Note that $x_t^G = x_t(\hat{w}_1, \hat{\tau}_1, \mathcal{B})$ for $t \in [0, \hat{\tau}_1)$. Now I want to establish that, without any short-term goals, the agent has a strict incentive to deviate to an alternative work schedule. To see this, I first fix any time $s \in (0, \hat{\tau}_1)$. Suppose the agent follows the proposed work schedule $(\boldsymbol{x}^G, \boldsymbol{y}^G)$ up to time s. Then her perceived overall cost if she continues to follow this schedule is

$$\begin{aligned}
W^G(x_s^G, s; w, T, \mathcal{B}) =& C(\hat{w}_1 - x_s(\hat{w}_1, \hat{\tau}_1, \mathcal{B}), \hat{\tau}_1 - s, \mathcal{B}) \\
&+ \sum_{i=2}^{k} C(\hat{\tau}_i - \hat{\tau}_{i-1}, \hat{w}_i - \hat{w}_{i-1}, \mathcal{B}) \\
=& \frac{\gamma B^{\alpha-1}\hat{w}_1^{\alpha}(1-\frac{s}{\hat{\tau}_1})^{\alpha B}}{(\hat{\tau}_1 - s)^{\alpha-1}} + \frac{\gamma B^{\alpha-1}(w-\hat{w}_1)^{\alpha}}{(T-\hat{\tau}_1)^{\alpha-1}} \\
=& \gamma B^{\alpha-1}\left(\frac{w}{T}\right)^{\alpha}\left[\frac{(\hat{\tau}_1 - s)^{\alpha B - \alpha + 1}}{\hat{\tau}_1^{\alpha B - \alpha}} + T - \hat{\tau}_1\right],
\end{aligned}$$

where the last equation holds since $\hat{w}_1 = w\hat{\tau}_1/T$.

However, if she instead follows the work schedule $\tilde{x}_t = x_{t-s}(w - x_s^G, T-s, \mathcal{B})$ for $t \in (s, T)$, then her perceived overall cost becomes

$$\begin{aligned}\tilde{W}(x_s^G, s; w, T, \mathcal{B}) &= C(w - x_s(\hat{w}_1, \hat{\tau}_1, \mathcal{B}), T-s) = \gamma B^{\alpha-1} \frac{[w - \hat{w}_1 + \hat{w}_1(1 - s/\hat{\tau}_1)^B]^\alpha}{(T-s)^{\alpha-1}} \\ &= \gamma B^{\alpha-1} \left(\frac{w}{T}\right)^\alpha \frac{\left[T - \hat{\tau}_1 + \hat{\tau}_1 \left(1 - s/\hat{\tau}_1\right)^B\right]^\alpha}{(T-s)^{\alpha-1}} \leq W^G(x_s^G, s; w, T, \mathcal{B}),\end{aligned}$$

where the equality holds if and only if $\beta = 1$. Therefore, for any present-biased agent, the alternative work schedule is strictly preferred, and thus the proposed work schedule $(\boldsymbol{x}^G, \boldsymbol{y}^G)$ is not implementable without short-term goals.

B.5. *Procrastination in the Discrete Time*

B.5.1. *Special Case: Time-Consistent Agent ($\beta = 1$).* When $\beta = 1$, for $k = 1, \cdots, T$, $A_k = 1/\left\{1 + [A_{k-1}^{\alpha-1} - (1-\beta)A_{k-1}^{\alpha}]^{-\frac{1}{\alpha-1}}\right\} = 1/(1 + A_{k-1}^{-1})$, and $A_0 = 1$. Now I prove by mathematical induction that $A_k = \frac{1}{k+1}$ for $k = 0, 1, \cdots, T$.

(i) For $k = 0$, $A_0 = 1/(0+1) = 1$.

(ii) Suppose $A_m = 1/(m+1)$ for $m \in \{0, 1, \cdots, T-1\}$. Then we have

$$A_{m+1} = 1/\left(1 + A_m^{-1}\right) = 1/(m+2).$$

Therefore, $A_k = \frac{1}{k+1}$ for $k = 0, 1, \cdots, T$ if $\beta = 1$.

B.5.2. *Comparative Static of A_k (with regard to β).* Given $A_{k-1} \in (0, 1]$, for all $k = 1, \cdots, T$, $A_k = 1/\left\{1 + [A_{k-1}^{\alpha-1} - (1-\beta)A_{k-1}^{\alpha}]^{-\frac{1}{\alpha-1}}\right\} \in (0, 1)$ strictly increases in β. Given $\beta \in (0, 1]$, A_k also strictly increases in A_{k-1}. Therefore, as β grows, A_k grows for $k = 1, \cdots, T$. In other words, with a smaller present bias, the agent finishes a larger share of remaining work at every period.

B.6. *Non-Identifiability of Rush Aversion α*

Rush aversion $\alpha > 1$ affects both the time preference measure $B = (\beta/\hat{\beta})^{\frac{1}{\alpha-1}}(\alpha-1)/(\alpha-\hat{\beta})$ and the prior belief measure $\lambda = w_L + (1-\mu)^{\frac{1}{\alpha}}(w_H - w_L)$. Given the possible workload (w_H, w_L), (B, λ) summarizes what we can learn from a work trajectory.

In this section, I will show that we cannot recover α from (B, λ), that is, α is not identifiable even with complete work trajectory data.

Formally, for any parameter vector $(\beta, \hat{\beta}, \alpha, \mu) \in (0,1]^2 \times (1, \infty) \times [0, 1]$, there always exists an alternative vector $(\beta', \hat{\beta}', \alpha', \mu') \in (0,1]^2 \times (1, \infty) \times [0, 1]$ yielding the same work trajectory whereas $\alpha' \neq \alpha$, that is,

$$B = \left(\beta/\hat{\beta}\right)^{\frac{1}{\alpha-1}} (\alpha-1)/(\alpha-\hat{\beta}) = \left(\beta'/\hat{\beta}'\right)^{\frac{1}{\alpha'-1}} (\alpha'-1)/(\alpha'-\hat{\beta}'), \tag{B.6}$$

$$\lambda = w_L + (1-\mu)^{\frac{1}{\alpha}}(w_H - w_L) = w_L + (1-\mu')^{\frac{1}{\alpha'}}(w_H - w_L), \tag{B.7}$$

$$\alpha' \neq \alpha \text{ and } (\beta', \hat{\beta}', \alpha', \mu') \in (0,1]^2 \times (1,\infty) \times [0,1]. \tag{B.8}$$

Here is a proposed alternative vector $(\bar{\beta}', \bar{\hat{\beta}}', \bar{\alpha}', \bar{\mu}') \in (0,1]^2 \times (1,\infty) \times [0,1]$ that satisfies the non-identifiable conditions (B.6), (B.7) and (B.8). For any $\varepsilon \in (0, \frac{\hat{\beta}}{1-\hat{\beta}})$, take

$$\begin{aligned}
&\bar{\alpha}' = \alpha + (\alpha-1)\varepsilon > \alpha, \\
&\bar{\mu}' = 1 - (1-\mu)^{\frac{\varepsilon}{\alpha}+1} \in (\mu, 1), \\
&\bar{\hat{\beta}}' = \hat{\beta} - (1-\hat{\beta})\varepsilon \in (0, \hat{\beta}], \\
&\bar{\beta}' = \hat{\beta}' \left(\beta/\hat{\beta}\right)^{\varepsilon+1} = [\hat{\beta} - (1-\hat{\beta})\varepsilon] \left(\beta/\hat{\beta}\right)^{\varepsilon+1} \in (0, \beta).
\end{aligned}$$

If we make no assumption on sophistication and observe only one kink in the work trajectory, then the proposed vector $(\bar{\beta}', \bar{\hat{\beta}}', \bar{\alpha}', \bar{\mu}')$ shows that α is not identifiable. Even if we assume that the agent is sophisticated, the same still holds true (as $\bar{\beta}' = \bar{\hat{\beta}}'$ if $\beta = \hat{\beta}$). Additionally, if there is no kink in the observed work trajectory, we can identify $\mu = 0$ if $w_L = x_T$, and $\mu = 1$ if $w_H = x_T$. In this case, $(\bar{\beta}', \bar{\hat{\beta}}', \bar{\alpha}', \mu)$ shows that α is not identifiable.

To sum up, present bias and prior workload belief can always be adjusted and generate the same effect on work trajectory as rush aversion. With present bias and prior workload belief unobserved, rush aversion cannot be identified.