Online Appendix

DEBRAJ RAY, NIKHIL VELLODI, RUQU WANG

1. DISCUSSION ON THE PROFILE OF PATIENCE

In this section, we elaborate on the specification of weights placed by the agent on her various selves. As mentioned, our results go through relatively unscathed with even more general schemes, and in fact some interesting insights can be drawn from such considerations.

Recall that our date t agent places weight α on her current self, β on her retirement self at date T, and uniform weight ω on all other future selves. Consider the following variations.

1.1. **Weights Only On Future Selves Between Current Age and Retirement.**

Suppose our agent spreads uniform weight over all her "intermediate" future selves, i.e. weight ω is placed on all selves in (t, T), in addition to the weights on her current and retirement selves, which remain α and β respectively. Make exactly the same assumption A.1 in the main text. Then very little changes in Proposition 1. Intuitively, a weight on future selves beyond T, as in the main text, only serves to enhance the importance of the retirement self, thereby raising the effective value of β but otherwise causing no difference at all to the pre-retirement results. This intuition is fully borne out. The instantaneous rate of impatience $i(t,s)$ for $s \in [t, T)$ now reads

$$i(t,s) = \Big[\frac{(\rho_f \alpha - \omega)e^{-\rho_f(s-t)} - (\rho_b \beta - \omega)e^{-\rho_b(T-s)}}{(\alpha - \frac{\omega}{\rho_f})e^{-\rho_f(s-t)} + (\beta - \frac{\omega}{\rho_b})e^{-\rho_b(T-s)} + \frac{\omega}{\rho_f} + \frac{\omega}{\rho_b}}\Big]$$

and in particular,

$$i(t,t) = \Big[\frac{(\rho_f \alpha - \omega) - (\rho_b \beta - \omega)e^{-\rho_b(T-t)}}{(\alpha - \frac{\omega}{\rho_f}) + (\beta - \frac{\omega}{\rho_b})e^{-\rho_b(T-t)} + \frac{\omega}{\rho_f} + \frac{\omega}{\rho_b}}\Big]$$

Just as before, our assumptions guarantee that the numerator of $i(t,s)$ is decreasing in s while the denominator is increasing, and the same is true of $i(t,t)$. Thus, the results regarding how $i(t,s)$ and $i(t,t)$ fall in the approach to retirement remain entirely unchanged.

Part 5 of the Proposition is now sharper. There is still a jump in impatience as t crosses T, and moreover there will be an immediate reversion to standard geometric discounting post-retirement, at the rate ρ_f.

1.2. **Time-Varying Weights.** We might also think it reasonable to allow the weights α, β, ω to depend on time. We use here the formulation that weights are placed on the current self, the retirement self and all future selves. (The other case in which weights are placed on all "intermediate" future selves works in exactly the same way.) We can without any loss of generality choose a particular normalization. So we presume that at any date $t < T$, our individual places non-negative weights $\{\alpha(t), \beta(t), \omega(t)\}$ on the three sets of selves, and that these sum to 1; i.e., $\alpha(t) + \beta(t) + (N - t)\omega(t) = 1$. We impose the following restrictions:

[A.2] $\min\{\rho_f, \rho_b\} \min\{\alpha(t), \beta(t)\} > 2\omega(t)$ for all $t < T$.

[A.3] $\alpha(t)$ and $\beta(t)$ are non-decreasing.

Proposition 1 in the main text remains valid under these assumptions. As a matter of fact, observe that the time-varying nature of the weights only enters the picture when t is changing, which means that we need only reconsider Parts 4 and 5. To see why $i(t,t)$ is decreasing before retirement under these assumptions, observe that:

$$i(t,t) = \Big[\frac{\rho_f \alpha(t) - \rho_b \beta(t) e^{-\rho_b(T-t} - \omega(t)[1 - e^{-\rho_b(T-t}]}{\rho_b \alpha(t) + \rho_b \beta(t) e^{-\rho_b(T-t} + \omega(t)[1 - e^{-\rho_b(T-t}]}\Big]$$

$$(1) \qquad = \Big[\frac{(\rho_f \alpha(t) - \omega(t)) - (\rho_b \beta(t) - \omega(t)) e^{-\rho_b(T-t)}}{(\alpha(t) + \frac{\omega(t)}{\rho_b}) + (\beta(t) - \frac{\omega(t)}{\rho_b}) e^{-\rho_b(T-t)}}\Big]$$

By inspecting the formula (1) above and recalling the normalization, notice that $i(t,t)$ increases most rapidly when all the extra weight is handed over to $\alpha(t)$ as t increases. That is, it suffices to examine the case in which $\beta(t)$ and $\omega(t)$ are constant in t, and $\alpha'(t) = \omega(t)$. Differentiating $i(t,t)$ with respect to t and imposing these restrictions on weights, it can be shown that $i(t,t)$ is decreasing in t if and only if

$$(\rho_b \alpha(t) - \omega)(\rho_b \beta - \omega) e^{-\rho_b(T-t)} \geq \omega^2$$

from which it is clear that Condition A.2 is sufficient for this expression to hold.

For $t > T$, follows the analysis of the main text, or is exponential and constant if no weight is additionally placed on selves beyond T.

1.3. **Weights on Past Selves.** Yet another alternative is to place a uniform weight *all* selves, past or future, in addition to the privileged weights α and β on the current and retirement selves. Now, at any date, the "effective weight" on the current self t is α *plus* the discounted effect of all weights on "ancestral selves." As time passes, these ancestral selves will accumulate, effectively lending weight to α. Meanwhile, the effective weight on all other selves remains steady up to retirement. We therefore have a special case of the previous section on changing weights, where β and ω are unchanged, but where the cumulated weight $\alpha(t)$ on the current self can be viewed as

$$\alpha(t) = \alpha + \omega \int_0^t e^{-\rho_f(t-i)} di \tag{2}$$

It is possible to embed this model in the case of time-varying weights. First make Assumption A.2 on the original weights; i.e. assume $\min\{\rho_f, \rho_b\} \min\{\alpha, \beta\} > 2\omega$. Next, renormalize all weights by dividing by the sum $\alpha(t) + \beta + (N-t)\omega$, where $\alpha(t)$ is defined by (2). Then it is easy to see that A.2 and A.3 are both satisfied for the renormalized set of weights, and the earlier analysis applies.[1]

1.4. **No Retirement Self.** Finally, consider the case in which the agent places uniform weight on all future selves, without the presence of a privileged future retirement self. This is a model that might seem to natural to some: after all, if the agent is forward-looking enough to consider one future self, why not all consider all such selves equally? We claim, however, that such equal treatment is at odds with the basic premise in our paper, which places special emphasis on a stock-taking self in the future. Thus our weights are fundamentally "bi-modal" in nature, with privilege accorded to the shadow parent, as well as the current self.

That said, we consider briefly this alternative model. Indeed, the predictions in Proposition 1 generally not hold. Consider for instance, the expressions for $i(t,s)$ and $i(t,t)$:

$$i(t,s)\Big[\frac{(\rho_f\alpha - \omega)e^{-\rho_f(s-t)} + \omega e^{-\rho_b(N-s)}}{(\alpha - \frac{\omega}{\rho_f})e^{-\rho_f(s-t)} - \frac{\omega}{\rho_b}e^{-\rho_b(N-s)} + \frac{\omega}{\rho_f} + \frac{\omega}{\rho_b}}\Big]$$

and

$$i(t,t) = \Big[\frac{(\rho_f\alpha - \omega) + \omega e^{-\rho_b(N-t)}}{(\alpha + \frac{\omega}{\rho_b}) - \frac{\omega}{\rho_b}e^{-\rho_b(N-t)}}\Big].$$

Looking first at $i(t,s)$, it is clear that both the denominator and numerator are always positive, and hence negative discounting cannot occur. Further, note that the denominator of the expression for $i(t,s)$ above is decreasing in s, and hence $i(t,s)$ is *increasing* if and only if

$$e^{-\rho_f(s-t)}e^{\rho_b(N-s)} \geq \frac{\rho_f(\rho_f\alpha - \omega)}{\rho_b\omega}$$

For s close to t, this expression will hold for N sufficiently large, or alternatively for ω close enough to its upper bound $\rho_f\alpha$. Turning to $i(t,t)$, it is immediate that $i(t,t)$

[1]It is easy to check that $\alpha'(t) < \omega$ for all t, so that $\alpha(t) + \beta + \omega(N-t)$ falls with t. Therefore, if we renormalize, the new weights will satisfy [A.3], while [A.2] is unaffected by the renormalization and continues to apply.

is now in fact *increasing* in t. Previously, the conflict was largely between the current and retirement self. Given that the latter is initially situated far away, its influence was initially discounted. Over time, the conflict was felt to a greater degree as the influence of this future self became a proximate reality. Without such a focal self, the passage of time simply erodes the influence of future selves as they are passed, and thus the rate of impatience simply increases.