Online Appendix: Mistakes in Future Consumption, High MPCs Now

Chen Lian*

October 19, 2023

A Online Appendix A: Proofs

Proposition 1. Based on the consumption rules in (3) and (4) and the definition of $V_1(w_1)$ in (7), we know

$$V_1'(w_1) = \frac{1}{2}(1-\lambda_1)\,u'\left(\frac{1}{2}(1-\lambda_1)\,w_1\right) + \frac{1}{2}(1+\lambda_1)\,u'\left(\frac{1}{2}(1+\lambda_1)\,w_1\right).$$

Since u is quadratic, we know that u'' is a constant and

$$V_1'' = u'' \cdot \left[\frac{1}{4}(1-\lambda_1)^2 + \frac{1}{4}(1+\lambda_1)^2\right] = \frac{1}{2}u'' \cdot \left(1+\lambda_1^2\right). \tag{A.1}$$

This proves the first part of Proposition 1. From (6), we know

$$u'\left(c_0^{\text{Deliberate}}(\Delta)\right) = V_1'(w_1) \quad \text{with} \quad w_1 = \Delta - c_0^{\text{Deliberate}}(\Delta). \tag{A.2}$$

Taking a partial derivative with respect to Δ, we have

$$\phi_0^{\text{Deliberate}} = \frac{1}{2}\left(1+\lambda_1^2\right)\left(1-\phi_0^{\text{Deliberate}}\right) = \frac{\frac{1}{2}(1+\lambda_1^2)}{1+\frac{1}{2}(1+\lambda_1^2)} \tag{A.3}$$

This proves the second part of Proposition 1.

*Lian: UC Berkeley and NBER (email: chen_lian@berkeley.edu).

A generalization of Proposition 1. Consider the more general specification of $t = 1$ consumption rule in (10). Based on (7), we have

$$V_1'(w_1) = \frac{1}{2}(1-\lambda_1)u'\left(\frac{1}{2}(1-\lambda_1)w_1 - \bar{\lambda}_1\right) + \frac{1}{2}(1+\lambda_1)u'\left(\frac{1}{2}(1+\lambda_1)w_1 + \bar{\lambda}_1\right).$$

Because u is quadratic, we know that V_1'' shares the same formula as (A.1). As a result, $\phi_0^{\text{Deliberate}}$ shares the same formula as (A.3). Proposition 1 again follows. This explains that the key to the high-MPC result is mistakes in the future consumption's response to saving changes, λ_1.

Proof of Proposition 2. Based on each self's actual consumption rules $\{c_t(w_t)\}_{t=0}^{T-1}$, I can define the value function $V_t(w_t)$ as a function of the current state, w_t, for each $t \in \{0, \cdots, T-1\}$,

$$V_t(w_t) \equiv u(c_t(w_t)) + \sum_{k=1}^{T-t-1} \delta^k u(c_{t+k}(w_{t+k})) + \delta^{T-t} v(w_T), \tag{A.4}$$

subject to the budget in (13). For the last period T, we have $V_T(w_T) = v(w_T)$. Given (A.4), each self t's deliberate consumption rule defined in (14) satisfies

$$c_t^{\text{Deliberate}}(w_t) = \arg\max_{c_t} \; u(c_t) + \delta V_{t+1}(R(w_t - c_t)). \tag{A.5}$$

Moreover, for $t \in \{0, \cdots, T-1\}$, the value function $V_t(w_t)$ defined in (A.4) satisfies

$$V_t(w_t) = u(c_t(w_t)) + \delta V_{t+1}(R(w_t - c_t(w_t))). \tag{A.6}$$

Note that because I assume u, v, and c_t are third-order continuously differentiable, V_t is third-order continuously differentiable too.

The optimal deliberate consumption now is given by[1]

$$u'\left(c_t^{\text{Deliberate}}(w_t)\right) = R\delta V_{t+1}'\left(R\left(w_t - c_t^{\text{Deliberate}}(w_t)\right)\right). \tag{A.7}$$

We henceforth have:

$$u''\left(c_t^{\text{Deliberate}}(\bar{w}_t)\right)\frac{\partial c_t^{\text{Deliberate}}(\bar{w}_t)}{\partial w_t} = R^2\delta\left(1 - \frac{\partial c_t^{\text{Deliberate}}(\bar{w}_t)}{\partial w_t}\right)V_{t+1}''(\bar{w}_{t+1}),$$

where $\bar{w}_{t+1} = R(\bar{w}_t - \bar{c}_t) = R\left(\bar{w}_t - c_t^{\text{Deliberate}}(\bar{w}_t)\right)$ and

[1] This equation imposes the concavity of the continuation value $V_{t+1}(w_{t+1})$. This is true around the path $\{\bar{w}_t, \bar{c}_t\}_{t=0}^{T-1}$ because $V_{t+1}''(\bar{w}_{t+1}) = u''(\bar{c}_{t+1}) \cdot \Gamma_{t+1} < 0$, as proved below.

$$\frac{\partial c_t^{\text{Deliberate}}\left(\bar{w}_t\right)}{\partial w_t}=\frac{R^2\delta V_{t+1}^{''}\left(\bar{w}_{t+1}\right)}{u''\left(c_t^{\text{Deliberate}}\left(\bar{w}_t\right)\right)+R^2\delta V_{t+1}^{''}\left(\bar{w}_{t+1}\right)}. \tag{A.8}$$

From (A.6):

$$V_t^{'}\left(w_t\right)=\frac{\partial c_t\left(w_t\right)}{\partial w_t}u'\left(c_t\left(w_t\right)\right)+\left(1-\frac{\partial c_t\left(w_t\right)}{\partial w_t}\right)\delta RV_{t+1}^{'}\left(w_{t+1}\right), \tag{A.9}$$

and

$$\begin{aligned}V_t^{''}\left(\bar{w}_t\right)=&\left(\frac{\partial c_t\left(\bar{w}_t\right)}{\partial w_t}\right)^2u''\left(c_t\left(\bar{w}_t\right)\right)+\left(1-\frac{\partial c_t\left(\bar{w}_t\right)}{\partial w_t}\right)^2\delta R^2V_{t+1}^{''}\left(\bar{w}_{t+1}\right),\\&+\frac{\partial^2c_t\left(\bar{w}_t\right)}{\partial w_t^2}\left[u'\left(c_t\left(\bar{w}_t\right)\right)-\delta RV_{t+1}^{'}\left(\bar{w}_{t+1}\right)\right].\end{aligned}$$

At $\bar{w}_t$, because $c_t\left(\bar{w}_t\right)=c_t^{\text{Deliberate}}\left(\bar{w}_t\right)=\bar{c}_t$, from (A.7), we have $u'\left(c_t\left(\bar{w}_t\right)\right)=\delta RV_{t+1}^{'}\left(\bar{w}_{t+1}\right)$. As a result,

$$V_t^{''}\left(\bar{w}_t\right)=\left(\frac{\partial c_t\left(\bar{w}_t\right)}{\partial w_t}\right)^2u''\left(c_t\left(\bar{w}_t\right)\right)+\left(1-\frac{\partial c_t\left(\bar{w}_t\right)}{\partial w_t}\right)^2\delta R^2V_{t+1}^{''}\left(\bar{w}_{t+1}\right). \tag{A.10}$$

Define $\Gamma_t\equiv V_t^{''}\left(\bar{w}_t\right)/u''\left(c_t\left(\bar{w}_t\right)\right)$, $\phi_t^{\text{Deliberate}}\equiv\frac{\partial c_t^{\text{Deliberate}}\left(\bar{w}_t\right)}{\partial w_t}$, and

$$\phi_t\equiv\frac{\partial c_t\left(\bar{w}_t\right)}{\partial w_t}=\left(1-\lambda_t\right)\phi_t^{\text{Deliberate}}. \tag{A.11}$$

From (A.8) and (A.10), we have

$$\phi_t^{\text{Deliberate}}=\frac{\delta R^2\Gamma_{t+1}\frac{u''\left(\bar{c}_{t+1}\right)}{u''\left(\bar{c}_t\right)}}{1+\delta R^2\Gamma_{t+1}\frac{u''\left(\bar{c}_{t+1}\right)}{u''\left(\bar{c}_t\right)}} \tag{A.12}$$

and

$$\begin{aligned}\Gamma_t&=\phi_t^2+\left(1-\phi_t\right)^2\delta R^2\Gamma_{t+1}\frac{u''\left(\bar{c}_{t+1}\right)}{u''\left(\bar{c}_t\right)}\\&=\left(1-\lambda_t\right)^2\frac{\left(\delta R^2\Gamma_{t+1}\frac{u''\left(\bar{c}_{t+1}\right)}{u''\left(\bar{c}_t\right)}\right)^2}{\left(1+\delta R^2\Gamma_{t+1}\frac{u''\left(\bar{c}_{t+1}\right)}{u''\left(\bar{c}_t\right)}\right)^2}+\left(\frac{1+\lambda_t\delta R^2\Gamma_{t+1}\frac{u''\left(\bar{c}_{t+1}\right)}{u''\left(\bar{c}_t\right)}}{1+\delta R^2\Gamma_{t+1}\frac{u''\left(\bar{c}_{t+1}\right)}{u''\left(\bar{c}_t\right)}}\right)^2\delta R^2\Gamma_{t+1}\frac{u''\left(\bar{c}_{t+1}\right)}{u''\left(\bar{c}_t\right)}\\&=\frac{\left(\delta R^2\Gamma_{t+1}\frac{u''\left(\bar{c}_{t+1}\right)}{u''\left(\bar{c}_t\right)}\right)^2}{1+\delta R^2\Gamma_{t+1}\frac{u''\left(\bar{c}_{t+1}\right)}{u''\left(\bar{c}_t\right)}}\lambda_t^2+\frac{\delta R^2\Gamma_{t+1}\frac{u''\left(\bar{c}_{t+1}\right)}{u''\left(\bar{c}_t\right)}}{1+\delta R^2\Gamma_{t+1}\frac{u''\left(\bar{c}_{t+1}\right)}{u''\left(\bar{c}_t\right)}}.\end{aligned} \tag{A.13}$$

We know Γ_{t+1} and $\phi_t^{\text{Deliberate}}$ increases with $\{|\lambda_{t+k}|\}_{k=1}^{T-t-1}$ for all $t\in\{0,\cdots,T-1\}$. Proposition 2 then follows.

Proof of Corollary 1. For the pre-shock ($\bar{\Delta} = 0$) outcome, from (15), we have

$$u'(\bar{c}_t) = \delta R u'(\bar{c}_{t+1}).$$

As a result, for all $t \in \{0, \cdots, T-1\}$,

$$\frac{\bar{c}_{t+1}}{\bar{c}_t} = (\delta R)^{\frac{1}{\gamma}}. \tag{A.14}$$

Substituting it into (A.12) and (A.13), we have

$$\phi_t^{\text{Deliberate}} = \frac{\delta R^2 \Gamma_{t+1} (\delta R)^{-\frac{\gamma+1}{\gamma}}}{1 + \delta R^2 \Gamma_{t+1} (\delta R)^{-\frac{\gamma+1}{\gamma}}} = \frac{\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} \Gamma_{t+1}}{1 + \delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} \Gamma_{t+1}} \tag{A.15}$$

and

$$\Gamma_t = \frac{\left(\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} \Gamma_{t+1}\right)^2}{1 + \delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} \Gamma_{t+1}} \lambda_t^2 + \frac{\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} \Gamma_{t+1}}{1 + \delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} \Gamma_{t+1}} \equiv f(\Gamma_{t+1}),$$

with

$$f(x) \equiv \frac{\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} x}{1 + \delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} x} + \frac{\left(\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} x\right)^2}{1 + \delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} x} \lambda^2 = \frac{\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} x}{1 + \delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} x} \left(1 + \lambda^2 \delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} x\right).$$

We also know that $\Gamma_T = \frac{v''(\bar{w}_T)}{u''(\bar{c}_T)} = \kappa > 0$, where I use $\bar{w}_T = \bar{c}_T$.

Let $\Gamma = \frac{\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} - 1}{\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} \left[1 - \left(\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}}\right) \lambda^2\right]}$ denote the fixed point of f. That is $f(\Gamma) = \Gamma$. Moreover, as long as $\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} > 1$ and $|\lambda| < \left(\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}}\right)^{-\frac{1}{2}}$, we have $\Gamma > f(x) > x$ if $0 < x < \Gamma$; and $\Gamma < f(x) < x$ if $x > \Gamma$. We then have two cases:

1) If $\Gamma > \kappa$. We have $\Gamma > \Gamma_0 = f^T(\kappa) > f^{(T-1)}(\kappa) > \cdots > \kappa = \Gamma_T$. As a result, $\Gamma_0 = f^T(\kappa)$ converges to the fixed point Γ when $T \to +\infty$.

2) If $\Gamma < \kappa$. We have $\Gamma < \Gamma_0 = f^T(\kappa) < f^{(T-1)}(\kappa) < \cdots < \kappa = \Gamma_T$. As a result, $\Gamma_0 = f^T(\kappa)$ converges to the fixed point Γ when $T \to +\infty$.

Together, one way or another, as long as $\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} > 1$ and $|\lambda| < \left(\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}}\right)^{-\frac{1}{2}}$, $\Gamma_0 \to \Gamma$ when $T \to +\infty$. From (A.15), we have, when $T \to +\infty$,

$$\phi_0^{\text{Deliberate}} \to \phi^{\text{Deliberate}} \equiv \frac{\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} - 1}{\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} (1 - \lambda^2)}.$$

Proof of Proposition 3. Based on (7) and (21), we have

$$
\begin{aligned}
V_1'\left(w_1;\bar{\lambda}_1\right) &= \frac{1}{2}u'\left(\frac{1}{2}w_1-\bar{\lambda}_1\right)+\frac{1}{2}u'\left(\frac{1}{2}w_1+\bar{\lambda}_1\right)\\
\frac{\partial V_1'\left(w_1;\bar{\lambda}_1\right)}{\partial\bar{\lambda}_1} &= -\frac{1}{2}u''\left(\frac{1}{2}w_1-\bar{\lambda}_1\right)+\frac{1}{2}u''\left(\frac{1}{2}w_1+\bar{\lambda}_1\right)\\
\frac{\partial^2 V_1'\left(w_1;\bar{\lambda}_1\right)}{\partial\bar{\lambda}_1^2} &= \frac{1}{2}u'''\left(\frac{1}{2}w_1-\bar{\lambda}_1\right)+\frac{1}{2}u'''\left(\frac{1}{2}w_1+\bar{\lambda}_1\right).
\end{aligned}
$$

We have

$$
\frac{\partial V_1'\left(w_1;0\right)}{\partial\bar{\lambda}_1}=0 \quad \text{and} \quad \frac{\partial^2 V_1'\left(w_1;0\right)}{\partial\bar{\lambda}_1^2}>0. \tag{A.16}
$$

Based on (6) and (7), we have

$$
\begin{aligned}
u'\left(c_0^{\text{Deliberate}}\left(\Delta;\bar{\lambda}_1\right)\right) &= V_1'\left(\Delta-c_0^{\text{Deliberate}}\left(\Delta;\bar{\lambda}_1\right);\bar{\lambda}_1\right),\\
u''\left(c_0^{\text{Deliberate}}\left(\Delta;\bar{\lambda}_1\right)\right)\frac{\partial c_0^{\text{Deliberate}}\left(\Delta;\bar{\lambda}_1\right)}{\partial\bar{\lambda}_1} &= -V_1''\left(\Delta-c_0^{\text{Deliberate}}\left(\Delta;\bar{\lambda}_1\right);\bar{\lambda}_1\right)\frac{\partial c_0^{\text{Deliberate}}\left(\Delta;\bar{\lambda}_1\right)}{\partial\bar{\lambda}_1}\\
&\quad+\frac{\partial V_1'\left(\Delta-c_0^{\text{Deliberate}}\left(\Delta;\bar{\lambda}_1\right);\bar{\lambda}_1\right)}{\partial\bar{\lambda}_1}.
\end{aligned} \tag{A.17}
$$

Together with (A.16), we have

$$
\frac{\partial c_0^{\text{Deliberate}}\left(\Delta;0\right)}{\partial\bar{\lambda}_1}=\frac{\frac{\partial V_1'\left(\Delta-c_0^{\text{Deliberate}}(\Delta;0);0\right)}{\partial\bar{\lambda}_1}}{u''\left(c_0^{\text{Deliberate}}\left(\Delta;0\right)\right)+V_1''\left(\Delta-c_0^{\text{Deliberate}}\left(\Delta;0\right);0\right)}=0.
$$

and

$$
\begin{aligned}
u''\left(c_0^{\text{Deliberate}}\left(\Delta;0\right)\right)\frac{\partial^2 c_0^{\text{Deliberate}}\left(\Delta;0\right)}{\partial\bar{\lambda}_1^2} &= -V_1''\left(\Delta-c_0^{\text{Deliberate}}\left(\Delta;0\right);0\right)\frac{\partial^2 c_0^{\text{Deliberate}}\left(\Delta;0\right)}{\partial\bar{\lambda}_1^2}\\
&\quad+\frac{\partial^2 V_1'\left(\Delta-c_0^{\text{Deliberate}}\left(\Delta;0\right);0\right)}{\partial\bar{\lambda}_1^2}.
\end{aligned}
$$

As a result,

$$
\frac{\partial^2 c_0^{\text{Deliberate}}\left(\Delta;0\right)}{\partial\bar{\lambda}_1^2}=\frac{\frac{\partial^2 V_1'\left(\Delta-c_0^{\text{Deliberate}}(\Delta;0);0\right)}{\partial\bar{\lambda}_1^2}}{u''\left(c_0^{\text{Deliberate}}\left(\Delta\right);0\right)+V_1''\left(\Delta-c_0^{\text{Deliberate}}\left(\Delta;0\right);0\right)}<0.
$$

This proves Proposition 3.

B Online Appendix B: Additional Results

B.1 Partial Sophistication and the Role of Perceived Dynamic Inconsistency

The main analysis can accommodate a more general interpretation if I re-define deliberate consumption (14) based on current self 0′s perceived future consumption rules $\{\tilde{c}_t(w_t)\}_{t=1}^{T-1}$. That is, for $t \in \{0, \cdots, T-1\}$,

$$c_t^{\text{Deliberate}}(w_t) \equiv \arg\max_{c_t} u(c_t) + \sum_{k=1}^{T-t-1} \delta^k u(\tilde{c}_{t+k}(w_{t+k})) + \delta^{T-t} v(w_T), \tag{B.1}$$

subject to the budget (13). Future $\left\{c_t^{\text{Deliberate}}(w_t)\right\}_{t=1}^{T-1}$ can then be interpreted as the consumption that self 0 thinks is optimal at each future period $t \in \{1, \cdots, T-1\}$, given utility (11) and her perceived future consumption rules $\{\tilde{c}_t(w_t)\}_{t=1}^{T-1}$.

I can then define self 0′s perceived future mistakes $\left\{\tilde{\lambda}_t\right\}_{t=1}^{T-1}$ as how her perceived future consumption rules $\{\tilde{c}_t(w_t)\}_{t=1}^{T-1}$ deviate from what she deems optimal $\left\{c_t^{\text{Deliberate}}(w_t)\right\}_{t=1}^{T-1}$. Specifically, similar to Proposition 1, I impose that perceived mistakes in future consumption only take the form of mistakes in response to saving changes, while there are no mistakes in the absence of the shock Δ. That is, there are sequences $\{\bar{c}_t\}_{t=0}^{T-1}$ and $\{\bar{w}_t\}_{t=0}^{T}$ such that

$$(15) \quad \text{holds} \quad \text{and} \quad \tilde{c}_t(\bar{w}_t) = \bar{c}_t \quad \forall t \in \{1, \cdots, T-1\}. \tag{B.2}$$

We can then define self 0′s perceived future mistakes in response to saving changes $\left\{\tilde{\lambda}_t\right\}_{t=1}^{T-1}$ similar to (16):

$$\tilde{\phi}_t = \left(1 - \tilde{\lambda}_t\right) \phi_t^{\text{Deliberate}} \quad \forall t \in \{1, \cdots, T-1\}, \tag{B.3}$$

where $\tilde{\phi}_t \equiv \frac{\partial \tilde{c}_t(\bar{w}_t)}{\partial w_t}$ and $\phi_t^{\text{Deliberate}} = \frac{\partial \tilde{c}_t^{\text{Deliberate}}(\bar{w}_t)}{\partial w_t}$. We can then re-state Proposition 2 as how self 0′s perceived future mistakes $\left\{\tilde{\lambda}_t\right\}_{t=1}^{T-1}$ increase the current MPC, $\phi_0^{\text{Deliberate}}$.

Corollary B.1. *Based on the definition in (B.1) and (B.3), $\phi_0^{Deliberate} \equiv \frac{\partial c_0^{Deliberate}(\bar{w}_0)}{\partial w_0}$ increases with perceived future mistakes $\left|\tilde{\lambda}_t\right|$ for each $t \in \{1, \cdots, T-1\}$, as long as (B.2) holds.*

From this reinterpretation, the key to the high-MPC result is: the current self thinks that her future consumption will deviate from what she deems optimal. In other words, the essence is a form of *perceived dynamic inconsistency*. For the specific behavioral foundations considered in Section 4, such dynamic inconsistency can come from two sources. First, perceived differences in different selves' decision utility (such as present bias Corollaries B.6 and B.7). Second, violations of the law of iterated expectations (such as versions of inattention and diagnostic expectations in Corollaries B.3 – B.5).

One important example of how perceived future mistakes are determined is the case of partial sophistication as in O'Donoghue and Rabin (1999, 2001). That is, the current self has a partial understanding of future mistakes, and her perceived future mistake at t is given by:

$$\tilde{\lambda}_t = s\lambda_t, \tag{B.4}$$

where $s \in [0, 1]$ captures current self 0's degree of sophistication. There are two immediate lessons. First, partial sophistication suffices for all qualitative results about how future mistakes increase current MPCs. Second, current MPCs increase with the degree of sophistication.

Corollary B.2. *With (B.4), $\phi_0^{Deliberate}$ increases with current self 0's degree of sophistication s.*

Proof of Corollary B.1 and Corollary B.2. The proof of Proposition 2 goes through exactly, with perceived future mistakes $\tilde{\lambda}_t$ replacing the role of actual future mistakes λ_t. Corollary B.1 then follows. Corollary B.2 then follows directly from Corollary B.1 and (B.4).

B.2 Robustness Checks for the Numerical Illustration

Here, I conduct robustness checks with other parameterizations of the numerical exercise described in Section 3.

In Figure B.1, I first consider a higher relative degree of risk aversion $\gamma = 2$, while keeping other parameters constant. We can see that the deliberate MPC is still very similar to the one calculated analytically in Corollary 1. The main lesson on how future mistakes in response to saving changes increase the current MPC is unchanged.

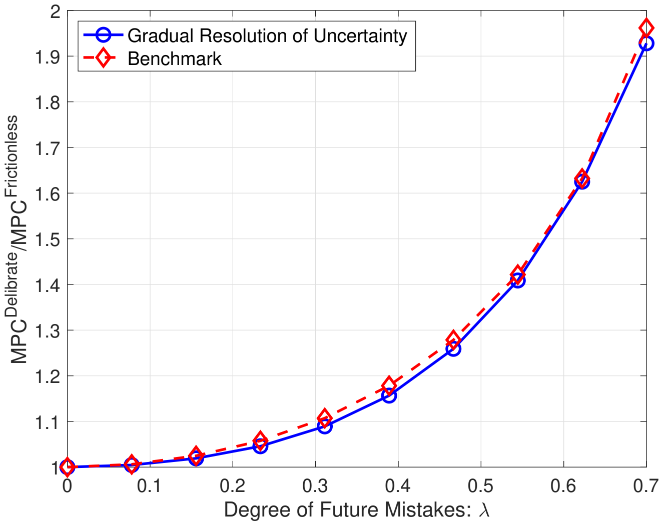

Figure B.1: Robustness Checks: $\gamma = 2$.

In Figure B.2, I then consider a higher return on saving $R = 1.07$, while keeping other parameters constant. We can see that the deliberate MPC is still very similar to the one calculated analytically in Corollary 1. The main lesson on how future mistakes in response to saving changes increase the current MPC is unchanged.

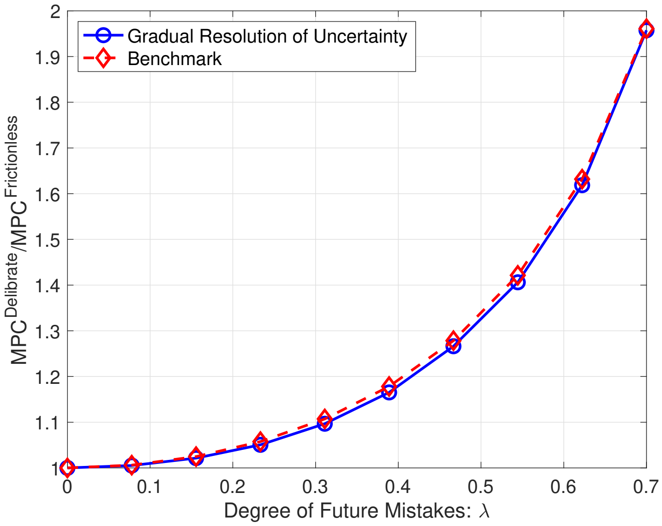

Figure B.2: Robustness Checks: $R = 1.07$.

In Figure B.3, I then consider a higher discount factor $\delta = 0.93$, while keeping other parameters constant. We can see that the deliberate MPC is still very similar to the one calculated analytically in Corollary 1. The main lesson on how future mistakes in response to saving changes increase the current MPC is unchanged.

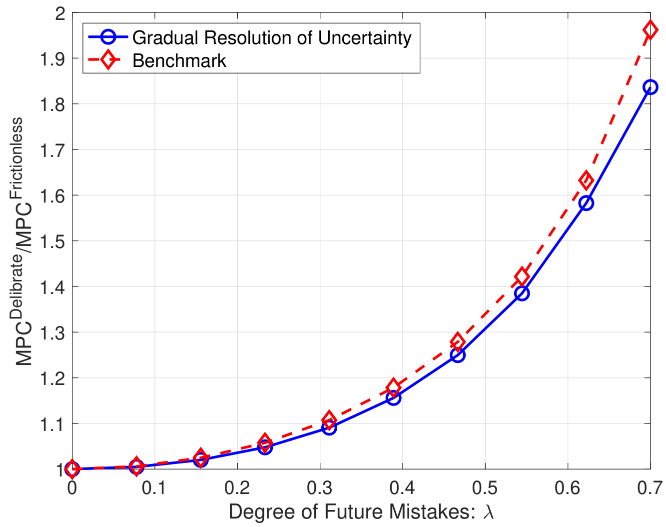

Figure B.3: Robustness Checks: $\delta = 0.93$.

B.3 The Precautionary Saving Motive and MPCs

A natural question is whether the precautionary saving motive driven by future mistakes in overall consumption level can also impact current MPCs. To illustrate this, consider the same environment as in Figure 1, with $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$; $\gamma = 1.1$; $\sigma = 1$; $\delta = 0.902$; $R = 1.04$; and $\underline{a} = 0$. Instead of mistakes in response to saving changes in (18), I focus on mistakes in the overall consumption level $\{\bar{\lambda}_t\}_{t=1}^{T-1}$. Specifically, similar to (21), these mistakes take the form of an additive deviation from the deliberate counterpart,

$$c_t(x_t) = \min\left\{-\frac{\underline{a}}{R} + x_t, c_t^{\text{Deliberate}}\left(x_t - \bar{\lambda}_t\right)\right\},$$

which makes sure the consumer will not violate her borrowing constraints despite her mistakes. As in Figure 1, with uncertainty, it is easier to write the actual consumption rule as a function of cash on hand x_t. When $\bar{\lambda}_t > 0$, self t's under-consumes (even in absence of the shock Δ). When $\bar{\lambda}_t < 0$, self t's over-consumes. From Figure B.4, we can see this type of additive future mistakes in overall consumption levels $\{\bar{\lambda}_t\}_{t=1}^{T-1}$ effectively does not matter for the current MPC $\phi_0^{\text{Deliberate}}$.[2]

[2] In Figure B.4, the x-axis is $\bar{\lambda}_t$ (in the unit of the standard deviation of the income risk $\sigma = 1$).

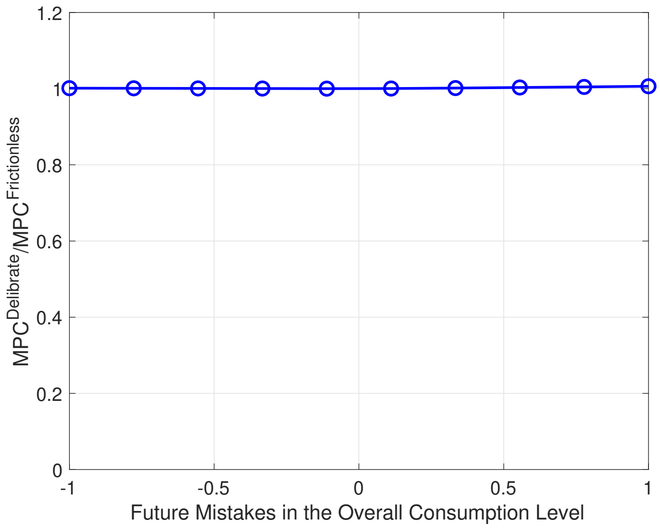

Figure B.4: The Precautionary Saving Motive and MPCs.

In applications, the essentially only possibility that future mistakes in overall consumption levels are large enough to matter for MPCs in (23) is that these mistakes take a multiplicative form

$$c_t(w_t) = c_t^{\text{Deliberate}}((1 - \Lambda_t) w_t), \tag{B.5}$$

where $\Lambda_t \neq 0$ captures self t's mistake. In this case, mistakes in overall consumption level can be very large: at w_t, self t behaves as if her wealth level were $(1 - \Lambda_t) w_t$, which can deviate significantly from w_t if Λ_t is away from zero. The precautionary saving motive due to those future mistakes can be large, which can impact MPCs nontrivially. Below, I provide a thorough analysis of this case. When the utility function is not that concave (EIS> 1), the high-MPC channel focused in the paper in Proposition 2 still dominates and future mistakes still unambiguously lead to high MPCs. When the utility function is very concave (EIS< 1), the precautionary saving channel may dominate.

B.4 Combined Multiplicative Mistakes

In some popular behavioral foundations, mistakes in response to saving changes come together with mistakes in the overall consumption level. The most classical example is the plain-vanilla version of hyperbolic discounting without commitment devices. In a homothetic case, such a combined mistake take a multiplicative form. This allows me to provide a sharp characterization on how such "combined" mistakes impact current MPCs.

Specifically, let the utility be given by the CRRA form with $u(c) = v(c) = \frac{c^{1-\gamma}}{1-\gamma}$. In this homothetic case, the frictionless consumption rule will be a multiple of the wealth w_t. Consider the case that the actual consumption rules inherit this property: for $t \in \{0, \cdots, T-1\}$,

$$c_t(w_t) = \Phi_t w_t \quad \text{and} \quad c_t^{\text{Deliberate}}(w_t) = \Phi_t^{\text{Deliberate}} w_t, \tag{B.6}$$

where, similar to (16), self t's mistake Λ_t is given by

$$\Phi_t = (1 - \Lambda_t)\, \Phi_t^{\text{Deliberate}}, \tag{B.7}$$

where $c_t^{\text{Deliberate}}(w_t)$ is defined based on Definition 1 as usual. In the homothetic environment here, future mistake Λ_t takes a multiplicative as in (B.5) and plays a dual role. When $\Lambda_t > 0$, self t both under-consumes overall and under-reacts to changes in w_t. When $\Lambda_t < 0$, self t both over-consumes overall and over-reacts to changes in w_t. In other words, Λ_t combines mistakes in response to saving changes with mistakes in the overall consumption level.

I can now study how these "combined" future mistakes $\{\Lambda_t\}_{t=1}^{T-1}$ impact the current consumption $c_0^{\text{Deliberate}}(w_0)$. In the homothetic environment here, $\Phi_0^{\text{Deliberate}}$ in (B.6) also plays a dual role. It determines both the current MPC and the overall current consumption level. Future mistakes' impact on $\Phi_0^{\text{Deliberate}}$ then combines the high-MPC effect in Proposition 2 and the low-consumption-level effect in Proposition 3.

Proposition B.1. *(1) When $\gamma < 1$, $\Phi_0^{Deliberate}$ increases with the future mistake $|\Lambda_t|$ in a neighborhood of $\Lambda_t = 0$ for each $t \in \{1, \cdots, T-1\}$.*

(2) When $\gamma > 1$, $\Phi_0^{Deliberate}$ decreases with the future mistake $|\Lambda_t|$ in a neighborhood of $\Lambda_t = 0$ for each $t \in \{1, \cdots, T-1\}$.

When the utility function is not that concave ($\gamma < 1$), the high-MPC channel in Proposition 2, which pushes $\Phi_t^{\text{Deliberate}}$ higher, dominates the precautionary saving channel in Proposition 3, which pushes $\Phi_t^{\text{Deliberate}}$ lower. When the utility function is very concave ($\gamma > 1$), the precautionary saving channel in Proposition 3, which pushes $\Phi_t^{\text{Deliberate}}$ lower, dominates.[3]

Proof of Proposition B.1. I guess and verify the continuation value function defined in (A.4) takes the form of

$$V_t(w_t) = \kappa_t \frac{w_t^{1-\gamma}}{1-\gamma}$$

[3]One may wonder how to reconcile Proposition B.1 with Figure B.4, where the precautionary saving motive does not matter much for the MPC. Note that, in Figure B.4, as the rest of the paper, mistakes in overall consumption level take the form of an "additive" deviation from the deliberate counterpart, similar to (21). Figure B.4 shows that the precautionary saving motive driven by those types of mistakes is unlikely to matter for the MPC. On the other hand, mistakes in (B.7) take a multiplicative form. It leads to large deviations from the deliberation counterpart and large precautionary saving motives in Proposition B.1.

for $t \in \{0, \cdots, T\}$. I work with backward induction. At T, I have:

$$V_T(w_T) = \frac{w_T^{1-\gamma}}{1-\gamma} \quad \text{and} \quad \kappa_T = 1.$$

For each $t \leq T-1$, from (A.7), the deliberate consumption is given by

$$\left(c_t^{\text{Deliberate}}(w_t)\right)^{-\gamma} = \delta R \kappa_{t+1} \left(R\left(w_t - c_t^{\text{Deliberate}}(w_t)\right)\right)^{-\gamma}$$
$$\Phi_t^{\text{Deliberate}} = \frac{(\delta\kappa_{t+1})^{-\frac{1}{\gamma}}(R)^{1-\frac{1}{\gamma}}}{1+(\delta\kappa_{t+1})^{-\frac{1}{\gamma}}(R)^{1-\frac{1}{\gamma}}} \tag{B.8}$$

From (B.7), the actual consumption is given by

$$\Phi_t = \frac{(1-\Lambda_t)(\delta\kappa_{t+1})^{-\frac{1}{\gamma}}(R)^{1-\frac{1}{\gamma}}}{1+(\delta\kappa_{t+1})^{-\frac{1}{\gamma}}(R)^{1-\frac{1}{\gamma}}}.$$

From the recursive formulation of the value function in (A.6), we have:

$$\kappa_t = \left((1-\Lambda_t)\frac{(\delta\kappa_{t+1})^{-\frac{1}{\gamma}}(R)^{1-\frac{1}{\gamma}}}{1+(\delta\kappa_{t+1})^{-\frac{1}{\gamma}}(R)^{1-\frac{1}{\gamma}}}\right)^{1-\gamma} + \delta\kappa_{t+1}R^{1-\gamma}\left(1-(1-\Lambda_t)\frac{(\delta\kappa_{t+1})^{-\frac{1}{\gamma}}(R)^{1-\frac{1}{\gamma}}}{1+(\delta\kappa_{t+1})^{-\frac{1}{\gamma}}(R)^{1-\frac{1}{\gamma}}}\right)^{1-\gamma}.$$

Define

$$f(\Lambda,\kappa) \equiv \left((1-\Lambda)\frac{(\delta\kappa)^{-\frac{1}{\gamma}}(R)^{1-\frac{1}{\gamma}}}{1+(\delta\kappa)^{-\frac{1}{\gamma}}(R)^{1-\frac{1}{\gamma}}}\right)^{1-\gamma} + \delta\kappa R^{1-\gamma}\left(1-(1-\Lambda)\frac{(\delta\kappa)^{-\frac{1}{\gamma}}(R)^{1-\frac{1}{\gamma}}}{1+(\delta\kappa)^{-\frac{1}{\gamma}}(R)^{1-\frac{1}{\gamma}}}\right)^{1-\gamma}.$$

We have

$$\frac{\partial f(\Lambda,\kappa)}{\partial\Lambda} = -(1-\gamma)\left(\Phi^{\text{Deliberate}}\right)^{1-\gamma}(1-\Lambda)^{-\gamma} + (1-\gamma)\Phi^{\text{Deliberate}}\delta\kappa R^{1-\gamma}\left(1-(1-\Lambda)\Phi^{\text{Deliberate}}\right)^{-\gamma},$$

where

$$\Phi^{\text{Deliberate}} = \frac{(\delta\kappa)^{-\frac{1}{\gamma}}(R)^{1-\frac{1}{\gamma}}}{1+(\delta\kappa)^{-\frac{1}{\gamma}}(R)^{1-\frac{1}{\gamma}}} \in (0,1). \tag{B.9}$$

Moreover,

$$\frac{\partial^2 f(\Lambda,\kappa)}{\partial\Lambda^2} = -\gamma(1-\gamma)\left(\Phi^{\text{Deliberate}}\right)^{1-\gamma}(1-\Lambda)^{-\gamma-1} - \gamma(1-\gamma)\left(\Phi^{\text{Deliberate}}\right)^2\delta\kappa R^{1-\gamma}\left(1-(1-\Lambda)\Phi^{\text{Deliberate}}\right)^{-\gamma-1}.$$

Together with (B.9), we have

$$
\begin{aligned}
\frac{\partial f(0,\kappa)}{\partial \Lambda} &= -(1-\gamma)\left(\Phi^{\text{Deliberate}}\right)^{1-\gamma} + (1-\gamma)\,\Phi^{\text{Deliberate}}\delta\kappa R^{1-\gamma}\left(1-\Phi^{\text{Deliberate}}\right)^{-\gamma} = 0 \\
\frac{\partial^2 f(0,\kappa)}{\partial \Lambda^2} &= -\gamma(1-\gamma)\left(\Phi^{\text{Deliberate}}\right)^{1-\gamma} - \gamma(1-\gamma)\left(\Phi^{\text{Deliberate}}\right)^{2}\delta\kappa R^{1-\gamma}\left(1-\Phi^{\text{Deliberate}}\right)^{-\gamma-1} \\
&= -\gamma(1-\gamma)\left(\Phi^{\text{Deliberate}}\right)^{2-\gamma}\left[\left(\Phi^{\text{Deliberate}}\right)^{-1} + \left(\Phi^{\text{Deliberate}}\right)^{\gamma}\delta\kappa R^{1-\gamma}\left(1-\Phi^{\text{Deliberate}}\right)^{-\gamma-1}\right] \\
&= -\gamma(1-\gamma)\left(\Phi^{\text{Deliberate}}\right)^{2-\gamma}\left[\left(\Phi^{\text{Deliberate}}\right)^{-1} + \left(1-\Phi^{\text{Deliberate}}\right)^{-1}\right].
\end{aligned}
$$

So

$$
\frac{\partial^2 f(0,\kappa)}{\partial \Lambda^2} > 0 \iff \gamma > 1.
$$

Moreover,

$$
\frac{\partial f(0,\kappa)}{\partial \kappa} > 0.
$$

Together, this means

1. When $\gamma < 1$, $\kappa_t^{\text{Deliberate}}$ decreases with mistake $|\Lambda_{t+k}|$ in a neighborhood of $\Lambda_{t+k} = 0$ for each $k \in \{0, \cdots, T-t-1\}$.
2. When $\gamma > 1$, $\kappa_t^{\text{Deliberate}}$ increases with mistake $|\Lambda_{t+k}|$ in a neighborhood of $\Lambda_{t+k} = 0$ for each $k \in \{0, \cdots, T-t-1\}$.

Together with (B.8), we arrive at Proposition B.1.

B.5 Inattention

Based on the perceived $w_t^p(w_t)$ in (24), the actual consumption rule for each self $t \in \{0, \cdots, T-1\}$ is given by

$$
c_t(w_t) = \arg\max_{c_t}\ u(c_t) + \delta V_{t+1}\left(R\left(w_t^p(w_t) - c_t\right)\right), \tag{B.10}
$$

where the continuation value function V_{t+1} is defined as in (A.4). To isolate the impact of future inattention on current consumption, the deliberate consumption is defined as in (A.5). As a corollary of Proposition 2, future consumption mistakes in the form of inattention lead to higher current MPCs.

Corollary B.3. *$\phi_0^{Deliberate}$ increases with future selves' degrees of inattention $\{\lambda_t\}_{t=1}^{T-1}$ if the default wealth w_t^d is the pre-shock value $\bar{w}_t$ for all t.*

In the inattention case studied in Corollary B.3, each self's perceived w_t is given by a deterministic weighted average between the actual w_t and the default. This follows the sparsity approach in Gabaix (2014). An alternative way to model inattention is through noisy signals (Sims, 2003). These two approaches will lead to similar predictions on MPCs.

Specifically, it's well known that one needs linear consumption rules (quadratic utility) and Normally distributed fundamentals to obtain tractability with noisy signals. I hence consider the quadratic utility case of the problem set up in Section 2. I assume a Normally distributed exogenous shock, i.e., $\Delta \sim \mathcal{N}\left(0, \sigma^2\right)$.[4] Unlike the main analysis, each self t's knowledge of the current w_t is now summarized by a noisy signal $x_t = w_t + \epsilon_t$, while $\epsilon_t \sim \mathcal{N}\left(0, \sigma^2_{\epsilon_t}\right)$ and is independent of Δ and other ϵ_t. In this case, each self understands that her signal is noisy and tries to infer her actual w_t from the signal.

$$E\left[w_t \mid x_t\right] = (1-\lambda_t)x_t + \lambda_t \bar{w}_t, \tag{B.11}$$

where $\lambda_t = \frac{Var(\epsilon_t)}{Var(w_t)+Var(\epsilon_t)} \in [0,1]$ depends negatively on the signal-to-noise ratio of her signal about w_t.

Based on this signal, the actual consumption rule of each self t is given by

$$c_t\left(x_t\right) = \arg\max_{c_t}\ u\left(c_t\right) + \delta E\left[V_{t+1}\left(R\left(w_t - c_t\right)\right)|x_t\right], \tag{B.12}$$

where the continuation value function V_{t+1} is defined in (A.4) and the deliberate consumption is defined in (A.5), taking future selves' inattention to permanent income as given. The deliberate MPC is given by $\phi_t^{\text{Deliberate}} \equiv \frac{\partial c_t^{\text{Deliberate}}(w_t)}{\partial w_t}$.[5] We have

Corollary B.4. *Each self t's deliberate MPC $\phi_t^{Deliberate}$ increases with future selves' degrees of inattention $\{\lambda_{t+k}\}_{k=1}^{T-t-1}$.*

As discussed in the main text, the essence of the high-MPC result is that the current self thinks that her future consumption will deviate from what she deems optimal. For the belief-based distortion considered in Corollaries B.3 and B.4, such perceived dynamic inconsistency comes in the form of violations of the law of iterated expectations. That is,

$$E_t\left[E_{t+1}\left[w_{t+1}\right]\right] \neq E_t\left[w_{t+1}\right], \tag{B.13}$$

where $E_t[\cdot]$ captures self t's belief. To see this, note that, in the sparsity case (24), we have $E_t\left[w_{t+1}\right] = R\left(w_t^p\left(w_t\right) - c_t\right)$ and $E_t\left[E_{t+1}\left[w_{t+1}\right]\right] = R\left(1-\lambda_{t+1}\right)\left(w_t^p\left(w_t\right) - c_t\right) + \lambda_{t+1} w_{t+1}^d$, which leads to (B.13). In the noisy signal case, $E_t\left[w_{t+1}\right] = R\left(E\left[w_t \mid x_t\right] - c_t\right)$ and $E_t\left[E_{t+1}\left[w_{t+1}\right]\right] = R\left(1-\lambda_{t+1}\right)\left(E\left[w_t \mid x_t\right] - c_t\right) + \lambda_{t+1}\bar{w}_{t+1}$, which leads to (B.13).

[4] This together with the linear actual consumption rule from (B.12) guarantees that each w_t is Normally distributed too.

[5] Since $c_t^{\text{Deliberate}}\left(w_t\right)$ is linear with quadratic utility, $\phi_t^{\text{Deliberate}}$ does not depend on w_t.

This discussion also helps illustrate what forms of future inattention generate relevant mistakes that lead to higher current MPCs. To break law of iterated expectations, it is crucial that the latter self's information set does not nest the earlier self's information set. In other words, some forms of bounded recall is needed. The classical formulation of Rational Inattention (Sims, 2003), which maintains perfect recall and law of iterated expectations, will not generate relevant mistakes that lead to higher current MPCs. On the other hand, modern formulations of Rational Inattention incorporating bounded recall (Da Silveira, Sung and Woodford, 2020; Afrouzi et al., 2020) and the sparsity model studied above break law of iterated expectations and will generate relevant mistakes leading to higher current MPCs. For example, in the noisy signal case in Corollary B.4, self $t+1$'s information (summarized by x_{t+1}) does not nest self t's information x_t.

Proof of Corollary B.3. From (24) and (B.10), we know the degree of inattention λ_t here corresponds to the degree of mistake in (16). Corollary B.3 then follows from Proposition 2.

Proof of Corollary B.4. The value in (A.6) is now given by

$$V_t(w_t) = \int \left[u\left(c_t\left(w_t+\epsilon_t\right)\right) + \delta V_{t+1}\left(R\left(w_t - c_t\left(w_t+\epsilon_t\right)\right)\right)\right] f_t\left(\epsilon_t\right) d\epsilon_t, \tag{B.14}$$

where $f_t(\cdot)$ is the p.d.f. for $\epsilon_t \sim \mathcal{N}\left(0, \sigma^2_{\epsilon_t}\right)$. Similar to the proof of Proposition 2, I use $\Gamma_t \equiv V_t''/u'' > 0$ to define the "concavity" of the continuation value function. From (A.5), the deliberate MPC is then given by

$$\phi_t^{\text{Deliberate}} = \frac{\delta R^2 \Gamma_{t+1}}{1+\delta R^2 \Gamma_{t+1}}.$$

From the actual consumption in (B.12), we have[6]

$$\phi_t = \left(1-\lambda_t\right)\phi_t^{\text{Deliberate}} = \frac{\left(1-\lambda_t\right)\delta R^2\Gamma_{t+1}}{1+\delta R^2\Gamma_{t+1}}, \tag{B.15}$$

where From (B.14), we have

$$\frac{\partial V_t(w_t)}{\partial w_t} = \int \left[\phi_t u'\left(c_t\left(w_t+\epsilon_t\right)\right) + \left(1-\phi_t\right)\delta R \frac{\partial V_{t+1}\left(w_{t+1}\right)}{\partial w_{t+1}}\right] f_t\left(\epsilon_t\right) d\epsilon_t,$$

[6] $\phi_t \equiv \frac{\partial c_t(x_t)}{\partial x_t}$. Since $c_t(\cdot)$ is linear with quadratic utility, ϕ_t does not depend on x_t.

where $w_{t+1} = R\left(w_t - c_t\left(w_t + \epsilon_t\right)\right)$. The recursive formulation of Γ_t is then given by

$$\begin{aligned}\Gamma_t &= \left(\phi_t\right)^2 + \left(1-\phi_t\right)^2 \Gamma_{t+1}\delta R^2 \\ &= \frac{\left(\delta R^2 \Gamma_{t+1}\right)^2}{1+\delta R^2\Gamma_{t+1}}\lambda_t^2 + \frac{\delta R^2\Gamma_{t+1}}{1+\delta R^2\Gamma_{t+1}}.\end{aligned}$$

We have Γ_t increases in $\{\lambda_{t+k}\}_{k=0}^{T-t-1}$. Corollary B.4 then follows.

B.6 Diagnostic Expectations

To follow closely Bianchi, Ilut and Saijo (2023), I use the three-period example with quadratic utility in Section 1. In the final period $t = 2$, as in (3), the consumer consumes out of all her remaining saving, $c_2(w_2) = w_2$. In the middle period $t = 1$, a higher saving w_1 triggers more vivid memories of good times for the consumer, which leads her to become overly optimistic about c_2. On the other hand, a lower saving w_1 triggers more vivid memories of bad times for the consumer, which leads her to become overly pessimistic about c_2. Mathematically, the consumer's consumption $c_1(w_1)$ at $t = 1$ is given by

$$u'\left(c_1\left(w_1\right)\right) = E_1^{\theta}\left[u'\left(c_2\left(w_2\right)\right)\right], \tag{B.16}$$

where $E_1^{\theta}[\cdot]$ captures her diagnostic expectation given by[7]

$$E_1^{\theta}\left[c_2\left(w_2\right)\right] = \left(1+\theta\right)c_2\left(w_2\right), \tag{B.17}$$

and $\theta > 0$ measures the degree of over-reaction in expectation, i.e., the representativeness distortion. Together, we have

$$c_1\left(w_1\right) = \frac{1+\theta}{2+\theta}w_1. \tag{B.18}$$

In other words, since the diagnostic expectation at $t = 1$ about c_2 over-reacts to saving changes in w_1, consumption c_1 also over-reacts to saving changes. Based on (B.18), one can then define the deliberate consumption at $t = 0$ as in (5). As a corollary to Proposition 1, future diagnostic expectations increase the current MPC.

Corollary B.5. *The current MPC* $\phi_0^{Deliberate}$ *strictly increases with the degree of future diagnostic expectations* θ.

[7]The case studied here corresponds to the "distant memory" $J \geq 2$ case in Bianchi, Ilut and Saijo (2023). That is, the reference point for $E_1^{\theta}[\cdot]$ is invariant to the shock Δ and decisions at $t = 0$. It is instead given by the pre-shock outcome $\bar{c}_2 = \bar{w}_2 = 0$ in Section 1.

The result can also be easily extended to the concave case in Proposition 2. This is because diagnostic expectations are precisely about belief over-reaction to shocks, while there are no mistakes in the overall expectations level. As a result, Proposition 2 applies.

As discussed in the main text, the essence of the high-MPC result is that the current self thinks that her future consumption will deviate from what she deems optimal. For the belief-based distortion considered in Corollary B.5, such perceived dynamic inconsistency comes in the form of violations of the law of iterated expectations. From (B.17), we can see the violation easily:[8]

$$E_0^{\theta}\left[E_1^{\theta}\left[c_2\left(w_2\right)\right]\right] = \left(1+\theta\right) E_0^{\theta}\left[c_2\left(w_2\right)\right] \neq E_0^{\theta}\left[c_2\left(w_2\right)\right].$$

Proof of Corollary B.5. From (4) and (B.18), we know $\lambda = -\frac{\theta}{2+\theta}$. And Corollary B.5 follows from Proposition 1.

B.7 Hyperbolic Discounting

My framework can also accommodate hyperbolic discounting (e.g. Laibson, 1997; Barro, 1999; Angeletos et al., 2001; Harris and Laibson, 2001). Let me start with the case with commitment devices, e.g., the original Laibson (1997) and Angeletos et al. (2001). This case only introduces mistakes in response to saving changes and will map to Proposition 2.

Specifically, the consumer can put her saving in illiquid assets with costly withdrawals to avoid over-consumption driven by the present bias. In absence of shocks, she can achieve optimal consumption through this commitment device. That is, (15) holds. On the other hand, in response to shocks, the commitment device no longer prevents her from consuming sub-optimally. In this case, a presently biased future self t's consumption will be given by

$$c_t\left(w_t\right) = \bar{c}_t + 1\cdot\left(w_t - \bar{w}_t\right), \tag{B.19}$$

for all w_t in a neighborhood of $\bar{w}_t$.[9]

Given (B.19), I can define the deliberate consumption rule $c_t^{\text{Deliberate}}\left(w_t\right)$ as usual. As a corollary of Proposition 2, mistakes in future consumption driven by future present biases will necessarily increase the current MPC.

[8] The "recent memory" $J = 1$ case in Bianchi, Ilut and Saijo (2023) instead does not break law of iterated expectations and does not lead to perceived dynamic inconsistency. This is because, in this case, the reference point for $E_1^{\theta}[\cdot]$ moves with decisions at $t = 0$.

[9] To derive (B.19). First, consider a small positive deviation of w_t away from $\bar{w}_t$. Because $u'\left(\bar{c}_t\right) = \delta V'\left(\bar{w}_{t+1}\right)$, $u'\left(\bar{c}_t\right) > \beta_t \delta V'\left(\bar{w}_{t+1}\right)$ for all $\beta_t < 1$. As a result, present bias will prompt self t to consume out of all the positive deviation $w_t - \bar{w}_t$ and (B.19) holds. Second, consider a small negative deviation of w_t away from $\bar{w}_t$. Because of the costly withdrawals from the illiquid assets, self t can only use c_t to absorb the negative deviation $w_t - \bar{w}_t$ and (B.19) again holds.

Corollary B.6. *Given any strictly concave utility functions u and v, (15), and the hyperbolic-discounting future consumption rules (B.19), $\phi_0^{Deliberate} \equiv \frac{\partial c_0^{Deliberate}(\bar{w}_0)}{\partial w_0} \geq \phi_0^{Frictionless}$, where $\phi_0^{Frictionless}$ is the frictionless MPC at $\bar{w}_0$.*

Now let us turn to the plain vanilla beta-delta model without access to illiquid assets as a commitment device (Barro, 1999; Harris and Laibson, 2001). Here, hyperbolic discounting leads to both mistakes in response to saving changes and mistakes in overall consumption levels. Specifically, the actual future consumption rule of self t is given by

$$c_t(w_t) = \arg\max_{c_t} u(c_t) + \delta\beta_t V_{t+1}(R(w_t - c_t)), \tag{B.20}$$

where $\beta_t \in [0,1]$ captures self t's present bias, which leads to both types of mistakes. Both the focused high-MPC channel in Proposition 2 (because of future mistakes in response to saving changes) and the precautionary saving channel in Proposition 3 (because of mistakes in overall consumption levels) are at force. With CRRA utility, this case maps to the multiplicative case in Proposition B.1.

Corollary B.7. *When $u(x) = v(x) = \frac{x^{1-\gamma}}{1-\gamma}$, the hyperbolic discounting case in (B.20) is nested by Proposition B.1. When $\gamma < 1$, the current MPC $\phi_0^{Deliberate}$ increases with future selves' present bias, i.e., decreases with each $\{\beta_t\}_{t=1}^{T-1}$.*

Similar to the discussion after Proposition B.1, when the utility function is not that concave (EIS>1), the high-MPC channel focused in the paper in Proposition 2 dominates and future mistakes still unambiguously lead to high MPCs. When the utility function is very concave (EIS< 1), the precautionary saving channel may dominate. This is consistent with the result in Maxted (2022).

Proof of Corollary B.6. This follows directly from (B.19) and Proposition 2.

Proof of Corollary B.7. From (B.20), we have

$$u'(c_t(w_t)) = \delta\beta_t R V'_{t+1}(R(w_t - c_t(w_t))). \tag{B.21}$$

From (A.5), we have

$$u'\left(c_t^{\text{Deliberate}}(w_t)\right) = \delta R V'_{t+1}\left(R\left(w_t - c_t^{\text{Deliberate}}(w_t)\right)\right). \tag{B.22}$$

Comparing (B.21) and (B.22), we have:

$$\phi_t = \beta_t^{-\frac{1}{\gamma}} \phi_t^{\text{Deliberate}}.$$

Corollary B.7 then follows directly from Proposition B.1.

B.8 Stochastic Epsilon-mistakes

Here, we study stochastic mistakes that do not bias the consumer's response to saving changes in a particular way. That is, $\lambda_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_t^2)$ in (16). Define the deliberate consumption $c_t^{\text{Deliberate}}(w_t)$ as usual given (16). Similar to Proposition 2, future stochastic mistakes in response to saving changes lead to higher current MPCs.

Corollary B.8. *If* $\lambda_t \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_t^2)$, $\phi_0^{Deliberate} \equiv \frac{\partial c_0^{Deliberate}(\bar{w}_0)}{\partial w_0}$ *increases with the variances in future selves' stochastic mistakes,* σ_t^2, *for* $t \in \{1, \cdots, T-1\}$,

This result means that, even if future consumption's response may be correct on average, stochastic mistakes in response to saving changes still increase current MPCs.

Proof of Corollary B.8. This case is not directly nested in Proposition 2, as the actual consumption rule is stochastic. But the proof is essentially unchanged.

The value function in (A.6) is now given by

$$V_t(w_t) = E_t\left[u(c_t(w_t)) + \delta V_{t+1}(R(w_t - c_t(w_t)))\right],$$

where $E_t[\cdot]$ averages over the potential realizations of λs. The deliberate consumption in (A.5) is unchanged.

In the proof of Proposition 2, the deliberate MPC is still given by (A.12), but (A.13) becomes

$$\begin{aligned}
\Gamma_t &= \mathbb{E}^{\lambda_t}\left[\left(\phi_t^{\text{Deliberate}}(1-\lambda_t)\right)^2 + \left(1 - \phi_t^{\text{Deliberate}}(1-\lambda_t)\right)^2 \Gamma_{t+1}\delta R^2 \frac{u''(\bar{c}_{t+1})}{u''(\bar{c}_t)}\right] \\
&= \frac{\delta R^2 \Gamma_{t+1} \frac{u''(\bar{c}_{t+1})}{u''(\bar{c}_t)}}{1 + \delta R^2 \Gamma_{t+1} \frac{u''(\bar{c}_{t+1})}{u''(\bar{c}_t)}} + \sigma_t^2 \frac{\left(\delta R^2 \Gamma_{t+1} \frac{u''(\bar{c}_{t+1})}{u''(\bar{c}_t)}\right)^2}{1 + \delta R^2 \Gamma_{t+1} \frac{u''(\bar{c}_{t+1})}{u''(\bar{c}_t)}}.
\end{aligned}$$

As a result, Γ_t increases with $\left\{\sigma_{t+k}^2\right\}_{k=0}^{T-t-1}$. Corollary B.8 then follows directly from (A.12).

References

Afrouzi, Hassan, Spencer Yongwook Kwon, Augustin Landier, Yueran Ma, and David Thesmar. (2020) "Overreaction in expectations: Evidence and theory." *NBER Working Paper.*

Angeletos, George-Marios, David Laibson, Andrea Repetto, Jeremy Tobacman, and Stephen Weinberg. (2001) "The hyperbolic consumption model: Calibration, simulation, and empirical evaluation." *Journal of Economic Perspectives*, 15(3): 47–68.

Barro, Robert. (1999) "Ramsey meets Laibson in the Neoclassical growth model." *The Quarterly Journal of Economics*, 114(4): 1125–1152.

Bianchi, Francesco, Cosmin Ilut, and Hikaru Saijo. (2023) "Diagnostic business cycles." *Review of Economic Studies.*

Da Silveira, Rava Azeredo, Yeji Sung, and Michael Woodford. (2020) "Optimally imprecise memory and biased forecasts." *NBER Working Paper.*

Gabaix, Xavier. (2014) "A sparsity-based model of bounded rationality." *Quarterly Journal of Economics*, 129(4).

Harris, Christopher, and David Laibson. (2001) "Dynamic choices of hyperbolic consumers." *Econometrica*, 69(4): 935–957.

Laibson, David. (1997) "Golden eggs and hyperbolic discounting." *The Quarterly Journal of Economics*, 112(2): 443–478.

Maxted, Peter. (2022) "Present bias in consumption-saving models: A tractable continuous-time approach."

O'Donoghue, Ted, and Matthew Rabin. (1999) "Doing it now or later." *American Economic Review*, 89(1): 103–124.

O'Donoghue, Ted, and Matthew Rabin. (2001) "Choice and procrastination." *Quarterly Journal of Economics*, 116(1): 121–160.

Sims, Christopher. (2003) "Implications of rational inattention." *Journal of Monetary Economics*, 50(3): 665–690.