NBER WORKING PAPER SERIES

MISTAKES IN FUTURE CONSUMPTION, HIGH MPCS NOW

Chen Lian

Working Paper 29517
http://www.nber.org/papers/w29517

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2021

I am extremely grateful to Marios Angeletos, Alp Simsek and Ricardo Caballero for continuous guidance through the project. I am grateful to Hassan Afrouzi (discussant), Nick Barberis, Stefano Dellavigna, Yuriy Gorodnichenko, Pierre-Olivier Gourinchas, Jonathon Hazell, Martin Holm, Julian Kozlowski, Yueran Ma, Emi Nakamura, Michaela Pagel, Jonathan Parker, Karthik Sastry, Benjamin Schoefer, Frank Schilbach, Jon Steinsson, Dmitry Taubinsky, Adam Szeidl, Andrei Shleifer, Ludwig Straub, Ivan Werning, and Muhamet Yildiz, and seminar participants at Berkeley, Berkeley-Chicago Behavioral Seminar, BI Norwegian Business School, China Star Tour, CICM 2021, Chicago Booth, Columbia, Duke, Duke Fuqua, MIT, Yale SOM, SED 2021, Stanford GSB for very helpful comments and discussions. I am grateful to Martin Holm for sharing his data. This paper is a more general version of and replaces the job market version named "Consumption with Imperfect Perception of Wealth." I acknowledge the financial support from Alfred P. Sloan Foundation Pre-doctoral Fellowship in Behavioral Macroeconomics, awarded through the NBER. Bruno Smaniotto provides excellent research assistance. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Mistakes in Future Consumption, High MPCs Now
Chen Lian
NBER Working Paper No. 29517
November 2021
JEL No. D90,E21,E70,G40,G51

ABSTRACT

In a canonical intertemporal consumption problem, I show how anticipation of future consumption mistakes leads to higher current marginal propensities to consume (MPCs). This result is driven by mistakes in future consumption's response to saving changes (i.e., changes in asset balances) and is independent of their specific behavioral foundations. My framework can accommodate many widely studied behavioral biases, such as inattention, present bias, diagnostic expectations, and near-rationality (epsilon-mistakes). This channel helps explain the empirical puzzle on high-liquidity consumers' high MPCs and can be significant. The same channel can also help explain other puzzles in intertemporal choices, such as violations of the fungibility principle, excess discounting of future income, and large risk aversion. Methodologically, I develop a general approach to study predictions of sophistication (i.e., the anticipation of future mistakes) independent of the underlying behavioral biases.

Chen Lian
Department of Economics
University of California, Berkeley
Evans Hall
Berkeley, CA 94720
and NBER
chen_lian@berkeley.edu

1 Introduction

Motivation. There is increasing consensus that behavioral biases play an important role in explaining consumption behavior. For example, recent evidence shows that consumers exhibit high marginal propensities to consume (MPCs) *away* from liquidity constraints (Parker, 2017; Kueng, 2018; Olafsson and Pagel, 2018; Fagereng, Holm and Natvik, 2019; McDowall, 2020).[1] It is hard for the canonical liquidity-constraints-based models (Carroll, 1997; Gourinchas and Parker, 2002) to explain this evidence and it points toward behavioral explanations.

But macroeconomists face a challenge in how to systematically incorporate behavioral biases into the mainstream model. There are many potential behavioral biases in intertemporal consumption problems, such as inattention (Sims, 2003; Maćkowiak and Wiederholt, 2015; Gabaix, 2014, 2016; Mackowiak, Matejka and Wiederholt, 2021; Carroll et al., 2020), present bias (Angeletos et al., 2001; Laibson, Maxted and Moll 2020), mental accounting (Shefrin and Thaler, 1988; Thaler, 1990), diagnostic expectations (Bordalo et al., 2020, Bianchi, Ilut and Saijo 2021), and imperfect problem solving (Ilut and Valchev, 2020). It is unclear whether there are robust, consistent predictions on how behavioral biases impact MPCs. Addressing this challenge is important, since consumption behavior away from liquidity constraints plays a key role in determining the macroeconomic impact of monetary and fiscal policies (Auclert, 2019; Holm, Paul and Tischbirek, 2020).

Methodology. In this paper, I introduce a wedge-based approach to systematically study the impact of behavioral mistakes on MPCs. The wedge approach is popular in macroeconomics to study financial and labor market frictions (Chari, Kehoe and McGrattan 2007). Here, I instead use wedges to capture how behavioral consumption rules deviate from their optimal counterparts (Mullainathan, Schwartzstein and Congdon, 2012; Farhi and Gabaix, 2020).[2] This approach can nest many widely studied behavioral biases, such as inattention, hyperbolic discounting, diagnostic expectations, and near-rationality (ϵ-mistakes).

I then use this approach to study how anticipation of future mistakes, i.e., sophistication, impacts current MPCs. This sophistication channel is very important in other contexts, such as "doing it now or later" in O'Donoghue and Rabin (1999, 2001).[3] There is also ample empirical

[1]For example, Fagereng, Holm and Natvik (2019) study consumption responses to unexpected Norwegian lottery prizes, and find high MPCs even among liquid winners: their estimates of the MPC for the group with the highest liquid asset balance is much higher than the prediction of standard liquidity-constraints-based models. Kueng (2018) documents excess sensitivity of the consumption response to the Alaska Permanent Fund payments, and finds the excess sensitivity is largely driven by high-income households with substantial liquid assets.

[2]The behavioral literature has used the wedge-approach to study normative questions, such as optimal policies with behavioral agents (Mullainathan, Schwartzstein and Congdon, 2012; Farhi and Gabaix, 2020). The methodological innovation of this paper is to show how it can be applied to study an important positive question, i.e., the robust prediction of sophistication on behavior.

[3]My approach is different from O'Donoghue and Rabin (1999, 2001) in two ways. First, I do not take an exact

evidence that consumers have a nontrivial degree of sophistication (e.g., Allcott et al., 2020; Carrera et al., 2021; Le Yaouanq and Schwardmann, 2019).[4] But its implications for MPCs are not well understood. Moreover, the impact of sophistication is often studied in the context of a specific mistake, instead of more broadly, independent of the exact mistakes. The main goal of the paper is to show that anticipation of future consumption mistakes in response to saving changes (i.e. changes in asset balances) leads to higher current MPCs. This result holds no matter the behavioral cause of these mistakes. This result contrasts with the direct impact of current behavioral mistakes on current MPCs, e.g., how current inattention or current present bias impacts current MPCs. This direct channel can lead to either higher MPCs (e.g., hyperbolic discounting) or lower MPCs (e.g., inattention).

The main result. The key result of the paper is that future consumption mistakes in response to saving changes lead to higher current MPCs. With these future mistakes, the consumer is less willing to adjust her saving and more willing to adjust her current consumption. Hence she displays higher current MPCs. This result is true no matter whether future consumption mistakes take the form of over-reaction or under-reaction to saving changes. This result can also be easily extended to the case of partial sophistication, i.e., partial understanding of future mistakes.[5]

To explain the intuition, consider the response to a positive current income shock. If the consumer increases her saving, the additional saving will not be spent optimally, because she cannot perfectly smooth increases in her future consumption. As a result, the value of increasing saving relative to the value of increasing current consumption diminishes. The consumer is then more willing to increase her current consumption and exhibits a higher current MPC.[6]

It is important to clarify that my high MPC result does not come from the precautionary saving motive. This motive is about how the dispersion of the levels of future consumption across states or time decreases the *level* of current consumption (increases the level of saving), but not directly about *MPCs*.

To isolate my mechanism, I first establish my high-MPC result in the case of quadratic utility, which a fortiori shuts down the precautionary saving motive. My high-MPC result extends to arbitrary concave utility, under an additional condition: there are no mistakes in the absence of shocks, i.e., mistakes occur only in the response of future consumption to saving changes. Many

stand on where future mistakes come from, while they focus on the impact of future present bias. Second, I focus on continuous decisions, e.g., standard intertemporal consumption and saving problems, while they focus on discrete choices.

[4] For example, in the context of hyperbolic discounting, Allcott et al. (2020) find that the degree of sophistication is close to 1.

[5] An interesting comparative statics result is that current MPCs increase with the degree of sophistication.

[6] By the same token, for a negative current income shock, the value of decreasing saving is extra negative, again because her future selves cannot perfectly smooth their consumption decreases in response to the saving decrease. The consumer is then more willing to decrease her current consumption and again exhibits a higher current MPC.

popular behavioral foundations satisfy this condition. For example, in models of beliefs-driven behavioral biases such as inattention and diagnostic expectations, belief mistakes only happen when the underlying fundamental deviates from the pre-shock default value (Sims, 2003; Maćkowiak and Wiederholt, 2015; Bianchi, Ilut and Saijo, 2021). As another example, for present bias agents with access to a commitment device (Laibson 1997; Angeletos et al. 2001), they can achieve optimal consumption through the commitment device in absence of the shock, but not in response to it. Finally, I provide additional results regarding how future mistakes in response to saving changes can still lead to higher current MPCs even when the previous condition does not hold.

Applications. I conduct a wedge-based calibration exercise to gauge the magnitudes of future mistakes' impact on current MPCs, accommodating different possible behavioral foundations. Using the consumption impulse responses to unexpected temporary incomes shocks (lottery prizes) from Fagereng, Holm and Natvik (2019) and Auclert, Rognlie and Straub (2018), I calculate the relevant wedges capturing future consumption mistakes in response to saving changes. I show the impact of these future mistakes can be quite sizable, by itself doubling the magnitude of the current MPC. Alternative calibration strategies based on specific behavioral foundations are also considered.

The key intermediate step to prove the high-MPC result is to establish the excess concavity of the continuation value function: mistakes in response to saving changes mean that saving changes are extra costly. The same excess concavity can help explain other well-known puzzles in intertemporal decisions. For example, future mistakes in response to saving changes lead to higher risk aversion and help explain the equity premium puzzle. As another example, these future mistakes can also explain the violation of the fungibility principle (Shefrin and Thaler, 1988), i.e., the prediction that all components of permanent income matter equally through the total present value. In particular, future mistakes in response to saving changes lead to a muted current consumption response to news of future income. This prediction captures Thaler (1990)'s observation about excess discounting of future income and is consistent with the limited "announcement" effect documented by Stephens and Unayama (2011), Parker (2017), and Kueng (2018).

My result also accommodates an alternative intra-household interpretation. The unitary model of household spending has long been rejected and it has been widely documented that the wife and husband exhibit different consumption behavior (Thomas, 1990; Browning et al., 1994; Anderson and Baland, 2002; Duflo, 2003; Duflo and Udry, 2004; Ashraf, 2009). In the intertemporal setting, there is strong evidence against the assumption that the wife and husband can commit their consumption and saving behavior to achieve an ex-ante collective Pareto efficient outcome (Mazzocco, 2007; Lise and Yamada, 2019). Instead, household consumption behavior fluctuates over time, depending on which spouse has a temporarily higher decision weight. From the lens of my model, this means that future consumption (e.g., determined by the husband) may deviate from what the current consumer (e.g., the wife) deems optimal. She then displays a higher MPC because, from her perspective, future consumption will respond inefficiently to saving changes.

Related literature. The evidence of high MPCs away from liquidity constraints (e.g., Fagereng, Holm and Natvik, 2019) is not easily explained by the traditional liquidity-constraint-based models and points towards behavioral forces. Compared to the existing behavioral literature on intertemporal consumption (e.g. Maćkowiak and Wiederholt, 2015; Matejka, 2016; Gabaix, 2016; Mackowiak, Matejka and Wiederholt, 2021; Carroll et al., 2020), the key difference is that this paper does not focus on a specific behavioral bias. I instead try to establish predictions independent of the specific behavioral biases.

The most related papers are Ilut and Valchev (2020) and Bianchi, Ilut and Saijo (2021). They also try to develop behavioral explanations of high-liquidity consumers' high MPCs. Ilut and Valchev (2020)'s theory is based on the consumer's imperfect problem solving. The high-MPC result there comes from the consumer's difficulty in calculating her optimal consumption rule. Also, Ilut and Valchev (2020) focus on the case of full naivety and does not study the impact of future mistakes on current consumption. Bianchi, Ilut and Saijo (2021) instead generate high MPCs from diagnostic expectations. In fact, Bianchi, Ilut and Saijo (2021) show that MPCs under sophistication are higher than MPCs under naivety. Through the lens of my paper, this result arises because diagnostic expectations lead to inefficient responses to saving changes in their model.

Mullainathan (2002) and Azeredo da Silveira and Woodford (2019) generate high MPCs because consumers' expectation of future income over-reacts to changes in current income. On the other hand, the channel emphasized in my paper can lead to high MPCs even if consumers form rational expectations about their current and future incomes. Thakral and To (2020) show that consumers' MPCs out of unanticipated income shocks may be higher than those out of anticipated income shocks in an environment with anticipatory utility.[7]

A recent literature studies the implications of sophistication, i.e., anticipation of future mistakes, in intertemporal consumption decisions (e.g., Allcott et al., 2020; Maxted, 2020). This literature focuses on the present bias only. I instead show that the implications of sophistication do not necessarily depend on specific mistakes. For this purpose, I develop a wedge-based approach to capture the impact of future mistakes. The wedge approach is widely used to study financial and labor market frictions in macroeconomics (Chari, Kehoe and McGrattan, 2007). In behavioral economics, it has been used to study normative questions, such as optimal policies

[7] Kőszegi and Rabin (2009) study a two-period consumption and saving model in which the consumer is loss averse to changes in her beliefs about her present and future consumption. The model generates high MPCs out of positive income shocks, but low MPCs out of negative income shocks.

with behavioral agents (Mullainathan, Schwartzstein and Congdon, 2012; Baicker, Mullainathan and Schwartzstein, 2015; Bernheim and Taubinsky, 2018; Farhi and Gabaix, 2020). The methodological innovation of this paper is to show how the wedge approach can be applied to study an important positive question, i.e., the robust prediction of sophistication on behavior.

Holm (2018) and Carroll, Holm and Kimball (2021) study how future liquidity constraints impact the concavity of consumption and value functions and the implications for consumer behavior. I instead focus on how future behavioral biases impact the shape of consumption and value functions.

2 A Simple Example

I will start with the simplest example of how future mistakes can lead to high current MPCs. The consumer lives for three periods, $t \in \{0, 1, 2\}$. Her experienced utility is given by

$$u(c_0) + u(c_1) + u(c_2), \tag{1}$$

where $u(\cdot) : \mathbb{R} \to \mathbb{R}$ is a strictly concave, increasing, utility function, and the discount factor is set to be 1 for simplicity. For the illustration purpose, I let $u(\cdot)$ be quadratic so the consumption rule is linear in the simple example here. The result can be generalized to the case with general concave utilities, which is the focus of Section 5.

The consumer can save and borrow through a risk-free asset with a gross interest rate $R = 1$. To isolate the friction of interest, she is not subject to borrowing constraints.

The question of interest is how consumption c_0 responds to the $t = 0$ income shock Δ. That is, the current MPC. For illustration purposes, in this section, the shock Δ is the only source of the consumer's income. Without the shock, income in each $t \in \{0, 1, 2\}$ is normalized to zero and the initial wealth is also normalized to zero. Together, her intertemporal budget is given by

$$w_1 = R(\Delta - c_0) \quad \text{and} \quad w_2 = R(w_1 - c_1), \tag{2}$$

where w_t is the consumer's wealth/saving at the start of period t. Without the shock ($\Delta = 0$), the optimal consumption at each period $t \in \{0, 1, 2\}$ is simply given by $\bar{c}_t = 0$.

Now let us turn to the consumer's consumption policy. At period $t = 2$, the consumer consumes out of her remaining saving,[8]

$$c_2(w_2) = w_2. \tag{3}$$

[8] Note that here c_2 can be negative. This makes sure that the problem is always well defined.

At period $t = 1$, the consumer's consumption rule is given exogenously by

$$c_1(w_1) = \frac{1}{2}(1 - \lambda_1) w_1. \tag{4}$$

Compared to the frictionless consumption rule $c_1^{\text{Frictionless}}(w_1) = \frac{1}{2} w_1$, λ_1 captures the mistake in response to changes in saving w_1.[9] When $\lambda_1 > 0$, c_1 under-reacts to w_1. When $\lambda_1 < 0$, c_1 over-reacts to w_1. As illustrated shortly, this is the type of future mistakes that leads to a higher MPC at $t = 0$.

I then study how the future mistake λ_1 impacts the current MPC at $t = 0$. $c_0^{\text{Deliberate}}(\Delta)$ captures self $t = 0$'s optimal consumption taking her future consumption rules (3) and (4) as given:

$$c_0^{\text{Deliberate}}(\Delta) = \arg\max_{c_0} u(c_0) + u(c_1(w_1)) + u(c_2(w_2)) \tag{5}$$

subject to the budget (2). $c_0^{\text{Deliberate}}(\Delta)$ isolates the impact of future mistakes λ_1 since it is the consumption that the consumer would choose at $t = 0$ if she were not subject to any current mistake but took her future mistakes as given. I hence term it "deliberate consumption." The current MPC is then given by $\phi_0^{\text{Deliberate}} \equiv \frac{\partial c_0^{\text{Deliberate}}(\Delta)}{\partial \Delta}$.

To establish the high-MPC result and better understand the intuition, I write (5) in a recursive form. Self 0 trades off between the utility of current consumption and the continuation value of saving:

$$c_0^{\text{Deliberate}}(\Delta) = \arg\max_{c_0} u(c_0) + V_1(w_1),$$

where $w_1 = R(\Delta - c_0)$ as in (2) and $V_1(w_1)$ captures the continuation value function, defined based on future consumption rules in (3) and (4):

$$\begin{aligned} V_1(w_1) &\equiv u(c_1(w_1)) + u(c_2(w_1 - c_1(w_1))) \\ &= u\left(\frac{1}{2}(1 - \lambda_1) w_1\right) + u\left(\frac{1}{2}(1 + \lambda_1) w_1\right). \end{aligned} \tag{6}$$

I can then establish the main result.

Proposition 1. *1.* ***Excess concavity of the continuation value:*** *The concavity of the continuation value function* $\left|V_1''\right| = \frac{1}{2}\left|u''\right|(1 + \lambda_1^2)$ *strictly increases with the future mistake* $|\lambda_1|$.

[9] Here, by "mistakes," I mean deviations from the optimal decision rule derived based on the experienced utility in (1), in the language of Kahneman, Wakker and Sarin (1997). More generally, mistakes matter for current consumption because the current self anticipates that her future selves will deviate from what she deems optimal.

2. ***High current MPCs:*** *The current MPC* $\phi_0^{Deliberate} = \frac{\frac{1}{2}\left(1+\lambda_1^2\right)}{1+\frac{1}{2}\left(1+\lambda_1^2\right)}$ *strictly increases with the future mistake* $|\lambda_1|$ *.*[10]

Proposition 1 shows that the future consumption mistake in response to saving changes (a larger $|\lambda_1|$) leads to a higher current MPC (a higher $\phi_0^{\text{Deliberate}}$). When future consumption responds inefficiently to saving changes, the consumer is less willing to adjust her saving. In response to changes in current income, she is then more willing to adjust her current consumption and displays a higher MPC. The high MPC result holds regardless of whether the future consumption mistake takes the form of under-reaction ($\lambda_1 > 0$) or over-reaction ($\lambda_1 < 0$) and regardless of the exact behavioral causes of the future mistake λ_1.

To better understand the high MPC result, note that the value of changing saving w_1 by δ is[11]

$$V_1(\delta) - V_1(0) \approx u'(0)\,\delta - \frac{1}{2}\left|V_1''\right| \cdot \delta^2, \tag{7}$$

which decreases with the future mistake $|\lambda_1|$ for any $\delta \neq 0$, because of the excess concavity in $\left|V_1''\right|$. Intuitively, because of future mistakes in response to saving changes, the consumer cannot perfectly smooth her future consumption responses to saving changes. The value of changing saving is then decreased (for both an increase in saving $\delta > 0$ and a decrease in saving $\delta < 0$).

On the other hand, the value of changing current consumption c_0 by δ is

$$u(\delta) - u(0) = u'(0)\,\delta - \frac{1}{2}\left|u''\right|\delta^2, \tag{8}$$

independent of the future mistake $|\lambda_1|$.

(7) and (8) together mean that the future mistake diminishes the value of changing saving relative to the value of changing current consumption. The consumer is then more willing to change her current consumption and exhibits a higher current MPC.

For example, consider a positive current income shock $\Delta > 0$. The value of increasing saving relative to the value of increasing current consumption diminishes because future selves cannot spend the additional saving optimally. The consumer is then more willing to increase her current consumption and exhibits a higher current MPC. By the same token, for a negative current income shock $\Delta < 0$, the value of decreasing saving in (7) is extra negative. The consumer is more willing to decrease her current consumption and again exhibits a higher current MPC.

[10] $\phi_0^{\text{Deliberate}}$ does not depend on Δ because $c_0^{\text{Deliberate}}(\Delta)$ is linear. V_1'' does not depend on w_1 because $V_1(w_1)$ is quadratic.

[11] Without the shock ($\Delta = 0$), the consumption c_0 and the saving w_1 are simply given by 0. That is why the baseline values in (7) and (8) are $V_1(0)$ and $u(0)$

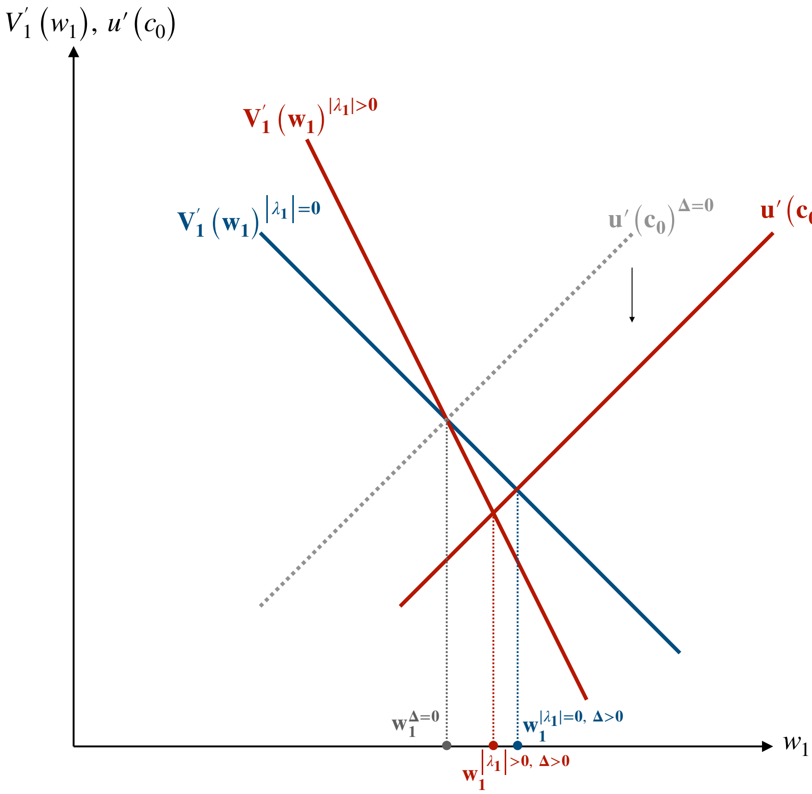

Figure 1: A Simple Example.

One can also visualize the high-MPC result based on the FOC:[12]

$$u'\left(c_0^{\text{Deliberate}}(\Delta)\right) = V_1'(w_1) \quad \text{with} \quad w_1 = R\left(\Delta - c_0^{\text{Deliberate}}(\Delta)\right). \tag{9}$$

Figure 1 plots the intersection between $u'(c_0)$ and $V_1'(w_1)$, as in (9). Because the future mistake ($|\lambda_1| > 0$) leads to the excess concavity of $V_1(w_1)$, the marginal value of saving $V_1'(w_1)$ becomes steeper with $|\lambda_1| > 0$. Now consider a positive shock to Δ as an example, which moves $u'(c_0) = u'(R(\Delta - c_0))$ downwards. One can see that, because $V_1'(w_1)$ is steeper with mistakes ($|\lambda_1| > 0$), the consumer is less willing to adjust her saving: $w_1 = R\left(\Delta - c_0^{\text{Deliberate}}(\Delta)\right)$ moves less with Δ. Equivalently, the consumer is more willing to adjust her consumption and displays a higher MPC: $c_0^{\text{Deliberate}}(\Delta) = \Delta - \frac{w_1}{R}$ moves more with Δ.

The key for the higher current MPC $\phi_0^{\text{Deliberate}}$ is that mistakes in future consumption's response to saving changes. To see this more clearly, we can extend the consumption rule in (4) to

$$c_1(w_1) = \frac{1}{2}(1 - \lambda_1) w_1 - \bar{\lambda}_1, \tag{10}$$

[12]From the FOC in (9), we can see that the current *MPC* $\phi_0^{\text{Deliberate}} \equiv \frac{\partial c_0^{\text{Deliberate}}(\Delta)}{\partial \Delta}$ is connected to the *second* derivative of the continuation value, i.e., the concavity: $\phi_0^{\text{Deliberate}} = \frac{V_1''}{u'' + V_1''}$. This is another way to see how the excess concavity of the continuation value leads to a higher current MPC.

which now allows two types of mistakes compared to the frictionless consumption rule $c_1^{\text{Frictionless}}(w_1) = \frac{1}{2}w_1$. First, same as in (4), λ_1 captures the mistake in response to changes in saving w_1. Second, (10) also allows the mistake in the overall consumption level in the absence of the shock ($\Delta = 0$). Specifically, $\bar{\lambda}_1$ captures how much the pre-shock ($\Delta = 0$) consumption level deviates from its frictionless level (0). When $\bar{\lambda}_1 > 0$, the consumer under-consumes at $t = 1$. When $\bar{\lambda}_1 < 0$, the consumer over-consumes at $t = 1$. In the quadratic environment here, $\phi_0^{\text{Deliberate}}$ is solely a function of the mistake in response to saving changes, λ_1, but is independent of the mistake in the pre-shock consumption level, $\bar{\lambda}_1$.[13] Intuitively, the MPC is about how the consumer responds to the income shock Δ, so it is directly connected to how future consumption responds to saving changes, λ_1, instead of its overall level, $\bar{\lambda}_1$.

It is important to clarify that the high MPC result in Proposition 1 does not come from the precautionary saving motive. This can be seen from two angles. First, there is simply no precautionary saving motive in Proposition 1, because the quadratic utility here a fortiori shuts down the precautionary saving motive. Second, as further explained in Section 5, the precautionary saving motive is about how the dispersion of the levels of future consumption across states or time[14] decreases the *level* of current consumption (increases the level of saving), but not directly about *MPCs*.

3 Set up

This section introduces a standard intertemporal consumption and saving problem. Then, I decompose the impact of behavioral biases on consumption rules into two parts: the effect of the current behavioral bias on current consumption; and the impact of the anticipation of future mistakes, i.e., sophistication in the language of O'Donoghue and Rabin (1999, 2001).

Utility and budget. I first introduce a canonical, single-agent, intertemporal consumption problem. The consumer's experienced utility is given by

$$U_0 \equiv \sum_{t=0}^{T-1} \delta^t u(c_t) + \delta^T v(a_T + y_T), \tag{11}$$

where c_t is her consumption at period $t \in \{0, 1, \ldots, T-1\}$, δ is her discount factor, $u(\cdot)$ captures the utility from consumption, and $v(\cdot) : \mathbb{R} \to \mathbb{R}$ captures the utility from retirement or bequests. Both $u(\cdot)$ and $v(\cdot)$ are strictly increasing and strictly concave.

[13] See the proof of Proposition 1 for details.

[14] The dispersion of the levels of future consumption across states or time can come from uncertainty, liquidity constraints, or mistake in the overall consumption level in the absence of the shock ($\bar{\lambda}_1$ in (10)). See Section 5 for a detailed discussion.

The consumer can save and borrow through a risk-free asset and is subject to the budget constraints

$$a_{t+1} = R\left(a_t + y_t - c_t\right) \quad \forall t \in \{0, \cdots, T-1\}, \tag{12}$$

where y_t is her exogenous income at period t, a_t is her wealth (i.e. saving/borrowing) at the start of period t, and R is the gross interest rate on the risk-free asset. The consumer can save and borrow through a risk-free asset.

To isolate the friction of interest, the consumer is not subject to any borrowing constraints. Her budget constraint (12) can then be rewritten as

$$w_{t+1} = R\left(w_t - c_t\right) \quad \forall t \in \{0, \cdots, T-1\}, \tag{13}$$

where $w_t = a_t + y_t + \sum_{k=1}^{T-t} R^{-k} y_{t+k}$ is her total wealth at period t, including her saving and the present value of her current and future income.

I study responses to an income shock Δ at $t = 0 : y_0 = \bar{y}_0 \to y_0 = \bar{y}_0 + \Delta$, or equivalently $w_0 = \bar{w}_0 \to w_0 = \bar{w}_0 + \Delta$, where bar over a variable captures its pre-shock value ($\Delta = 0$). For illustration purposes, I follow Chetty and Szeidl (2007) and let Δ be the only source of income uncertainty in the main analysis. Cases with gradual resolution of income uncertainty will be studied in Sections 4 and 5 below.

I use the widely adapted "multiple-selves" language as in Piccione and Rubinstein (1997) and Harris and Laibson (2001). That is, self $t \in \{0, \cdots, T-1\}$ is in charge of consumption and saving decisions at period t. In particular, I use $c_t(w_t)$ to denote each self t's *actual* consumption rule, subject to behavioral mistakes.[15]

Isolating the impact of future mistakes. Behavioral biases can impact self t's actual consumption rule $c_t(w_t)$ through two distinct channels. First, self t's own behavioral bias (parameterized by Λ_t) can directly impact her current consumption, e.g., the impact of current inattention or current present bias on current consumption. Second, anticipation of future selves' mistakes $\{\Lambda_{t+k}\}_{k=1}^{T-t-1}$, i.e., sophistication in the language of O'Donoghue and Rabin (1999, 2001), can also impact current consumption.

To isolate the second channel, I introduce the *deliberate* consumption rule $c_t^{\text{Deliberate}}(w_t)$, i.e., the consumption that self t would have chosen if she were not subject to any current behavioral mistake but took future selves' mistakes in their actual consumption rules as given.

[15] Embedded in this consumption rule is fungibility, i.e., the actual consumption remains a function of the total wealth w_t only. This helps isolate the channel of interest. And the main high-MPC result can be easily extended to the case where the fungibility principle is violated. See the discussion at the end of Section 4 below.

Definition 1. *For each* $t \in \{0, \cdots, T-1\}$, *self* t*'s deliberate consumption rule optimizes the consumer's utility in (11), taking future selves' actual consumption rules* $\{c_{t+k}(w_{t+k})\}_{k=1}^{T-k-1}$ *as given:*

$$c_t^{Deliberate}(w_t) \equiv \arg\max_{c_t} u(c_t) + \sum_{k=1}^{T-t-1} \delta^k u(c_{t+k}(w_{t+k})) + \delta^{T-t} v(w_T), \tag{14}$$

subject to the budget in (13).

With this definition, the following decomposition illustrates how the above two behavioral channels impact self t's actual consumption rule $c_t(w_t)$:

$$c_t(w_t) = \mathcal{S}\left(c_t^{\text{Deliberate}}(w_t), \Lambda_t\right). \tag{15}$$

Self t's own behavioral bias (parameterized by Λ_t) impacts actual consumption by letting it deviate from deliberate consumption $c_t^{\text{Deliberate}}(w_t)$, captured by the function $\mathcal{S}$.[16] On the other hand, the anticipation of future selves' mistakes $\{\Lambda_{t+k}\}_{k=1}^{T-t-1}$ impacts current actual consumption through the deliberate consumption $c_t^{\text{Deliberate}}(w_t)$.

The main theme for the rest of the paper is that, once I isolate the impact of future consumption mistakes through the deliberate consumption $c_t^{\text{Deliberate}}(w_t)$, I can show that future mistakes in response to saving changes robustly lead to high current MPCs, no matter the micro-foundations of these mistakes.

The deliberate consumption in (14) is defined based on correct knowledge of future actual consumptions rules (and future mistakes). This choice significantly simplifies the notation without changing the essence. The analysis can accommodate a more general interpretation if we define deliberate consumption (14) based on perceived future consumption rules. With this re-interpretation, the analysis here easily accommodates the case of partial sophistication in O'Donoghue and Rabin (1999, 2001). See Corollaries 3 and 4 for details.

A recursive formulation. 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 $t \in \{0, \cdots, T-1\}$,

$$V_t(w_t) \equiv u(c_t(w_t)) + \sum_{k=1}^{T-t-1} \delta^k u(c_{t+k}(w_{t+k})) + \delta^{T-t} v(w_T), \tag{16}$$

subject to the budget in (13). For the last period T, we have $V_T(w_T) = v(w_T)$.

Based on (16), I can express the deliberate consumption rule in (14) recursively. This recursive formulation paves the way for the analysis in the rest of the paper.

[16] We have $\mathcal{S}\left(c_t^{\text{Deliberate}}(w_t), 0\right) = c_t^{\text{Deliberate}}(w_t)$. That is, when the current self's is not subject to any behavioral bias ($\Lambda_t = 0$), she will choose the deliberate consumption rule as in (14).

Proposition 2. *For $t \in \{0, \cdots, T-1\}$, each self t's deliberate consumption rule defined in (14) satisfies*

$$c_t^{Deliberate}(w_t) = \max_{c_t} \; u(c_t) + \delta V_{t+1}(R(w_t - c_t)). \tag{17}$$

Moreover, for $t \in \{0, \cdots, T-1\}$, the value function $V_t(w_t)$ defined in (16) satisfies

$$V_t(w_t) = u(c_t(w_t)) + \delta V_{t+1}(R(w_t - c_t(w_t))), \tag{18}$$

where the actual consumption rule $c_t(w_t)$ is given by (15).

Finally, if consumption rules and value functions $\left\{c_t^{Deliberate}(w_t), c_t(w_t)\right\}_{t=0}^{T-1}$ and $\{V_t(w_t)\}_{t=0}^{T}$ satisfy (15), (17), (18), and the boundary condition $V_T(w_T) = v(w_T)$, they coincide with the corresponding objects defined sequentially in (14)–(16).

A note on budget constraints. It is worth noting that the final wealth $w_T = a_T + y_T$ is allowed to be negative, since the utility from retirement or bequests $v(\cdot)$ is defined on the entirety of $\mathbb{R}$. This guarantees that, even with consumption mistakes, the budget in (13) is always satisfied and the intrapersonal problem is always well defined. The final period does not play a special role: below, I show that the consumer's deliberate and actual consumption rules converge to simple limits when $T \to +\infty$.

4 The Benchmark Result

I will start the analysis with a benchmark case where u and v in (11) are arbitrary quadratic, strictly concave functions. In this case, consumptions functions are linear. Future consumption mistakes in response to saving changes always lead to higher current MPCs. This case also helps clarify that my main results do not come from the precautionary saving motive.

Mistakes in future consumption. Actual future consumption rules can take arbitrarily linear forms, which can be written as:[17]

$$c_{t+k}(w_{t+k}) = \phi_{t+k}(w_{t+k} - \bar{w}_{t+k}) + \bar{c}_{t+k}, \tag{19}$$

where ϕ_{t+k} captures how future consumption responds to changes in w_{t+k} and $\bar{c}_{t+k}$ and $\bar{w}_{t+k}$ capture the pre-shock ($\Delta = 0$) outcome, satisfying $\bar{c}_s = c_s(\bar{w}_s)$ and $\bar{w}_{s+1} = R(\bar{w}_s - \bar{c}_s)$ for $s \in \{0, 1, \cdots, T-1\}$.

There are two types of future mistakes embedded in (19) compared to its deliberate counterpart $c_{t+k}^{\text{Deliberate}}(w_{t+k})$ defined in Definition 1. The latter captures the consumption that self $t+k$ would

[17] (19) is defined for all $t + k \in \{0, \cdots, T-1\}$.

have chosen if she were not subject to any current mistake but took future mistakes as given:

$$c_{t+k}^{\text{Deliberate}}(w_{t+k}) = \phi_{t+k}^{\text{Deliberate}}(w_{t+k} - \bar{w}_{t+k}) + \bar{c}_{t+k}^{\text{Deliberate}}, \tag{20}$$

where $\phi_{t+k}^{\text{Deliberate}}$ captures how $t+k$ consumption should have responded to changes in w_{t+k} and $\bar{c}_{t+k}^{\text{Deliberate}} \equiv \bar{c}_{t+k}^{\text{Deliberate}}(\bar{w}_{t+k})$ captures what $t+k$ consumption should have been at the pre-shock wealth level $\bar{w}_{t+k}$.

First and most important for our purpose, future consumption may respond inefficiently to changes in wealth/saving w_{t+k}. That is, ϕ_{t+k} in (19) may deviate from its deliberate counterpart $\phi_{t+k}^{\text{Deliberate}}$ in (20). As in Section 2, I use λ_{t+k} to capture this mistake/behavioral wedge,

$$\phi_{t+k} = (1 - \lambda_{t+k})\,\phi_{t+k}^{\text{Deliberate}}. \tag{21}$$

When $\lambda_{t+k} > 0$, self $t+k$'s consumption under-reacts to changes in w_{t+k}. When $\lambda_{t+k} < 0$, self $t+k$'s consumption over-reacts to changes in w_{t+k}. These types of future mistakes are the key to the higher current MPC. As illustrated in Section 6, most popular behavioral foundations, such as inattention, hyperbolic discounting, and diagnostic expectations, generate these types of mistakes.

Second, mistakes in future consumption may also involve mistakes in the overall consumption level. That is, $\bar{c}_{t+k}$ in (19) may deviate from its deliberate counterpart $\bar{c}_{t+k}^{\text{Deliberate}}$ in (20).

$$\bar{\lambda}_{t+k} = \bar{c}_{t+k}^{\text{Deliberate}} - \bar{c}_{t+k}. \tag{22}$$

When $\bar{\lambda}_{t+k} > 0$, the consumer under-consumes at $t = 1$, no matter the value of w_{t+k}. When $\bar{\lambda}_{t+k} < 0$, the consumer over-consumes at $t+k$.

Here, each self's mistake $\Lambda_{t+k} = \left(\lambda_{t+k}, \bar{\lambda}_{t+k}\right)$ is exogenous. But I will connect these mistakes to the exact underlying behavioral biases in Section 6.

The main result. The main result of this section is that future mistakes in response to saving changes $\{\lambda_{t+k}\}_{k=1}^{T-t-1}$ increase the current MPC. Specifically, based on (19) and Definition 1, I can calculate current self t's deliberate consumption $c_t^{\text{Deliberate}}(w_t)$ and her deliberate MPC $\phi_t^{\text{Deliberate}} \equiv \frac{\partial c_t^{\text{Deliberate}}(w_t)}{\partial w_t}$.

Proposition 3. *For $t \in \{0, \cdots, T-2\}$, each self t's current MPC $\phi_t^{Deliberate}$ increases with each future self's mistake in response to saving changes $\{|\lambda_{t+k}|\}_{k=1}^{T-t-1}$ but is independent of $\left\{\left|\bar{\lambda}_{t+k}\right|\right\}_{k=1}^{T-t-1}$.*

Proposition 3 shows that future consumption mistakes in response to saving changes increase the current MPC. This means that $\phi_t^{\text{Deliberate}} \geq \phi_t^{\text{Frictionless}}$, where $\phi_t^{\text{Frictionless}}$ is the frictionless MPC of actual consumption when all λs are equal to 0. Moreover, regardless of whether future mistakes take the form of under-reaction ($\lambda_{t+k} > 0$) or over-reaction ($\lambda_{t+k} < 0$), these mistakes always increase the current MPC. On the other hand, mistakes in the overall consumption level $\{\bar{\lambda}_{t+k}\}_{k=1}^{T-t-1}$ do not matter for current MPCs.

Excess concavity of the continuation value function. Similar to Section 2, a simple way to understand Proposition 3 is through the continuation value function defined in (16). Specifically, let me use Γ_{t+1} to capture the concavity of the continuation value function $V_{t+1}(w_{t+1})$. That is, for $t \in \{0, \cdots, T-1\}$,

$$\Gamma_{t+1} \equiv \frac{\partial^2 V_{t+1}(w_{t+1})}{\partial w_{t+1}^2} / u'' > 0, \tag{23}$$

where a larger Γ_{t+1} means a more concave value function $V_{t+1}(w_{t+1})$.[18,19]

Lemma 1. *For each $t \in \{0, \cdots, T-2\}$, the concavity of the continuation value Γ_{t+1} increases with each future self's mistake in response to saving changes $\{|\lambda_{t+k}|\}_{k=1}^{T-t-1}$ but is independent of $\{|\bar{\lambda}_{t+k}|\}_{k=1}^{T-t-1}$.*

Because of future mistakes in response to saving changes, future selves cannot perfectly smooth their consumption responses to saving changes. The future consumption responses will be more concentrated in some periods. The value of changing saving is then decreased (for both an increase in saving and a decrease in saving). This leads to the excess concavity of the continuation value.

Importantly, the concavity of the continuation value function depends on the size of future consumption mistakes $|\lambda_{t+k}|$, but does not depend on whether mistakes take the form of under-reaction ($\lambda_{t+k} > 0$) or over-reaction ($\lambda_{t+k} < 0$). In this sense, future mistakes in response to changes in saving always increase the concavity of the continuation value function.

High current MPCs. I am now ready to explain the main Proposition 3. From the recursive formulation in (17), we know

$$c_t^{\text{Deliberate}}(w_t) = u(c_t) + \delta V_{t+1}(R(w_t - c_t)). \tag{24}$$

Similar to (7), the value of changing saving w_1 by δ is

$$V_{t+1}(w_{t+1}) - V_{t+1}(\bar{w}_{t+1}) = V_{t+1}'(\bar{w}_{t+1})\delta - \frac{1}{2}\left|u''\right| \cdot \Gamma_{t+1} \cdot \delta^2. \tag{25}$$

From the excess concavity in Γ_{t+1}, future mistakes $\{|\lambda_{t+k}|\}_{k=1}^{T-t-1}$ diminish the value of changing saving for any $\delta \neq 0$. On the other hand, the value of changing current consumption does not

[18] $u'' < 0$, a constant, is the second derivative of the utility function. Moreover, the definition in (23) can be extended to $\Gamma_0 \equiv \frac{\partial^2 V_0(w_0)}{\partial w_0^2}/u''$.

[19] Even with future consumption mistakes, the continuation value function here $V_{t+1}(w_{t+1})$ is always concave. This feature is guaranteed because my setup does not feature borrowing constraints. The pathological non-concave value function case arises when there is a kink in consumption rules due to borrowing constraints (e.g. Laibson, 1997 and Harris and Laibson, 2001).

depend on these future mistakes. As a result, the current self is then more willing to adjust her current consumption instead of her saving. She hence displays a higher MPC.[20]

Consider a positive shock to w_t. From (25), the value of increasing saving relative to the value of increasing current consumption diminishes. The consumer is then more willing to increase her current consumption and exhibits a higher current MPC. By the same token, for a negative shock to w_t, the value of decreasing saving is extra negative. The consumer is more willing to decrease her current consumption and again exhibits a higher current MPC.

One can also derive the high MPC result through the FOC.

$$u'\left(c_t^{\text{Deliberate}}\left(w_t\right)\right) = V_{t+1}'\left(R\left(w_t - c_t^{\text{Deliberate}}\left(w_t\right)\right)\right). \tag{26}$$

Taking a partial derivative with respect to w_t, the current *MPC* $\phi_t^{\text{Deliberate}} \equiv \frac{\partial c_t^{\text{Deliberate}}(w_t)}{\partial w_t}$ is linked with the second derivative of the continuation value, i.e., the concavity:

$$\phi_t^{\text{Deliberate}} = \frac{V_{t+1}''}{u'' + V_{t+1}''}.$$

The excess concavity of the continuation value in Lemma 1 then leads to a high current MPC.

In sum, Proposition 3 shows that, once we isolate the impact of future consumption mistakes on current MPCs, future mistakes in response to changes in saving always raise the current MPC, regardless of whether future selves over-react ($\lambda_{t+k} < 0$) or under-react ($\lambda_{t+k} > 0$). This result is in contrast with the impact of current behavioral biases (λ_t) on the current MPC, which can go either way.

On the other hand, mistakes in the overall consumption level $\left\{\bar{\lambda}_{t+k}\right\}_{k=1}^{T-t-1}$ do not matter for the current MPC. Intuitively, the MPC is about how the consumer responds to the income shock Δ, so it is directly connected to how future consumption responds to saving changes, $\{\lambda_{t+k}\}_{k=1}^{T-t-1}$, instead of its overall level, $\left\{\bar{\lambda}_{t+k}\right\}_{k=1}^{T-t-1}$.

The $T \to \infty$ limit. The deliberate MPC $\phi_t^{\text{Deliberate}}$ converges to simple limits when all future selves share the same friction $\lambda_{t+k} = \lambda$ and the consumer's horizon T goes to infinity.

Corollary 1. *Let $\delta R^2 < 1$ and $\lambda_{t+k} = \lambda$ with $|\lambda| < \left(\delta^{-1/2} R^{-1}\right)$ for all $k \geq 1$. We have, for $T \to +\infty$,*

$$\phi_t^{\textit{Deliberate}} \to \phi^{\textit{Deliberate}} = \frac{\delta R^2 - 1}{\delta R^2\left(1 - \lambda^2\right)}, \tag{27}$$

where the condition $\delta R^2 < 1$ and $|\lambda| < \left(\delta^{-1/2} R^{-1}\right)$ guarantees that the transversality condition

[20]The analogy of this result in price theory is that, in response to changes in wealth, the consumer is more willing to adjust the consumption of a good with a less concave utility function.

$\lim_{k\to+\infty}\delta^k u'(c_{t+k})=0$ *holds.*

When $\lambda\to\left(\delta^{-1/2}R^{-1}\right)^-$, the deliberate MPC $\phi^{\text{Deliberate}}$ achieves its upper bound,

$$\lim_{\lambda\to\left(\delta^{-1/2}R^{-1}\right)^-}\phi^{\text{Deliberate}}=1.$$

That is, when future selves' consumption mistakes are large enough, the current self t is so worried about her future selves' mistakes that she follows a simple rule of thumb: she consumes all changes in w_t.

4.1 Extensions

Gradual resolution of income uncertainty. Above, for illustration purposes, the only uncertainty is the income shock Δ, which is resolved in period 0. One may naturally wonder about the case with gradual resolution of income uncertainty. For the quadratic utility case studied here, following from the well-known certainty equivalence result, the high-MPC result in Proposition 3 can be easily recast with gradual resolution of uncertainty.

For clarity, I explicitly work with with different components of the budget constraint (12):

$$a_{t+1}=R\left(a_t+y_t-c_t\right),$$

where y_t is her income at period t and a_t is her saving/borrowing at the start of period t. I consider a gradual resolution of the income uncertainty, where the random income y_t is drawn i.i.d. across each period $t\in\{0,\cdots,T-1\}$ with a mean zero.

In this environment, it is easier to write the actual and deliberate consumption rule of each self $t\in\{0,\cdots,T-1\}$ as a function of cash on hand $x_t\equiv a_t+y_t$, which equals to the expected total wealth: $x_t=E_t\left[w_t\right]=E_t\left[a_t+y_t+\sum_{k=1}^{T-t}R^{-k}y_{t+k}\right]$. That is, c_t and $c_t^{\text{Deliberate}}$ are given by

$$c_t\left(x_t\right)=\phi_t x_t+\bar{c}_t\quad\text{and}\quad c_t^{\text{Deliberate}}\left(x_t\right)=\phi_t^{\text{Deliberate}}x_t+\bar{c}_t^{\text{Deliberate}},\tag{28}$$

and

$$x_{t+1}=R\left(x_t-c_t\right)+y_{t+1}.$$

I still use λ_t to capture how self t's actual MPC ϕ_t deviates from the deliberate MPC $\phi_t^{\text{Deliberate}}$:

$$\phi_t=\left(1-\lambda_t\right)\phi_t^{\text{Deliberate}}.$$

From future selves' actual consumption rules $\left\{c_{t+k}\left(x_{t+k}\right)\right\}_{k=1}^{T-1-t}$, one can calculate current self

t's deliberate consumption rule $c_t^{\text{Deliberate}}(x_t)$ and find her deliberate MPC $\phi_t^{\text{Deliberate}}$ as usual. Same as Proposition 3, future mistakes in response to saving changes increase the current MPC.

Corollary 2. *For $t \in \{0, \cdots, T-2\}$, $\phi_t^{Deliberate}$ shares the exact same formula as $\phi_t^{Deliberate}$ in Proposition 3.*

Partial sophistication. In essence, it is the anticipation of future mistakes that leads to higher current MPCs. In the main analysis, for notation simplicity, I define the deliberate consumption (14) based on correct anticipation of future actual consumption rules and future mistakes, i.e., full sophistication. But the high-MPC result can also be extended to the case of partial sophistication, i.e., partial understanding of future mistakes.

Specifically, the main analysis can accommodate a more general interpretation if I re-define deliberate consumption (14) based on perceived future consumption rules and perceived future mistakes. Let $\{\tilde{c}_{t,t+k}(w_{t+k})\}_{k=1}^{T-t-1}$ capture self t's perceived future consumption rules. We can redefine the deliberate consumption based on these perceived future consumption rules:

$$c_t^{\text{Deliberate}}(w_t) \equiv \arg\max_{c_t} u(c_t) + \sum_{k=1}^{T-t-1} \delta^k u(\tilde{c}_{t,t+k}(w_{t+k})) + \delta^{T-t} v(w_T), \tag{29}$$

subject to the budget $w_{t+k} = R(w_{t+k-1} - c_{t+k-1})$.

We can then re-state Propositions 3 as how perceived future mistakes in response to saving changes $\left\{\tilde{\lambda}_{t,t+k}\right\}_{k=1}^{T-t-1}$, defined based on $\{\tilde{c}_{t,t+k}(w_{t+k})\}_{k=1}^{T-t-1}$ similar to (21), increase the current MPC, $\phi_t^{\text{Deliberate}} \equiv \frac{\partial c_t^{\text{Deliberate}}(w_t)}{\partial w_t}$.[21]

Corollary 3. *Based on the definition in (29), Proposition 3 can be recast as $\phi_t^{Deliberate}$ increases with each perceived future mistake $\left\{\left|\tilde{\lambda}_{t,t+k}\right|\right\}_{k=1}^{T-t-1}$ for each for $t \in \{0, \cdots, T-2\}$ and $k \in \{0, \cdots, T-t-1\}$.*

In other words, the main analysis is exactly the same with this re-interpretation. One important example of how perceived future mistakes are determined is the case of partial sophistication in O'Donoghue and Rabin (1999, 2001). That is, the current self t has a partial understanding of future mistakes, and her perceived future mistake at $t+k$ are given by:

$$\tilde{\lambda}_{t,t+k} = s_t \lambda_{t+k}, \tag{30}$$

[21] Based on perceived future consumption rules $\{\tilde{c}_{t,t+k}(w_{t+k})\}_{k=1}^{T-t-1}$, we can define perceived future mistakes $\left\{\tilde{\lambda}_{t,t+k}\right\}_{k=1}^{T-t-1}$ similar to (21). We first find the consumption that would have been chosen if self $t+k$ were not subject to any behavioral mistake and takes future consumption rules as given by $\{\tilde{c}_{t,t+k+l}(w_{t+k+l})\}_{l=1}^{T-t-k-1}$: $c_{t,t+k}^{\text{Deliberate}}(w_{t+k}) \equiv \arg\max_{c_t} u(c_t) + \sum_{l=1}^{T-t-k-1} \delta^k u(\tilde{c}_{t,t+k+l}(w_{t+k+l})) + \delta^{T-t} v(w_T)$, subject to the budget. We can then define self t's perceived future mistake $\tilde{\lambda}_{t,t+k}$ at $t+k$ similar to (21).

where $s_t \in [0,1]$ captures the degree of self t's sophistication. From Corollary 3, 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. The second comparative statics prediction, formalized in the following Corollary, is empirically testable.

Corollary 4. *With (30), Corollary 3 can be recast as each self t's deliberate MPC* $\phi_t^{Deliberate}$ *increasing with the degree of sophistication* s_t.

The non-fungibility case and the importance of mistake in response to *saving* changes. In the main analysis above, I restrict consumption to be a function of the total wealth w_t only. In other words, I maintain the fungibility principle (Thaler, 1990): the consumer's saving and income do not matter separately; they only matter through the total wealth w_t based on the present value.

One may wonder what may happen in the non-fungibility case. That is, when consumption responds to saving and income differentially. I will refer the reader to Appendix C for a full analysis, since it will introduce many new notations. But let me preview the result here.

Specifically, to allow non-fungibility, I explicitly work with different components of the budget constraint (12):

$$a_{t+1} = R\left(a_t + y_t - c_t\right) \quad \forall t \in \{0, \cdots, T-1\}, \tag{31}$$

where y_t is her income at period t, a_t is her saving/borrowing at the start of period t, and R is the gross interest rate on the risk-free asset.[22]

I will then use λ^a_{t+k} to capture the mistakes of future consumption in response to changes in saving/borrowing and λ^y_{t+k} to capture the mistakes of future consumption in response to changes in income.[23] I show that the above high-MPC results remain true: as long as future consumption responds inefficiently to changes in saving, the current MPC is higher. That is, the deliberate MPC $\phi_t^{\text{Deliberate}}$ increases with each of $\left\{\left|\lambda^a_{t+k}\right|\right\}_{k=1}^{T-t-1}$. On the other hand, mistakes in future consumption in response to changes in income (λ^y_{t+k}s) do not matter for the current MPC. This also clarifies why I emphasize mistakes in future consumption's response to saving changes throughout the paper.

[22]The total wealth in (13) is then given by $w_t = a_t + y_t + \sum_{k=1}^{T-t} R^{-k} y_{t+k}$.

[23]In Appendix C, I study a fully general case where I also allow differential mistakes in response to different components of income: $\lambda^y_{t+k,0}$ to capture the mistakes of c_{t+k} in response to changes in current income y_{t+k} and $\lambda^y_{t+k,m-l}$ to capture the mistakes of future consumption c_{t+k} in response to changes in future income y_{t+m} for $m \in \{k+1, \cdots, T-t-1\}$.

5 General Concave Utilities and the Precautionary Saving Motive

In this Section, I relax a few key assumptions in the main analysis. I show when and how the main high-MPC result remains to hold with general concave utilities. Then, I explain why my high MPC result does not come from the precautionary saving motive.

5.1 The High-MPC Result with General Concave Utilities

Now I turn to analyze the problem set up in Section 3 with general concave utilities $u(\cdot)$ and $v(\cdot)$.[24] I first show the high-MPC result in Proposition 1 remains to be true, under an additional condition satisfied by many popular behavioral foundations: 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.

Specifically, similar to (21), I use λ_{t+k} to capture the mistakes in c_{t+k}'s response to changes in w_{t+k}, defined as

$$\phi_{t+k} = (1 - \lambda_{t+k}) \, \phi_{t+k}^{\text{Deliberate}}, \tag{32}$$

where $\phi_{t+k} \equiv \frac{\partial c_{t+k}(\bar{w}_{t+k})}{\partial w_{t+k}}$, $\phi_{t+k}^{\text{Deliberate}} \equiv \frac{\partial c_{t+k}^{\text{Deliberate}}(\bar{w}_{t+k})}{\partial w_{t+k}}$, and $\bar{c}_{t+k}$ and $\bar{w}_{t+k}$ capture the pre-shock ($\Delta = 0$) outcome as above, satisfying $\bar{c}_s = c_s(\bar{w}_s)$ and $\bar{w}_{s+1} = R(\bar{w}_s - \bar{c}_s)$, for $s \in \{0, 1, \cdots, T-1\}$. When $\lambda_{t+k} > 0$, self $t+k$'s consumption under-reacts to changes in w_{t+k}. When $\lambda_{t+k} < 0$, self $t+k$'s consumption over-reacts to changes in w_{t+k}.

On the other hand, I shut down mistakes in the overall future consumption level. That is, at the pre-shock ($\Delta = 0$) wealth level $\bar{w}_{t+k}$, self $t+k$ would choose the consumption that she should have chosen:[25]

$$\bar{c}_{t+k} = c_{t+k}(\bar{w}_{t+k}) = c_{t+k}^{\text{Deliberate}}(\bar{w}_{t+k}). \tag{33}$$

That is, $\bar{\lambda}_{t+k} \equiv c_{t+k}^{\text{Deliberate}}(\bar{w}_{t+k}) - \bar{c}_{t+k} = 0$. The impact of mistakes in the overall future consumption level will be discussed extensively in Section 5.2.

Based on future selves' actual consumption rules $\{c_{t+k}(w_{t+k})\}_{k=1}^{T-1-t}$, I can then calculate current self t's deliberate consumption $c_t^{\text{Deliberate}}(w_t)$ and her deliberate MPC $\phi_t^{\text{Deliberate}}$. I can now re-establish Proposition 3 in the case of general concave utility.

Proposition 4. *If (33) holds for all $k \in \{0, \cdots, T-t-1\}$,*

$$\phi_t^{Deliberate} \equiv \frac{\partial c_t^{Deliberate}(\bar{w}_t)}{\partial w_t}$$

[24] For technical reasons, I also assume u, v, and c_t are third-order continuously differentiable.

[25] (33) holds for for all $t+k \in \{0, \cdots, T-1\}$.

increases with the future mistake $|\lambda_{t+k}|$ *for each* $k \in \{1, \cdots, T-t-1\}$.

As a result, the above analysis about how future selves' mistakes in response to saving changes lead to higher current MPCs remain to hold in this more general setting. The intuition is exactly the same as in Proposition 3.

The case studied in Proposition 4 where mistakes in future consumption only take the form of mistakes in response to saving changes is economically meaningful. In fact, as studied in Section 6, many popular behavioral foundations satisfy this condition. For example, in models of beliefs-driven behavioral biases such as inattention and diagnostic expectations, belief mistakes often only happen when the underlying fundamental deviates from its pre-shock default value. As another example, for present bias agents with access to a commitment device (Laibson 1997; Angeletos et al. 2001), they can achieve optimal consumption through the commitment device for the pre-shock outcome in (33). But the commitment device cannot restrain their mistakes in response to the shock.

What happens if (33) is not satisfied and the consumer also exhibits future mistakes in the overall consumption level? Future mistakes in the overall future consumption level generate an additional channel: the precautionary saving motive. This is the focus of Section 5.2. But the precautionary saving motive is not directly about *MPCs*. It is about how the dispersion of the *levels* of future consumption across states or time, which can be driven by mistake in the overall consumption level ($\bar{\lambda}_{t+k}$s), decreases the *level* of current consumption (increases the level of saving).

The $T \to \infty$ limit. Similar to Corollary 1, the deliberate MPC $\phi_t^{\text{Deliberate}}$ converges to simple limits when all future selves share the same friction $\lambda_{t+k} = \lambda$ and the consumer's horizon T goes to infinity.

Corollary 5. *Consider the CRRA case with* $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$. *Let* $\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} > 1$ *and* $\lambda_{t+k} = \lambda$ *with* $|\lambda| < \left(\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}}\right)^{-\frac{1}{2}}$ *for all* $k \geq 1$. *We have, for* $T \to +\infty$,

$$\phi_t^{Deliberate} \to \phi^{Deliberate} = \frac{\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} - 1}{\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} \left(1-\lambda^2\right)}. \tag{34}$$

Despite the formula (34) here being slightly different from (27) for the quadratic case, in both cases, mistakes in future consumption raise the MPC by a factor of $\phi^{\text{Deliberate}}/\phi^{\text{Frictionless}} = \frac{1}{1-\lambda^2}$.

Gradual resolution of income uncertainty. For the general concave utility case studied here, things are more complicated with gradual resolution of income uncertainty and an analytical characterization seems impossible. In practice, however, the high-MPC result in Proposition

4 remains to hold as long as a condition akin to (33) holds: there are no mistakes in future consumptions when incomes are realized at their median levels.

I conduct a numerical exercise along this line in Figure 2. For clarity, similar to the quadratic case, I explicitly work with different components of the budget (31):

$$a_{t+1} = R\left(a_t + y_t - c_t\right),$$

where the random income $y_t \sim \log \mathcal{N}\left(0, \sigma^2\right)$ is drawn i.i.d. across each period $t \in \{0, \cdots, T-1\}$.[26] To illustrate the robustness of my result, I also introduce borrowing constraints: for all $t \in \{0, \cdots, T-1\}$,

$$a_{t+1} \geq \underline{a}.$$

Similar to the quadratic case in (28), the actual and deliberate consumption rule of each self $t \in \{0, \cdots, T-1\}$ can be written as a function of cash on hand $x_t = a_t + y_t$. But $c_t\left(x_t\right)$ and $c_t^{\text{Deliberate}}\left(x_t\right)$ are no longer linear.

Akin to (33), I let that actual consumptions coincide with their deliberate counterparts when the stochastic incomes are realized at their median levels. That is, for $t \in \{0, \cdots, T-1\}$,

$$c_t\left(\bar{x}_t\right) = c_t^{\text{Deliberate}}\left(\bar{x}_t\right),$$

where $\bar{x}_{t+1} = R\left(\bar{x}_t - c_t\left(\bar{x}_t\right) + 1\right)$, for $t \in \{0, 1, \cdots, T-1\}$.[27]

Similar to (32), actual consumptions respond inefficiently to saving changes

$$\phi_t = \left(1 - \lambda_t\right) \phi_t^{\text{Deliberate}}, \tag{35}$$

where $\phi_t \equiv \frac{\partial c_t(\bar{x}_t)}{\partial x_t}$ and $\phi_t^{\text{Deliberate}} \equiv \frac{\partial c_t^{\text{Deliberate}}(\bar{x}_t)}{\partial x_t}$ and λ_t captures self t's mistake. To extend (35) globally, the actual consumption of each self $t \in \{0, \cdots, T-1\}$ is given by

$$c_t\left(x_t\right) = \min\left\{\frac{\underline{a}}{R} + x_t, c_t^{\text{Deliberate}}\left(\left(1 - \lambda_t\right) x_t + \lambda_t \bar{x}_t\right)\right\}, \tag{36}$$

which makes sure the consumer will not violate her borrowing constraints despite her mistakes.

From future selves' actual consumption rules $\left\{c_{t+k}\left(x_{t+k}\right)\right\}_{t=1}^{T-k-1}$, one can calculate current self t's deliberate consumption rule $c_t^{\text{Deliberate}}\left(x_t\right)$ and find her deliberate MPC $\phi_t^{\text{Deliberate}}$ as usual. I numerically solved the following case: $T \to +\infty$; $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$; $\gamma = 2$; $\sigma = 1$; $\delta = 0.87$; $R = 1.06$; $\underline{a} = 0$, and $\lambda_t = \lambda$.[28]

[26]The income shock Δ considered in the main text can be viewed as a shock to y_0.
[27]Note that the median of y_t is 1.
[28]The value of γ, δ, and R are the same as those used in Section 7, which is from Fagereng et al. (2019).

In Figure 2, I plot a high-liquidity consumer's deliberate MPC $\phi_t^{\text{Deliberate}}$ as a function of λ.[29] I then compare it to the deliberate MPC $\phi_t^{\text{Deliberate}}$ calculated in Proposition 4 without gradual resolution of uncertainty. We can see that the deliberate MPC is effectively the same and the main lesson on how future mistakes in response to saving changes increase the current MPC is unchanged.

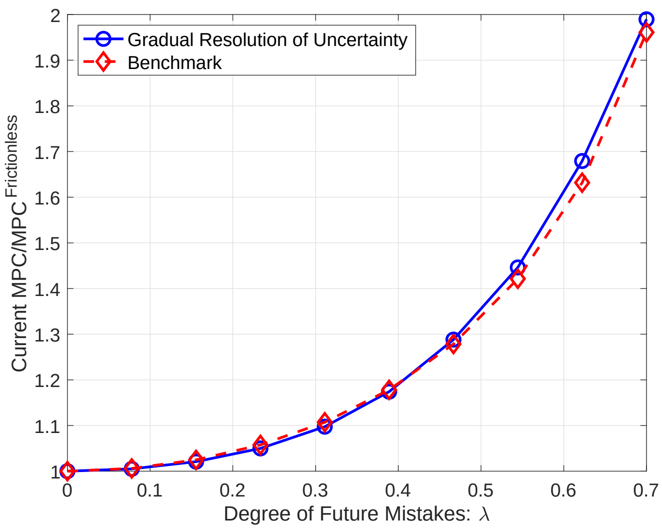

Figure 2: Gradual Resolution of Uncertainty.

5.2 The Precautionary Saving Motive

It is important to clarify that my high MPC result does not come from the precautionary saving motive. This can be easily seen from the high-MPC result in Proposition 3, because the quadratic utility case there a fortiori shuts down the precautionary saving motive. Moreover, as explained below, the precautionary saving motive is not directly about *MPCs.* It is about how the dispersion of the levels of future consumption across states or time decreases the *level* of current consumption (increases the level of saving).

The 3-period example. In the literature, the dispersion of the levels of future consumption behind the precautionary saving motive often comes from future uncertainty or liquidity constraints (Kimball, 1990; Carroll, 1997; Holm, 2018). In my framework, such dispersion can also come from future mistakes in the overall levels of consumption, i.e., $\bar{\lambda}_{t+k}$ in (10) and (22). In fact, this is related to a common intuition: future mistakes in the overall levels of consumption may prompt precautionary behavior and lead to a lower level of current consumption (a higher level of saving).

I now use the simple 3-period example in Section 2 to illustrate this. Similar to (37), the

[29]Since I am focusing on the behavior away from liquidity constraints, I focus on a consumer with initial cash on hand $\bar{x}_0 = 50 \cdot E[y_t]$.

consumer has a $t = 1$ consumption rule

$$c_1(w_1) = \frac{1}{2}w_1 - \bar{\lambda}_1, \tag{37}$$

where $\bar{\lambda}_1$ captures the mistake in the overall consumption level at $t = 1$ in the absence of the shock ($\Delta = 0$). When $\bar{\lambda}_1 > 0$, the consumer under-consumes at $t = 1$. When $\bar{\lambda}_1 < 0$, the consumer over-consumes at $t = 1$.

At $t = 2$, the consumer's consumption rule is then given by

$$c_2(w_2) = w_2 = w_1 - c_1(w_1) = \frac{1}{2}w_1 + \bar{\lambda}_1. \tag{38}$$

We can see that the mistake in the overall consumption level $\bar{\lambda}_1$ introduces the dispersion of consumption levels across time. With a "prudent" utility ($u''' > 0$), such a dispersion will decrease the current consumption level and increase the current saving level.

Proposition 5. *Consider the case with a prudent utility ($u''' > 0$) with (37). For each Δ, $c_0^{Deliberate}(\Delta)$ decreases with $\left|\bar{\lambda}_1\right|$ in a neighborhood of $\bar{\lambda}_1 = 0$.*

One may wonder how this result relates to the precautionary saving behavior driven by uncertainty in Kimball (1990). In Kimball (1990), there is dispersion of future consumption levels across different states, driven by uncertainty. Here, there is a dispersion of future consumption across different periods, driven by mistakes in the overall consumption level $\bar{\lambda}_1$. But the essence is the same: coupled with prudence, such a dispersion decreases the current consumption level and increases the current saving level.

To understand the proof behind Proposition 5, note that the level of current consumption $c_0^{\text{Deliberate}}(\Delta)$ is connected to the first-order derivative of the continuation value function:

$$u'\left(c_0^{\text{Deliberate}}(\Delta)\right) = V_1'\left(\Delta - c_0^{\text{Deliberate}}(\Delta)\right), \tag{39}$$

where

$$V_1'(w_1) = u'\left(\frac{1}{2}w_1 - \bar{\lambda}_1\right) + u'\left(\frac{1}{2}w_1 + \bar{\lambda}_1\right). \tag{40}$$

With prudence ($u''' > 0$), the dispersion of consumption level driven by $\bar{\lambda}_1$ increases the marginal value of saving $V_1'(w_1)$ in (40) and decreases the overall level of current consumption $c_0^{\text{Deliberate}}(\Delta)$ from (39).

Compared to the main high-MPC result in Propositions 3 and 4, Proposition 5 has two key differences. First, Proposition 5 is about the level of current consumption $c_0^{\text{Deliberate}}(\Delta)$ instead of the MPC $\frac{\partial c_0^{\text{Deliberate}}(\Delta)}{\partial \Delta}$. Second, Proposition 5 is about the impact of future mistakes in the overall consumption level $\bar{\lambda}_1$ instead of future mistakes in response to saving changes λ_1. A rough intuition is: the level of current consumption $c_0^{\text{Deliberate}}(\Delta)$ should be connected to future mistakes in the overall consumption level (Proposition 5). On the other hand, the MPC $\phi_0^{\text{Deliberate}}$ is about how the consumer responds to the income shock, so it is directly connected to how future consumption responds to saving changes (Propositions 3 and 4). Since this paper is about the MPC, the latter type of mistake plays a key role throughout.

The general T-period case. Proposition 5 can be easily extended to the general T-period case. Similar to Proposition 5, future mistakes in overall consumption level $\left\{\bar{\lambda}_{t+k}\right\}_{k=1}^{T-t-1}$ prompt precautionary behavior and lead to a lower current consumption level (and a higher current saving level).

To illustrate, similar to Proposition 5, I shut down mistakes in c_{t+k}'s response to saving changes $\{\lambda_{t+k}\}_{k=1}^{T-t-1}$, and focus on mistakes in the overall future consumption level $\left\{\bar{\lambda}_{t+k}\right\}_{k=1}^{T-t-1}$:

$$c_{t+k}\left(w_{t+k}\right)=c_{t+k}^{\text{Deliberate}}\left(w_{t+k}\right)-\bar{\lambda}_{t+k}. \tag{41}$$

When $\bar{\lambda}_{t+k}>0$, the consumer under-consumes at $t+k$, no matter the value of w_{t+k}. When $\bar{\lambda}_{t+k}<0$, the consumer over-consumes at $t+k$.

Proposition 6. *If future consumptions are given by (41) and utilities are prudent ($u'''>0$ and $v'''>0$), $c_t^{Deliberate}(w_t)$ decreases with $\left|\bar{\lambda}_{t+k}\right|$ in a neighborhood of $\bar{\lambda}_{t+k}=0$ for each $k\in\{1,\cdots,T-t-1\}$.*

The precautionary saving motive and MPCs. A natural question is whether the precautionary saving motive driven by future mistakes in overall consumption level $\left\{\bar{\lambda}_{t+k}\right\}_{k=1}^{T-t-1}$ can also impact current MPCs. In theory, this is possible. Taking a derivative with respect to w_t of the FOC in (26), the current MPC is given by:

$$\frac{\partial c_t^{\text{Deliberate}}\left(w_t\right)}{\partial w_t}=\frac{\delta R^2 V_{t+1}''\left(R\left(w_t-c_t^{\text{Deliberate}}\left(w_t\right)\right)\right)}{u''\left(c_t^{\text{Deliberate}}\left(w_t\right)\right)+\delta R^2 V_{t+1}''\left(R\left(w_t-c_t^{\text{Deliberate}}\left(w_t\right)\right)\right)}. \tag{42}$$

From Proposition 6, we know the precautionary saving motive driven by $\left\{\bar{\lambda}_{t+k}\right\}_{k=1}^{T-t-1}$ will decrease $c_t^{\text{Deliberate}}\left(w_t\right)$. Such a decrease in $c_t^{\text{Deliberate}}\left(w_t\right)$ may impact the MPC in (42) through third-order effects when $u'''\neq 0$ and/or $V'''\neq 0$. But such an effect is a degree of order higher than the high-MPC result in Proposition 4. Unless mistakes in overall consumption level $\left\{\bar{\lambda}_{t+k}\right\}_{k=1}^{T-t-1}$ are big, this type of mistake will not impact the MPC that much.

To illustrate this, consider the same environment as in Figure 2, with $u\left(c\right)=\frac{c^{1-\gamma}}{1-\gamma}$; $\gamma=2$; $\sigma=1$; $\delta=0.87$; $R=1.06$; and $\underline{a}=0$. Instead of mistakes in response to saving changes in (35), I focus on mistakes in the overall consumption level $\left\{\bar{\lambda}_{t+k}\right\}_{k=1}^{T-t-1}$. Specifically, similar to (41), these mistakes take the form of an additive deviation from the deliberate counterpart,[30]

$$c_{t+k}\left(w_{t+k}\right)=c_{t+k}^{\text{Deliberate}}\left(w_{t+k}\right)-\bar{\lambda}_{t+k}. \tag{43}$$

When $\bar{\lambda}_{t+k}>0$, self $t+k$'s under-consumes (even in absence of the shock Δ). When $\bar{\lambda}_{t+k}<0$, self $t+k$'s over-consumes. From Figure 3, we can see this type of additive future mistakes in overall consumption levels $\left\{\bar{\lambda}_{t+k}\right\}_{k=1}^{T-t-1}$ effectively does not matter for the current MPC.[31]

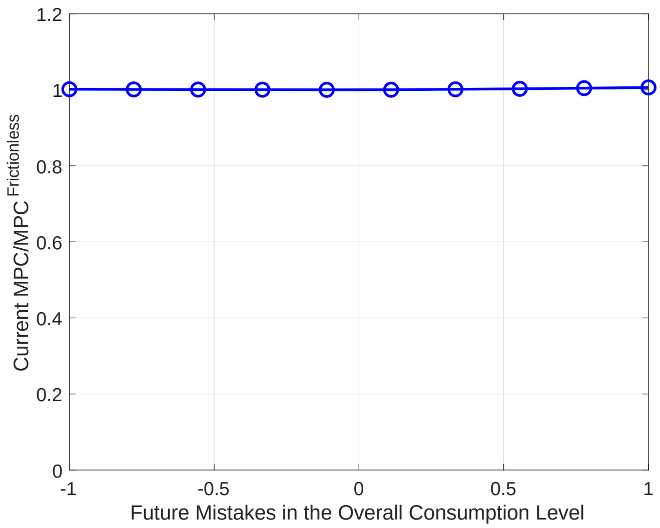

Figure 3: Gradual Resolution of Uncertainty.

In applications, the essentially only possibility that future mistakes in overall consumption levels are large enough to matter for MPCs in (42) is that these mistakes take a multiplicative form

$$c_{t+k}\left(w_{t+k}\right)=c_{t+k}^{\text{Deliberate}}\left(\left(1-\Lambda_{t+k}\right)w_{t+k}\right), \tag{44}$$

where $\Lambda_{t+k}\neq 0$ captures self $t+k$'s mistake. In this case, mistakes in overall consumption level can be very large: at w_{t+k}, self $t+k$ behaves as if her wealth level $\left(1-\Lambda_{t+k}\right)w_{t+k}$, which can deviate significantly from w_{t+k} if w_{t+k} is away from zero. The precautionary saving motive due to those future mistakes can be large, which can impact MPC nontrivially. In Proposition 10 in Appendix B, I provide a thorough analysis of this case. When the utility function is not that

[30]Rigorously, the actual consumption of each future self $k\in\{0,\cdots,T-t-1\}$ is given by

$$c_{t+k}\left(x_{t+k}\right)=\min\left\{\frac{a}{R}+x_{t+k},c_{t+k}^{\text{Deliberate}}\left(x_{t+k}\right)-\bar{\lambda}_{t+k}\right\},$$

which makes sure the consumer will not violate her borrowing constraints despite her mistakes as in (36). As in Figure 2, with uncertainty, it is easier to write the actual consumption rule as a function of cash on hand x_{t+k}.

[31]In Figure 3, the x-axis is $\bar{\lambda}_{t+k}$ (in the unit of the standard deviation of the income risk $\sigma=1$).

concave (EIS> 1), the high-MPC channel focused in the paper in Proposition 4 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.

6 Behavioral Foundations

The main results in the previous sections do not depend on the exact behavioral causes of future consumption mistakes. This section shows how my framework can accommodate many widely-studied behavioral foundations, such as inattention, diagnostic expectations, hyperbolic discounting, and near-rationality (ϵ-mistakes). I explain why these behavioral foundations lead to mistakes in response to saving changes, $\{\lambda_{t+k}\}_{k=1}^{T-t-1}$, the type of mistakes focused on in the paper. I also explain when mistakes only take this form and Proposition 4 applies.

Inattention. My framework can accommodate inattention (e.g. Sims, 2003; Gabaix, 2014; Maćkowiak and Wiederholt, 2015). Here, I follow the sparsity approach in Gabaix (2014) and let each self t's perceived w_t be given by

$$w_t^p(w_t) = (1-\lambda_t)\, w_t + \lambda_t w_t^d, \tag{45}$$

where $\lambda_t \in [0,1]$ captures self t's degree of inattention (a larger λ_t means more attention) and w_t^d captures the default. It is standard to set the default w_t^d to be the pre-shock value $\bar{w}_t$ (Gabaix, 2014).[32] Based on the perceived $w_t^p(w_t)$ in (45), the actual consumption rule for each self t is given by

$$c_t(w_t) = \arg\max_{c_t}\; u(c_t) + \delta V_{t+1}\left(R\left(w_t^p(w_t) - c_t\right)\right), \tag{46}$$

where the continuation value function V_{t+1} is defined as in (16), based on future selves' actual consumption rules.

To isolate the impact of future inattention on current consumption, the deliberate consumption is defined as in (17), taking future selves' inattention $\{\lambda_{t+k}\}_{k=1}^{T-t-1}$ as given:

$$c_t^{\text{Deliberate}}(w_t) = \arg\max_{c_t}\; u(c_t) + \delta V_{t+1}\left(R\left(w_t - c_t\right)\right) \qquad \forall t \in \{0,\cdots,T-1\}.$$

As a corollary of Propositions 3 and 4, future consumption mistakes in the form of inattention lead to higher current MPCs.

[32] An alternative way to model inattention is through noisy signals (Sims, 2003). Those two approaches often lead to similar predictions on behavior. For example, with linear consumption rules and Normally distributed incomes, the two approaches lead to the same predictions on MPCs (see Appendix B).

Corollary 6. *For* $t \in \{0, \cdots, T-2\}$, $\phi_t^{Deliberate}$ *increases with future selves' degrees of inattention* $\{\lambda_{t+k}\}_{k=1}^{T-t-1}$ *if (i) u and v are quadratic functions; or (ii) u and v in are general concave functions and the default wealth* w_t^d *is the pre-shock value* $\bar{w}_t$.

One interpretation of the above result is particularly worth mentioning: it is likely that the consumer is *currently* attentive to the stimulus check ($\lambda_t = 0$) but becomes inattentive to the saving changes driven by the stimulus check over time ($\lambda_{t+k} > 0$). In this case, the impact of future inattention unambiguously translates into a higher current actual MPC.

There is also ample empirical support for consumers' inattention to their saving/borrowing changes and its influence on economic decisions. The credit card literature, e.g., Agarwal et al. (2008) and Stango and Zinman (2014), finds that consumers often neglect their credit card balances, and this neglect often leads to suboptimal credit card usage. Moreover, the recent literature on Fintech shows that providing information about a consumer's total savings by aggregating her financial accounts will change her consumption behavior. Levi (2015) conducts an experiment in which he provides the participants with account aggregation tools that display their current total net saving. Participants significantly change their consumption after seeing the total net saving, implying that they have an imperfect perception of it without the tool. Likewise, Carlin, Olafsson and Pagel (2017) study a financial app that consolidates all of its users' bank account balances. They show that the app significantly reduces its users' interest expenses on consumer debt and other bank fees. In fact, such inattention to saving changes is the focus of the original version of Lian (2019), which the current, more general, paper replaces.

Let me now gauge the magnitudes of how much future inattention can increase current MPCs. As explained above, inattention to saving/wealth changes is the type of inattention that matters for the current MPC, instead of inattention to income.[33] In fact, one can argue that the income shock of interest, e.g. the stimulus check, is often salient. If we assume consumers pay full attention to income, one can then use the ratio between the MPC out of saving/wealth ϕ^a and the MPC out of the current income ϕ^y to gauge the degree of inattention to saving/wealth changes

$$\lambda_{t+k} = \lambda = 1 - \phi^a/\phi^y.$$

The MPC out of saving/wealth ϕ^a and the MPC out of current income ϕ^y are directly available from empirical studies. For example, Di Maggio, Kermani and Majlesi (2018) estimate ϕ^a and ϕ^y for rich households away from liquidity constraints. In their estimates, for consumers in the top half of wealth distribution, $\phi^a = \$0.05$ per year and $\phi^y = \$0.35$ per year. Together, these values imply $\lambda = 1 - 1/7 = 6/7$. Based on this estimated friction λ, the anticipation of future inattention

[33] See Proposition 11 in Appendix C for detail.

can raise the current MPC significantly: $\phi^{\text{Deliberate}}/\phi^{\text{Frictionless}} = 3.77$.[34]

One may argue the MPC out of current income $\phi^y = \$0.35$ used above might be at the higher range of the empirical estimates of high-liquidity consumers' MPCs. But it does reflect a general theme in the recent empirical literature: estimates of the MPC out of saving/wealth are typically much smaller than estimates of the MPC out of current income.[35] If we use a more moderate estimate of the MPC out of current income $\phi^y = 0.20$, this implies $\lambda = 3/4$. In this case, anticipation of future inattention can still double the magnitude of the current MPC compared to the frictionless benchmark: $\phi_0^{\text{Deliberate}}/\phi_0^{\text{Frictionless}} = 2.28$.

Diagnostic expectations. Above, inattention leads to under-reaction of future consumption in response to saving changes. Here I study the case of diagnostic expectation, which leads to over-reaction of future consumption in response to saving changes. Despite this difference, both types of future mistakes lead to higher current MPC as in Propositions 3 and 4.

Specifically, I follow the treatment in Bianchi, Ilut and Saijo (2021), which study how diagnostic expectations impact MPCs. In fact, Propositions 7 and 8 in Bianchi, Ilut and Saijo (2021) show that MPCs under sophistication are higher than MPCs under naivety. That is, the anticipation of future diagnostic expectations increases the current MPC. Through the lens of my paper, this result arises because diagnostic expectations lead to future mistakes in response to saving changes.

To follow closely with Bianchi, Ilut and Saijo (2021), I use the three-period example with quadratic utility in Section 2. This is also the setting of Propositions 7 and 8 in their paper.

In the final period $t = 2$, as in (3), the consumer consumes out of all her remaining savings, $c_2(w_2) = w_2$. 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.[36] 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. Such diagnostic expectations are based on the representativeness heuristic of probabilistic judgments in psychology (Kahneman and Tversky, 1972; Bordalo, Gennaioli and Shleifer, 2018; Bordalo et al., 2020).

Mathematically, the consumer's consumption $c_1(w_1)$ in $t = 1$ is given by

$$u'(c_1(w_1)) = E_1^{\theta}\left[u'(c_2(w_2))\right], \tag{47}$$

[34] This calculation is based on Corollaries 1 or 5.

[35] For example, Chodorow-Reich, Nenov and Simsek (2019)'s estimate of the MPC out of financial wealth is only \$0.028 per year, smaller than Di Maggio, Kermani and Majlesi (2018)'s. Fagereng et al. (2019) also find that rich households consume very little out of capital gains and have a saving rate out of capital gains close to one hundred percent.

[36] There is a delicate point about whether diagnostic expectations should only apply to exogenous variables or also to endogenous variables (e.g., consumption). Bianchi, Ilut and Saijo (2021) argue that the latter case explains the consumption behavior better, which I follow.

where $E_1^\theta[\cdot]$ captures her diagnostic expectation given by[37]

$$E_1^\theta\left[c_2\left(w_2\right)\right]=\left(1+\theta\right)c_2\left(w_2\right), \tag{48}$$

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

$$c_1\left(w_1\right)=\frac{1+\theta}{2+\theta}w_1. \tag{49}$$

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.

One can then define the deliberate consumption as in (5), taking the diagnostic-expectation-driven consumption rule (49) as given. As a corollary to Proposition 3, future diagnostic expectations increase the current MPC, which is exactly what Bianchi, Ilut and Saijo (2021) find in their Propositions 7 and 8: MPCs under sophistication are higher than MPCs under naivete.

Corollary 7. *The current MPC $\phi_0^{Deliberate}$ strictly increases with the degree of future diagnostic expectations θ.*

The result can also be easily extended to the concave case in Proposition 4. 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 4 applies.

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 4.

Specifically, the consumer can put her savings 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, (33) 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 $t+k$'s consumption will be given by

$$c_{t+k}\left(w_{t+k}\right)=\bar{c}_{t+k}+1\cdot\left(w_{t+k}-\bar{w}_{t+k}\right), \tag{50}$$

for all w_{t+k} in a neighborhood of $\bar{w}_{t+k}$.[38]

[37] Note that, in the example in Section 4, the pre-shock outcome $\bar{c}_2=\bar{w}_2=0$. Without this normalization, (48) can be written as $E_1^\theta\left[c_2\left(w_2\right)\right]=\left(1+\theta\right)\left[c_2\left(w_2\right)-c_2\left(\bar{w}_2\right)\right]+\bar{c}_2$.

[38] To derive (50). First, consider a small positive deviation of w_{t+k} away from $\bar{w}_{t+k}$. Because $u'\left(\bar{c}_{t+k}\right)=\delta V'\left(\bar{w}_{t+k+1}\right)$, $u'\left(\bar{c}_{t+k}\right)>\beta_{t+k}\delta V'\left(\bar{w}_{t+k+1}\right)$ for all $\beta_{t+k}<1$. As a result, the present bias will prompt the $t+k$ self to consume out of all the positive deviation $w_{t+k}-\bar{w}_{t+k}$ and (50) holds. Second, consider a small negative

Given (50), I can define the deliberate consumption rule $c_t^{\text{Deliberate}}(w_t)$ as usual. As a corollary of Proposition 4, mistakes in future consumption driven by future present biases will necessarily increase the current MPC.

Corollary 8. *Given any strictly concave utility functions u and v, (33), and the hyperbolic-discounting future consumption rules (50), $\phi_t^{Deliberate} \equiv \frac{\partial c_t^{Deliberate}(\bar{w}_t)}{\partial w_t} \geq \phi_t^{Frictionless}$, where $\phi_t^{Frictionless}$ is the frictionless MPC at $\bar{w}_t$.*

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 $t+k$ is given by

$$c_{t+k}(w_{t+k}) = \arg\max_{c_{t+k}} u(c_{t+k}) + \delta\beta_{t+k} V_{t+k+1}(R(w_{t+k} - c_{t+k})),$$

where $\beta_{t+k} \in [0,1]$ captures self $t+k$'s present bias, which leads to both types of mistakes. Both the focused high-MPC channel in Proposition 4 (because of future mistakes in response to saving changes) and the precautionary saving channel in Proposition 6 (because of mistakes in overall consumption levels) are at force. With CRRA utility, this case maps to the multiplicative case in (43). In Corollary 8 in Appendix B, I provide a thorough analysis of this case. Similar to the discussion after (43), when the utility function is not that concave (EIS>1), the high-MPC channel focused in the paper in Proposition 4 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 (2020).

Near-rationality and $\epsilon-$mistakes. The main mechanism studied in the paper focuses on mistakes in response to saving changes. A natural question is why the consumer may exhibit such mistakes. It turns out that, if the consumer starts from a frictionless pre-shock outcome (33), the utility loss of mistakes in response to saving changes is small, second-order. This is the "near-rationality" argument laid out by Cochrane (1989) and Kueng (2018) about the small welfare loss of inefficient responses to shocks.

To illustrate, I calculate the equivalent monetary loss L, i.e., the amount of money the consumer is willing to pay to avoid mistakes in response to saving changes.

Proposition 7. *If consumption mistakes only come from responses to saving changes (33), the*

deviation of w_{t+k} away from $\bar{w}_{t+k}$. Because of the costly withdrawals from the illiquid assets, the $t+k$ self can only use c_{t+k} to absorb the negative deviation $w_{t+k} - \bar{w}_{t+k}$ and (50) again holds.

equivalent monetary loss is second-order:

$$L \sim \mathbb{O}^2\left(\{\lambda_t\}_{t=0}^{T-1}\right),$$

where $\{\lambda_t\}_{t=0}^{T-1}$ *is defined in (32) and* $\mathbb{O}^2$ *denotes second and higher order terms. This is true for both the quadratic case in Proposition 3 and the general concave utility case in Proposition 4.*

This near-rationality result implies that the consumer may be prone to "ϵ-mistakes." That is, stochastic mistakes that do not bias the consumer's response to saving changes in a particular way. For example, $\lambda_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_t^2)$ in (21) or (32).

We can then study how these future stochastic mistakes $\{\lambda_{t+k}\}_{k=1}^{T-t-1}$ impact the current MPC. Define the deliberate consumption $c_t^{\text{Deliberate}}(w_t)$ as usual given (21) or (32). Similar to Proposition 3 or 4, future stochastic mistakes in response to saving changes lead to higher current MPCs.

Corollary 9. *If* $\lambda_{t+k} \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma_{t+k}^2\right)$*, for* $k \in \{0, \cdots, T-t-1\}$, $\phi_t^{Deliberate} \equiv \frac{\partial c_t^{Deliberate}}{\partial w_t}$ *increases with the variances in future selves' stochastic mistakes,* σ_{t+k}^2*. This is true for both the quadratic case in Proposition 3 and the general concave utility case in Proposition 4.*

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.

Endogenous mistakes and higher MPCs for richer consumers. In the above analysis. I treat the degree of mistakes λs as exogenous. In many behavioral models (e.g., inattention), there is an additional ex ante "stage-0" where λs are endogeneized, balancing the utility loss from mistakes and the cognitive cost of not making mistakes. Then, there is a "stage-1" where the decision maker makes actual consumption decisions given the degree of mistakes λs. The above analysis applies verbatim for such a "stage-1."

There is one additional point worth mentioning. For given λs, richer consumers (with lower $|u''|$) suffer smaller utility losses for a given degree of mistakes (λs). This can be formalized based on the equivalent monetary loss L studied in Proposition 7.

Proposition 8. *If* $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$ *and consumption mistakes come from responses to saving changes ((32) and (33)), the equivalent monetary loss* L *is decreasing in the consumer's initial wealth* $\bar{w}_0$*.*

On the other hand, the cognitive cost of making the correct decision (the cost of decreasing $\{\lambda_s\}_{s=0}^{T-1}$) does not necessarily decrease in the consumer's wealth. For example, in the case of rational inattention, the entropy cost is independent of the consumer's wealth. Together, richer consumers will endogenously choose a higher degree of mistakes $\{\lambda_s\}_{s=0}^{T-1}$. Coupled with the main high-MPC result in Proposition 4, richer consumers then have higher MPCs because they endogenously choose a higher degree of mistakes. This channel may be particularly useful for explaining the puzzling evidence in Kueng (2018) that a consumer's MPC may increase with her income/wealth.

An intra-household interpretation. In fact, most empirical evidence on MPCs is about consumption at the household level (e.g., Kueng, 2018; Fagereng, Holm and Natvik, 2019). Household consumption is decided jointly by the wife and husband. In fact, the unitary model of household spending has long been rejected and it has been widely documented that the wife and husband exhibit different consumption behavior (Thomas, 1990; Browning et al., 1994; Anderson and Baland, 2002; Duflo, 2003; Duflo and Udry, 2004; Ashraf, 2009). In the intertemporal setting here, there is strong evidence against the assumption that the wife and husband can commit their consumption and saving behavior to achieve the ex-ante collective Pareto frontier (Mazzocco, 2007; Lise and Yamada, 2019). Instead, the household consumption behavior fluctuates over time, depending on which spouse has a temporarily higher decision weight.

From the lens of my model, this means that consumption of future selves (e.g., the husband) may deviate from what the current self (e.g., the wife) deems optimal. That is, from the current self's perspective, mistakes in future consumption's response to saving changes $\{\lambda_{t+k}\}_{k=0}^{T-t-1}$ may come from the intra-household friction. From Propositions 3 and 4, we then know that the current self displays a higher MPC.

An interpretation independent of specific biases. Beyond the specific biases studied above, let me provide another interpretation independent of specific biases. From her life experiences, the consumer knows that she has cognitive limitations and her future consumption may not respond efficiently to changes in saving. With this knowledge and even without knowledge of the exact mistakes of her future selves, the consumer will have a higher current MPC.

7 Gauging the Magnitudes

I now gauge the magnitudes of future mistakes' impact on current MPCs. There are two ways to do this exercise. First, one can calibrate λ based on a specific friction and gauge how much anticipation of this particular friction can increase the current MPCs. This is what I did for the case of inattention after Corollary 6. Second, and perhaps more interestingly, one can leverage the "wedge-based" strategy in the paper to gauge the magnitudes of future mistakes' total impact on current MPCs, accommodating different possible behavioral foundations.

Specifically, I will conduct a calibration exercise using the impulse responses of household consumption to unexpected lottery prizes from Fagereng, Holm and Natvik (2019),[39] termed as the intertemporal MPCs in Auclert, Rognlie and Straub (2018). I use these intertemporal MPCs to calculate the relevant wedges capturing future consumption mistakes in response to saving changes.

In Figure 4, I plot the average household consumption response at year k to an unexpected lottery prize at year 0, denoted as $i\text{MPC}_k$.

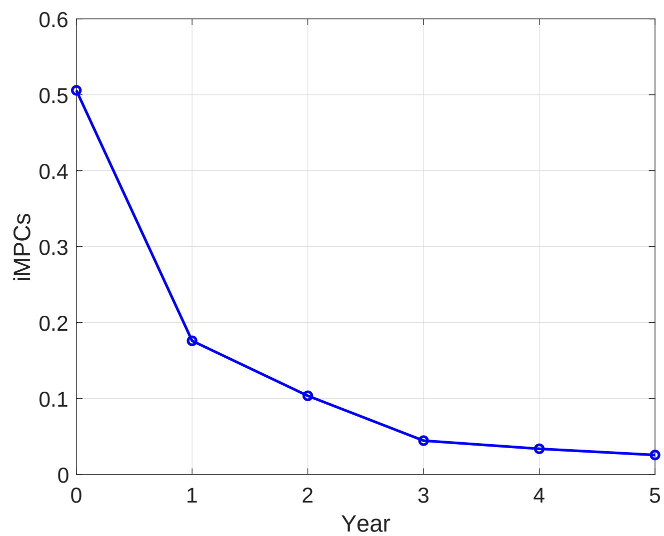

Figure 4: Impulse responses of household consumption to unexpected lottery prizes.

Based on the $i\text{MPC}_k$, I can construct how much actual future consumption responds to saving changes, ϕ_k, for each $k \geq 1$:

$$\phi_k = \frac{i\text{MPC}_k}{1 - \sum_{s=0}^{k-1} i\text{MPC}_s}, \tag{51}$$

which corresponds to ϕ_ks in the main analysis. From Figure 4, we can easily see that those responses embed nontrivial mistakes, otherwise the $i\text{MPC}_k$ curve will be flatter.

I then use these ϕ_ks to calculate the relevant future mistakes in response to saving changes $\{\lambda_k\}_{k\geq 1}$. Specifically, following Fagereng, Holm and Natvik (2019), I consider the CRRA utility with $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$. I use the (δ, R, γ) estimated in Fagereng, Holm and Natvik (2019), $(0.87, 1.06, 2)$. Based on this utility and the actual consumption responses in (51), I can then calculate how much the future deliberate consumption should respond to saving changes, $\phi_k^{\text{Deliberate}}$, for each $k \geq 1$.[40] Together with (51), I can then use $\phi_k = (1 - \lambda_k)\, \phi_k^{\text{Deliberate}}$ to calculate future mistakes in response to saving changes $\{\lambda_k\}_{k\geq 1}$, based on the difference between $\phi_k^{\text{Deliberate}}$ and ϕ_k.

I then study how these future mistakes $\{\lambda_k\}_{k\geq 1}$ increase the current MPC $\phi_0^{\text{Deliberate}}$. The impact of future mistakes can be quite sizable. Based on (65) and (66), anticipation of these future mistakes

[39] I am very grateful to Martin Holm for sharing his data.

[40] Specifically, I use (64), (65), and (66) in the proof of Proposition 4.

can double the magnitude of the current MPC compared to the frictionless benchmark:

$$\phi_0^{\text{Deliberate}} \approx 0.19 \approx 2\phi_0^{\text{Frictionless}} \tag{52}$$

It is true that $\phi_0^{\text{Deliberate}}$ cannot fully explain the high MPC at year 0, $\phi_0 \approx 0.5$, in Figure 4. But that is by design, because $\phi_0^{\text{Deliberate}}$ isolates the impact of future mistakes while shutting down the current mistake λ_0. Of course, the current mistake λ_0 is also important. The goal of the exercise is to show that the impact of future mistakes can be quantitively meaningful.

The exercise here assumes full knowledge of future mistakes, i.e., full sophistication. If the consumer is partially sophisticated, the result above can be viewed as an upper bound on how much future mistakes can increase the current MPC. There is also ample empirical evidence that consumers have at least partial knowledge about their future selves' mistakes and adjust behavior accordingly (e.g., Allcott et al., 2020; Carrera et al., 2021; Le Yaouanq and Schwardmann, 2019). For example, in the context of hyperbolic discounting, Allcott et al. (2020) find that the degree of sophistication (s_t in (30)) is close to 1.[41]

8 Other Applications

The main application in this paper is to show that future consumption mistakes in response to saving changes can help explain high-liquidity consumers' high MPCs. The same mechanism, through the excess concavity of the continuation value function driven by these future mistakes, can also help explain other well-known puzzles in intertemporal decisions. For example, these future mistakes lead to higher risk aversion and help explain the equity premium puzzle (Mehra and Prescott, 1985). Second, these future mistakes can also explain the violation of the fungibility principle (Shefrin and Thaler, 1988), i.e., the prediction that all components of permanent income matter equally through the total present value. In particular, these future mistakes lead to a muted current consumption response to news of future income. That is, the limited "announcement" effect documented by Stephens and Unayama (2011), Parker (2017), and Kueng (2018). Third, these future mistakes lead to a smaller elasticity of intertemporal substitution and speak to the empirical evidence on the small consumption responses to interest rate changes (Hall, 1988; Campbell and Mankiw, 1989; Havránek, 2015).

Risk aversion. A consumer's degree of risk aversion is proportional to the second-order derivatives of her value function: The degree of relative risk aversion is given by $-\frac{\frac{\partial^2 V_t}{\partial w_t^2}}{w_t \frac{\partial V_t}{\partial w_t}}$ and the

[41] Allcott et al. (2020) find a degree of sophistication s_t that ranges from 0.95 to 0.98 for the most experienced group and from 0.79 to 0.89 for the least experienced group.

degree of absolute risk aversion is given by $-\frac{\frac{\partial^2 V_t}{\partial w_t^2}}{\frac{\partial V_t}{\partial w_t}}$, both proportional to $\frac{\partial^2 V_t}{\partial w_t^2}$. From Lemma 1, we know that consumption mistakes lead to the excess concavity of the value function. We then know that consumption mistakes will also lead to a larger risk aversion. In terms of the magnitude, we can use the same calibration as in Section 7. From (52), we know that future consumption mistakes in response to saving changes can double the degree of risk aversion.

Excess discounting of future income. Empirical studies often find limited consumption responses to news about future income, i.e., a limited "announcement effect." Papers documenting this pattern away from liquidity constraints include Stephens and Unayama (2011), Parker (2017), and Kueng (2018). This is consistent with Thaler (1990)'s observation that the consumer often exhibits excess discounting of future income.

Future mistakes in response to saving changes can generate such an excess discounting of future income. To understand the intuition, note that, the response of current consumption to future income necessarily leads to saving changes, since the announced income has not arrived yet. When future selves respond inefficiently to changes in saving, the current self will be less willing to change her saving and display a muted consumption response to news about future income.

A detailed analysis involves the full non-fungibility case mentioned in Section 4 and requires the introduction of new notations. The detail can be found in Proposition 13 in Appendix C. Here, I will state the result for the simple case in which future selves only exhibit mistakes in response to saving changes ($\lambda^a_{t+k} = \lambda^a$), while they respond to current and future incomes perfectly ($\lambda^y_{t+k} = 0$). For the illustration purpose here, I consider the quadratic case in Section 4 and let $T \to +\infty$.

Proposition 9. *In this case, each self t's deliberate consumption rule is given by:*[42]

$$c_t^{Deliberate}(a_t, s_t) = \phi^{Deliberate}\left(a_t + y_t + \sum_{k=1}^{+\infty}\left(\omega^{Deliberate}\right)^k R^{-k} y_{t+k}\right) + \hat{c}_t^{Deliberate}, \tag{53}$$

where $\phi^{Deliberate} \equiv \frac{\delta R^2 - 1}{\delta R^2\left(1-(\lambda^a)^2\right)}$ *is the same as (27),* $\omega^{Deliberate} \equiv 1 - \frac{\left(\delta R^2-1\right)(\lambda^a)^2}{1-(\lambda^a)^2} \in [0,1]$, *and* $\hat{c}_t^{Deliberate}$ *is a scalar.*[43]

(53) shows that future mistakes in response to saving changes lead to excess discounting of future income. In response to a future income shock k period from now, the consumer behaves as if she discounts it by a factor $\left(\omega^{\text{Deliberate}}\right)^k$. Moreover, the consumer exhibits more discounting when her future selves exhibits larger mistakes in response to saving changes (a larger $|\lambda^a|$) or she responds to income shocks further in the future (a larger k).

[42] In (53), $s_t = (y_t, \cdots, y_T)$ is the income state.

[43] Similar to Corollary 1, for the transversality condition to hold, we need $\delta R^2 < 1$ and $|\lambda^a| < \left(\delta^{-1/2} R^{-1}\right)$.

Furthermore, when $\lambda^a \to \left(\delta^{-1/2}R^{-1}\right)^-$, (53) becomes effectively a "hand-to-mouth" limit:

$$c_t^{\text{Deliberate}}\left(a_t, s_t\right) = 1 \cdot \left(a_t + y_t\right) + 0 \cdot \left(\sum_{k=1}^{+\infty} R^{-k} y_{t+k}\right) + \hat{c}_t^{\text{Deliberate}}.$$

That is, when the current self is so worried about future mistakes in response to saving changes, she becomes unwilling to change her savings. As a result, she does not respond to changes in future income and absorbs all changes in current income. In other words, she is effectively "hand-to-mouth" with respect to *changes* in income, even though her pre-shock consumption level does not need to track her pre-shock cash on hand.

This simple "hand-to-mouth" limit also illustrates how my mechanism can explain the empirical evidence on excess sensitivity to *anticipated* income shocks away from liquidity constraint (e.g. Kueng, 2018). In this limit, consumption does not respond to future income until it arrives. At that point, consumption fully absorbs the anticipated income shock.

Future non-fungibility begets current non-fungibility. Thaler (1990) observes that the fungibility principle, i.e., the prediction of the permanent income hypothesis that consumption is only a function of the total present value of all components of income and savings, is often violated in practice. The above result about the excess discounting of future income illustrates a broader theme: the non-fungibility of future consumption by itself suffices to generate non-fungibility of current consumption. That is, even if the current self understands how to calculate permanent income correctly, as long as she anticipates future consumption mistakes in differentially responding to different components of permanent income, she will also respond differentially to different components of permanent income. In this sense, future non-fungibility begets current non-fungibility. See Proposition 12 in Appendix C about why this is generically true.

A smaller effect of interest rate changes. Another famous puzzle in intertemporal consumption is the empirical evidence on the weak intertemporal substitution motive and the small response of consumption to interest rate changes (Hall, 1988; Campbell and Mankiw, 1989; Havránek, 2015). My proposed channel, i.e., the anticipation of future consumption mistakes in response to saving changes, can also help resolve this puzzle.

The intuition is similar: the response of current consumption to interest rate changes leads to saving changes; with future consumption mistakes in response to saving changes, the current self is less willing to respond to interest rate changes. See Proposition 14 in Appendix C for a formalization.

9 Conclusion

In this paper, I show how inefficient responses of future consumption to saving changes lead to high marginal propensities to consume now. This channel is independent of liquidity constraints and helps explain the empirical puzzles on high liquidity consumers' high MPCs. The main approach, using wedges to capture behavioral mistakes and deriving robust predictions of sophistication independent of the exact psychological cause of these mistakes, can be useful in many other contexts.

Appendix A: Proofs

Proof of Proposition 1. We consider the more general specification of $t = 1$ consumption rule in (10). Based on (5), we have

$$u'\left(c_0^{\text{Deliberate}}(\Delta)\right) = \frac{1}{2}(1-\lambda_1)\,u'\left(\frac{1}{2}(1-\lambda_1)\,w_1 - \bar{\lambda}_1\right) + \frac{1}{2}(1+\lambda_1)\,u'\left(\frac{1}{2}(1+\lambda_1)\,w_1 + \bar{\lambda}_1\right),$$

where $w_1 = \Delta - c_0^{\text{Deliberate}}(\Delta)$. Since u is quadratic, we know $c_0^{\text{Deliberate}}(\Delta)$ is linear. As a result, we have

$$\phi_0^{\text{Deliberate}} = \frac{1}{4}\left[(1-\lambda_1)^2 + (1+\lambda_1)^2\right]\left(1 - \phi_0^{\text{Deliberate}}\right),$$

and

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

Proposition 1 follows directly.

Proof of Proposition 2. The definition of deliberate consumption in (14) at t together with the definition of the value function in (16) at $t+1$ lead to (17). The recursive formulation for the value function in (18) follows directly from the definition of the value function in (16).

Now, consider consumption rules and value functions $\left\{c_t^{\text{Deliberate}}(w_t), c_t(w_t)\right\}_{t=0}^{T-1}$ and $\{V_t(w_t)\}_{t=0}^{T}$ satisfy (15), (17), (18), and the boundary condition $V_T(w_T) = v(w_T)$. Since I am working with a finite horizon problem, I can iterate those conditions through backward induction and arrive at the sequential form in (14) – (16).

Proof of Lemma 1 and Proposition 3. I work with backward induction. At T, I have:

$$\Gamma_T = \frac{v''}{u''}.$$

For each $t \leq T-1$, from (17), the deliberate MPC is given by

$$\phi_t^{\text{Deliberate}} = \frac{\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}}. \quad (54)$$

From (21), the actual MPC is given by

$$\phi_t = \frac{(1-\lambda_t)\,\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}}. \quad (55)$$

From the recursive formulation of the value function in (17), we have:

$$\frac{\partial V_t(w_t)}{\partial w_t} = \phi_t u'(c_t(w_t)) + (1-\phi_t)\,\delta R \frac{\partial V_{t+1}(w_{t+1})}{\partial w_{t+1}}. \tag{56}$$

Together with the budget constraint $w_{t+1} = R(w_t - c_t)$, we have:

$$\begin{aligned}
\Gamma_t &= (\phi_t)^2 + (1-\phi_t)^2\,\Gamma_{t+1}\delta R^2 \\
&= \left(1+\Gamma_{t+1}\delta R^2\right)\left(\phi_t - \frac{\delta R^2\Gamma_{t+1}}{1+\delta R^2\Gamma_{t+1}}\right)^2 + \frac{\delta R^2\Gamma_{t+1}}{1+\delta R^2\Gamma_{t+1}} \\
&= \frac{\left(\delta R^2\Gamma_{t+1}\right)^2}{1+\delta R^2\Gamma_{t+1}}\lambda_t^2 + \frac{\delta R^2\Gamma_{t+1}}{1+\delta R^2\Gamma_{t+1}}.
\end{aligned} \tag{57}$$

Lemma 1 and Proposition 3 then follow directly.

Proof of Corollary 1. From (57), we know that $\Gamma_t = \frac{\left(\delta R^2\Gamma_{t+1}\right)^2}{1+\delta R^2\Gamma_{t+1}}\lambda^2 + \frac{\delta R^2\Gamma_{t+1}}{1+\delta R^2\Gamma_{t+1}} \equiv f(\Gamma_{t+1})$, with $f(x) \equiv \frac{\delta R^2 x}{1+\delta R^2 x} + \frac{\left(\delta R^2 x\right)^2}{1+\delta R^2 x}\lambda^2 = \frac{\delta R^2 x}{1+\delta R^2 x}\left(1+\lambda^2\delta R^2 x\right)$. We also know that $\Gamma_T = \frac{v''}{u''} > 0$.

Let $\Gamma = \frac{\delta R^2 - 1}{\delta R^2(1-\delta R^2\lambda^2)}$ denote the fix point of f. That is $f(\Gamma) = \Gamma$. Moreover, as long as $\delta R^2 < 1$ and $0 \le \lambda < \delta^{-1/2}R^{-1}$, 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 $\Gamma > \frac{v''}{u''} = \Gamma_T$. We have $\Gamma > \Gamma_t = f^{(T-t)}(\Gamma_T) > f^{(T-t-1)}(\Gamma_T) > \cdots > \frac{v''}{u''} = \Gamma_T$. As a result, $\Gamma_t = f^{(T-t)}(\Gamma_T)$ converges to the fix point Γ with $T \to +\infty$.

2) If $\Gamma < \frac{v''}{u''} = \Gamma_T$. We have $\Gamma < \Gamma_t = f^{(T-t)}(\Gamma_T) < f^{(T-t-1)}(\Gamma_T) < \cdots < \frac{v''}{u''} = \Gamma_T$. As a result, $\Gamma_t = f^{(T-t)}(\Gamma_T)$ converges to the fix point Γ with $T \to +\infty$.

Together, one way or another, as long as $\delta R^2 < 1$ and $0 \le \lambda < \delta^{-1/2}R^{-1}$, $\Gamma_t \to \Gamma$ with $T \to +\infty$. From (54), we then have, with $T \to +\infty$.

$$\phi_t^{\text{Deliberate}} \to \phi^{\text{Deliberate}} \equiv \frac{\delta R^2\Gamma}{1+\delta R^2\Gamma} = \frac{\delta R^2 - 1}{\delta R^2\left(1-\lambda^2\right)}.$$

Proof of Corollary 2. With graduate resolution of uncertainty, the optimal deliberate consumption in (17) becomes

$$c_t^{\text{Deliberate}}(x_t) = \max_{c_t}\; u(c_t) + \delta E_t\left[V_{t+1}\left(R(x_t - c_t) + y_{t+1}\right)\right],$$

while the recursive formulation for the value function in (18) becomes

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

where $E_t[\cdot]$ captures rational expectations based on period t's information.

The proof of Proposition 2 remains unchanged, except (56) becomes

$$\frac{\partial V_t(x_t)}{\partial w_t} = \phi_t u'(c_t(x_t)) + (1-\phi_t)\,\delta R E_t\left[\frac{\partial V_{t+1}((x_t - c_t(x_t)) + y_{t+1})}{\partial w_{t+1}}\right].$$

In particular, the formula (54), (56), and (57) remain unchanged. So Corollary 2 follows directly.

Proof of Corollary 3 and Corollary 4. Let $\{\tilde{c}_{t,t+k}(w_{t+k})\}_{k=1}^{T-t-1}$ capture self t's perceived future consumption rules. I redefine the deliberate consumption based on these perceived future consumption rules:

$$c_t^{\text{Deliberate}}(w_t) \equiv \arg\max_{c_t} u(c_t) + \sum_{k=1}^{T-t-1} \delta^k u(\tilde{c}_{t,t+k}(w_{t+k})) + \delta^{T-t} v(w_T) \tag{58}$$

subject to the budget $w_{t+k} = R(w_{t+k-1} - c_{t+k-1})$.

Based on self t's perceived future consumption rules $\{\tilde{c}_{t,t+k}(w_{t+k})\}_{k=1}^{T-t-1}$, we first find the consumption that self t believes that self $t+k$ would been chosen if self $t+k$ is not subject to any behavioral mistake and takes future consumption rules as given by $\{\tilde{c}_{t,t+k+l}(w_{t+k+l})\}_{l=1}^{T-t-k-1}$:

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

subject to the budget.

We can define perceived self $t+k$ future mistake $\tilde{\lambda}_{t,t+k}$ similar to (21):

$$\tilde{\phi}_{t,t+k} \equiv \frac{\partial \tilde{c}_{t,t+k}}{\partial w_{t+k}} = \left(1-\tilde{\lambda}_{t,t+k}\right)\frac{\partial \tilde{c}_{t,t+k}^{\text{Deliberate}}}{\partial w_{t+k}} \equiv \left(1-\tilde{\lambda}_{t,t+k}\right)\tilde{\phi}_{t,t+k}^{\text{Deliberate}}. \tag{59}$$

Based on (58) – (59), the proof of Proposition 3 goes through exactly, with perceived future mistakes $\tilde{\lambda}_{t,t+k}$ replacing the role of actual future mistakes λ_{t+k}. Corollary 3 then follows.

Corollary 4 then follows directly from Corollary 3 and (30).

Proof of Proposition 4. The recursive formulation in Proposition 2 remains to hold. 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[44]

$$u'\left(c_t^{\text{Deliberate}}\left(w_t\right)\right)=R\delta V_{t+1}'\left(R\left(w_t-c_t^{\text{Deliberate}}\left(w_t\right)\right)\right). \tag{60}$$

We henceforth have:

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

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

$$\frac{\partial c_t^{\text{Deliberate}}\left(\bar{w}_t\right)}{\partial w_t}=\frac{R^2\delta\frac{\partial^2 V_{t+1}\left(\bar{w}_{t+1}\right)}{\partial w_{t+1}^2}}{u''\left(c_t^{\text{Deliberate}}\left(\bar{w}_t\right)\right)+R^2\delta\frac{\partial^2 V_{t+1}\left(\bar{w}_{t+1}\right)}{\partial w_{t+1}^2}}. \tag{61}$$

From (18):

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

As a result,

$$\frac{\partial V_t\left(w_t\right)}{\partial w_t}=\frac{\partial c_t\left(w_t\right)}{\partial w_t}u'\left(c_t\left(w_t\right)\right)+\left(1-\frac{\partial c_t\left(w_t\right)}{\partial w_t}\right)\delta R\frac{\partial V_{t+1}\left(w_{t+1}\right)}{\partial w_{t+1}}, \tag{62}$$

and

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

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 (60), we have $u'\left(c_t\left(\bar{w}_t\right)\right)=\delta R\frac{\partial V_{t+1}\left(\bar{w}_{t+1}\right)}{\partial w_{t+1}}$. As a result,

$$\frac{\partial^2 V_t\left(\bar{w}_t\right)}{\partial w_t^2}=\left(\frac{\partial c_t\left(\bar{w}_t\right)}{\partial w_t}\right)^2u''\left(c_t\left(\bar{w}_t\right)\right)+\left(1-\frac{\partial c_t\left(\bar{w}_t\right)}{\partial w_t}\right)^2\delta R^2\frac{\partial^2 V_{t+1}\left(\bar{w}_{t+1}\right)}{\partial w_{t+1}^2}. \tag{63}$$

Define $\Gamma_t\equiv\frac{\partial^2 V_t\left(\bar{w}_t\right)}{\partial w_t^2}/u''\left(c_t\left(\bar{w}_t\right)\right)$, $\phi_t^{\text{Deliberate}}\equiv\frac{\partial c_t^{\text{Deliberate}}\left(\bar{w}_t\right)}{\partial w_t}$, and

$$\phi_t\equiv\frac{\partial c_t\left(\bar{w}_t\right)}{\partial w_t}\equiv\left(1-\lambda_t\right)\frac{\partial c_t^{\text{Deliberate}}\left(\bar{w}_t\right)}{\partial w_t}. \tag{64}$$

[44] This equation imposes the concavity of the continuation value $V_{t+1}\left(w_{t+1}\right)$. This is true around the path $\{\bar{w}_s,\bar{c}_s\}$ because $\frac{\partial^2 V_{t+1}\left(\bar{w}_{t+1}\right)}{\partial w_{t+1}^2}=u''\cdot\Gamma_{t+1}<0$, as proved below.

From (61) and (63), we have

$$\phi_t^{\text{Deliberate}} = \frac{\delta R^2 \Gamma_{t+1} \frac{u''(\bar{c}_{t+1})}{u''(\bar{c}_t)}}{1 + R^2 \delta \Gamma_{t+1} \frac{u''(\bar{c}_{t+1})}{u''(\bar{c}_t)}} \tag{65}$$

and

$$\begin{aligned}\Gamma_t &= \phi_t^2 + (1-\phi_t)^2 \Gamma_{t+1} \delta R^2 \frac{u''(\bar{c}_{t+1})}{u''(\bar{c}_t)}. \\ &= \frac{\left(\delta R^2 \Gamma_{t+1} \frac{u''(\bar{c}_{t+1})}{u''(\bar{c}_t)}\right)^2}{1 + \delta R^2 \Gamma_{t+1} \frac{u''(\bar{c}_{t+1})}{u''(\bar{c}_t)}} \lambda_t^2 + \frac{\delta R^2 \Gamma_{t+1} \frac{u''(\bar{c}_{t+1})}{u''(\bar{c}_t)}}{1 + \delta R^2 \Gamma_{t+1} \frac{u''(\bar{c}_{t+1})}{u''(\bar{c}_t)}}. \end{aligned} \tag{66}$$

Proposition 4 then follows.

Proof of Corollary 5. For the pre-shock ($\bar{\Delta} = 0$) outcome, from (33), we have

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

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

$$\frac{\bar{c}_{t+1}}{\bar{c}_t} = (\delta R)^{\frac{1}{\gamma}}. \tag{67}$$

Substituting it into (65) and (66), we have

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

and

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

Similar to the steps in Corollary 1, if $\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} > 1$ and $\lambda_{t+k} = \lambda$ with $|\lambda| < \left(\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}}\right)^{-\frac{1}{2}}$, we have, for $T \to +\infty$,

$$\begin{aligned}\Gamma_t^{\text{Deliberate}} &\to \Gamma \equiv \frac{\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} - 1}{\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} \left[1 - \left(\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}}\right) \lambda^2\right]}, \\ \phi_t^{\text{Deliberate}} &\to \phi^{\text{Deliberate}} \equiv \frac{\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} - 1}{\delta^{-\frac{1}{\gamma}} R^{1-\frac{1}{\gamma}} (1 - \lambda^2)}.\end{aligned}$$

Proof of Proposition 5. Based on (6), 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

$$
\frac{\partial V_1'\left(w_1;0\right)}{\partial\bar{\lambda}_1}=0 \quad \text{and} \quad \frac{\partial V_1'\left(w_1;0\right)}{\partial\bar{\lambda}_1^2}>0 \tag{68}
$$

Based on (5), we have

$$
\begin{aligned}
u'\left(c_0^{\text{Deliberate}}\left(\Delta;\bar{\lambda}_1\right)\right) &= V_1'\left(\Delta-c_0^{\text{Deliberate}}\left(\Delta;\bar{\lambda}_1\right);\bar{\lambda}_1\right),\\
u''\left(c_0^{\text{Deliberate}}\left(\Delta;\bar{\lambda}_1\right)\right)\frac{\partial c_0^{\text{Deliberate}}\left(\Delta;\bar{\lambda}_1\right)}{\partial\bar{\lambda}_1} &= -V_1''\left(\Delta-c_0^{\text{Deliberate}}\left(\Delta;\bar{\lambda}_1\right);\bar{\lambda}_1\right)\frac{\partial c_0^{\text{Deliberate}}\left(\Delta;\bar{\lambda}_1\right)}{\partial\bar{\lambda}_1}\\
&\quad+\frac{\partial V_1'\left(\Delta-c_0^{\text{Deliberate}}\left(\Delta;\bar{\lambda}_1\right);\bar{\lambda}_1\right)}{\partial\bar{\lambda}_1}.
\end{aligned} \tag{69}
$$

Together with (68), we have

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

and

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

As a result,

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

This proves Proposition 5.

Proof of Proposition 6. Because I assume u, v, and c_t are third-order continuously differentiable, V_t is third-order continuously differentiable too. Based on (26), we have

$$u'\left(c_t^{\text{Deliberate}}\left(w_t;\left\{\bar{\lambda}_{t+s}\right\}_{s=1}^{T-t-1}\right)\right)=\delta R\frac{\partial V_{t+1}}{\partial w_{t+1}}\left(R\left(w_t-c_t^{\text{Deliberate}}\left(w_t;\left\{\bar{\lambda}_{t+s}\right\}_{s=1}^{T-t-1}\right)\right);\left\{\bar{\lambda}_{t+s}\right\}_{s=1}^{T-t-1}\right).$$

From and (18) and(41) :

$$\begin{aligned}V_t\left(w_t;\left\{\bar{\lambda}_{t+s}\right\}_{s=0}^{T-t-1}\right)&=u\left(c_t^{\text{Deliberate}}\left(w_t;\left\{\bar{\lambda}_{t+s}\right\}_{s=1}^{T-t-1}\right)-\bar{\lambda}_t\right)\\&\quad+\delta V_{t+1}\left(R\left(w_t-c_t^{\text{Deliberate}}\left(w_t;\left\{\bar{\lambda}_{t+s}\right\}_{s=1}^{T-t-1}\right)+\bar{\lambda}_t\right);\left\{\bar{\lambda}_{t+s}\right\}_{s=1}^{T-t-1}\right).\end{aligned}$$

As a result,

$$\frac{\partial V_t}{\partial w_t}=\frac{\partial c_t^{\text{Deliberate}}}{\partial w_t}u'+\left(1-\frac{\partial c_t^{\text{Deliberate}}}{\partial w_t}\right)\delta R\frac{\partial V_{t+1}}{\partial w_{t+1}},\tag{70}$$

$$\frac{\partial^2 V_t}{\partial w_t^2}=\left(\frac{\partial c_t^{\text{Deliberate}}}{\partial w_t}\right)^2u''+\left(1-\frac{\partial c_t^{\text{Deliberate}}}{\partial w_t}\right)^2\delta R^2\frac{\partial^2 V_{t+1}}{\partial w_{t+1}^2}+\frac{\partial^2 c_t^{\text{Deliberate}}}{\partial w_t^2}\left(u'-\delta R\frac{\partial V_{t+1}}{\partial w_{t+1}}\right).$$

$$\begin{aligned}\frac{\partial^3 V_t}{\partial w_t^3}&=\left(\frac{\partial c_t^{\text{Deliberate}}}{\partial w_t}\right)^3u'''+\left(1-\frac{\partial c_t^{\text{Deliberate}}}{\partial w_t}\right)^3\delta R^3\frac{\partial^3 V_{t+1}}{\partial w_{t+1}^3}+\\&\frac{\partial^3 c_t^{\text{Deliberate}}}{\partial w_t^3}\left(u'-\delta R\frac{\partial V_{t+1}}{\partial w_{t+1}}\right)+\frac{\partial^2 c_t^{\text{Deliberate}}}{\partial w_t^2}\left(u''\frac{\partial c_t^{\text{Deliberate}}}{\partial w_t}-\left(1-\frac{\partial c_t^{\text{Deliberate}}}{\partial w_t}\right)\delta R^2\frac{\partial^2 V_{t+1}}{\partial w_{t+1}^2}\right)\end{aligned}$$

where I suppress the arguments of functions for notation simplicity. Evaluated at $\lambda=0$, we have

$$u'\left(c_t^{\text{Deliberate}}\left(w_t;\{0\}_{s=1}^{T-t-1}\right)\right)=\delta R\frac{\partial V_{t+1}\left(R\left(w_t-c_t^{\text{Deliberate}}\left(w_t;\{0\}_{s=1}^{T-t-1}\right)\right);\{0\}_{s=1}^{T-t-1}\right)}{\partial w_{t+1}}.$$

$$u''\left(c_t^{\text{Deliberate}}\left(w_t;\{0\}_{s=1}^{T-t-1}\right)\right)\frac{\frac{\partial c_t^{\text{Deliberate}}\left(w_t;\{0\}_{s=1}^{T-t-1}\right)}{\partial w_t}}{1-\frac{\partial c_t^{\text{Deliberate}}\left(w_t;\{0\}_{s=1}^{T-t-1}\right)}{\partial w_t}}=\delta R^2\frac{\partial^2 V_{t+1}\left(R\left(w_t-c_t^{\text{Deliberate}}\left(w_t;\{0\}_{s=1}^{T-t-1}\right)\right);\{0\}_{s=1}^{T-t-1}\right)}{\partial w_{t+1}^2}$$

$$\frac{\partial^2 V_t\left(w_t;\{0\}_{s=0}^{T-t-1}\right)}{\partial w_t^2}<0\quad\text{and}\quad\frac{\partial^3 V_t\left(w_t;\{0\}_{s=0}^{T-t-1}\right)}{\partial w_t^3}>0,$$

where I use prudence and induction for the last inequality.

From (70), we also have

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

$$\frac{\partial^3 V_t}{\partial w_t \partial^2 \bar{\lambda}_t} = u''' \frac{\partial c_t^{\text{Deliberate}}}{\partial w_t} + \delta R^2 \frac{\partial^3 V_{t+1}}{\partial w_{t+1}^3} \left(1 - \frac{\partial c_t^{\text{Deliberate}}}{\partial w_t}\right)$$

$$\frac{\partial^2 V_t}{\partial w_t \partial \bar{\lambda}_{t+k}} = \frac{\partial^2 c_t^{\text{Deliberate}}}{\partial w_t \partial \bar{\lambda}_{t+k}} \left(u' - \delta R \frac{\partial V_{t+1}}{\partial w_{t+1}}\right) + \frac{\partial c_t^{\text{Deliberate}}}{\partial \bar{\lambda}_{t+k}} \left[\frac{\partial c_t^{\text{Deliberate}}}{\partial w_t} u'' - \left(1 - \frac{\partial c_t^{\text{Deliberate}}}{\partial w_t}\right) \delta R^2 \frac{\partial^2 V_{t+1}}{\partial w_{t+1}^2}\right],$$

$$\begin{aligned} \frac{\partial^3 V_t}{\partial w_t \partial \bar{\lambda}_{t+k}^2} = {} & \frac{\partial^3 c_t^{\text{Deliberate}}}{\partial w_t \partial \bar{\lambda}_{t+k}^2} \left(u' - \delta R \frac{\partial V_{t+1}}{\partial w_{t+1}}\right) + \frac{\partial^2 c_t^{\text{Deliberate}}}{\partial w_t \partial \bar{\lambda}_{t+k}} \left(u'' \frac{\partial c_t^{\text{Deliberate}}}{\partial w_t} - \left(1 - \frac{\partial c_t^{\text{Deliberate}}}{\partial w_t}\right) \delta R^2 \frac{\partial^2 V_{t+1}}{\partial w_{t+1}^2}\right) \\ & + \frac{\partial^2 c_t^{\text{Deliberate}}}{\partial \bar{\lambda}_{t+k}^2} \left[\frac{\partial c_t^{\text{Deliberate}}}{\partial w_t} u'' + \left(1 - \frac{\partial c_t^{\text{Deliberate}}}{\partial w_t}\right) \delta R^2 \frac{\partial^2 V_{t+1}}{\partial w_{t+1}^2}\right] \\ & + \frac{\partial c_t^{\text{Deliberate}}}{\partial \bar{\lambda}_{t+k}} \left[\frac{\partial c_t^{\text{Deliberate}}}{\partial \bar{\lambda}_{t+k}} \left[\frac{\partial c_t^{\text{Deliberate}}}{\partial w_t} u''' - \left(1 - \frac{\partial c_t^{\text{Deliberate}}}{\partial w_t}\right) \delta R^3 \frac{\partial^3 V_{t+1}}{\partial w_{t+1}^3}\right] + \frac{\partial^2 c_t^{\text{Deliberate}}}{\partial w_t \partial \bar{\lambda}_{t+k}} \left(u'' + \delta R^2 \frac{\partial^2 V_{t+1}}{\partial w_{t+1}^2}\right)\right] \end{aligned}$$

Based on (26), we have

$$u'\left(c_t^{\text{Deliberate}}\left(w_t; \left\{\bar{\lambda}_{t+s}\right\}_{s=1}^{T-t-1}\right)\right) = \delta R \frac{\partial V_{t+1}}{\partial w_{t+1}} \left(R\left(w_t - c_t^{\text{Deliberate}}\left(w_t; \left\{\bar{\lambda}_{t+s}\right\}_{s=1}^{T-t-1}\right)\right); \left\{\bar{\lambda}_{t+s}\right\}_{s=1}^{T-t-1}\right).$$

$$\delta R \frac{\partial^2 V_{t+1}}{\partial w_{t+1} \partial \bar{\lambda}_{t+k}} = \left[u''\left(c_t^{\text{Deliberate}}\right) + \delta R^2 \frac{\partial^2 V_{t+1}}{\partial w_{t+1}^2}\right] \frac{\partial c_t^{\text{Deliberate}}}{\partial \bar{\lambda}_{t+k}}$$

$$\begin{aligned} \delta R \frac{\partial^3 V_{t+1}}{\partial w_{t+1} \partial \bar{\lambda}_{t+k}^2} = {} & \left[u''\left(c_t^{\text{Deliberate}}\right) + \delta R^2 \frac{\partial^2 V_{t+1}}{\partial w_{t+1}^2}\right] \frac{\partial^2 c_t^{\text{Deliberate}}}{\partial \bar{\lambda}_{t+k}^2} \\ & + \left[u'''\left(c_t^{\text{Deliberate}}\right) \frac{\partial c_t^{\text{Deliberate}}}{\partial \bar{\lambda}_{t+k}} + 2\delta R^2 \frac{\partial^3 V_{t+1}}{\partial \bar{\lambda}_{t+k} \partial w_{t+1}^2} - \delta R^3 \frac{\partial^3 V_{t+1}}{\partial w_{t+1}^3} \frac{\partial c_t^{\text{Deliberate}}}{\partial \bar{\lambda}_{t+k}}\right] \frac{\partial c_t^{\text{Deliberate}}}{\partial \bar{\lambda}_{t+k}}. \end{aligned}$$

Evaluated everything at $\lambda = 0$, we have

$$\frac{\partial^2 V_t\left(w_t; \{0\}_{s=0}^{T-t-1}\right)}{\partial w_t \partial \bar{\lambda}_t} = 0 \quad \text{and} \quad \frac{\partial^3 V_t\left(w_t; \{0\}_{s=0}^{T-t-1}\right)}{\partial w_t \partial^2 \bar{\lambda}_t} > 0.$$

$$\frac{\partial V_t^2\left(w_t; \{0\}_{s=0}^{T-t-1}\right)}{\partial w_t \partial \bar{\lambda}_{t+k}} = 0 \quad \text{and} \quad \frac{\partial c_t^{\text{Deliberate}} V_t\left(w_t; \{0\}_{s=1}^{T-t-1}\right)}{\partial \bar{\lambda}_{t+k}} = 0 \tag{71}$$

$$\frac{\partial^2 c_t^{\text{Deliberate}}\left(w_t; \{0\}_{s=1}^{T-t-1}\right)}{\partial \bar{\lambda}_{t+k}^2} = \frac{\delta R \frac{\partial^3 V_{t+1}\left(R\left(w_t - c_t^{\text{Deliberate}}\left(w_t; \{0\}_{s=1}^{T-t-1}\right)\right); \{0\}_{s=1}^{T-t-1}\right)}{\partial w_{t+1} \partial \bar{\lambda}_{t+k}^2}}{u''\left(c_t^{\text{Deliberate}}\left(w_t; \{0\}_{s=1}^{T-t-1}\right)\right) + \delta R^2 \frac{\partial^2 V_{t+1}\left(R\left(w_t - c_t^{\text{Deliberate}}\left(w_t; \{0\}_{s=1}^{T-t-1}\right)\right); \{0\}_{s=1}^{T-t-1}\right)}{\partial w_{t+1}^2}}.$$

$$\frac{\partial^3 V_t\left(w_t;\{0\}_{s=0}^{T-t-1}\right)}{\partial w_t \partial \bar{\lambda}_{t+k}^2} = \frac{\partial^2 c_t^{\text{Deliberate}}\left(w_t;\{0\}_{s=1}^{T-t-1}\right)}{\partial \bar{\lambda}_{t+k}^2}\left[\frac{\partial c_t^{\text{Deliberate}}\left(w_t;\{0\}_{s=1}^{T-t-1}\right)}{\partial w_t}u''\left(c_t^{\text{Deliberate}}\left(w_t;\{0\}_{s=1}^{T-t-1}\right)\right)\right.$$
$$\left.+\left(1-\frac{\partial c_t^{\text{Deliberate}}\left(w_t;\{0\}_{s=1}^{T-t-1}\right)}{\partial w_t}\right)\delta R^2\frac{\partial^2 V_{t+1}\left(R\left(w_t - c_t^{\text{Deliberate}}\left(w_t;\{0\}_{s=1}^{T-t-1}\right)\right);\{0\}_{s=1}^{T-t-1}\right)}{\partial w_{t+1}^2}\right].$$

Using the last expressions and through induction, we have, for all $k \in \{1, \cdots, T-t-1\}$,

$$\frac{\partial^2 c_t^{\text{Deliberate}}\left(w_t;\{0\}_{s=1}^{T-t-1}\right)}{\partial \bar{\lambda}_{t+k}^2} < 0 \quad \text{and} \quad \frac{\partial^3 V_t\left(w_t;\{0\}_{s=0}^{T-t-1}\right)}{\partial w_t \partial \bar{\lambda}_{t+k}^2} > 0.$$

Together with (71), this proves Proposition 6.

Proof of Corollary 6. From (45) and (46), we know the degree of inattention λ_t here corresponds to the degree of mistake in (21) and (32). Corollary 6 then follows from Propositions 3 and 4.

Proof of Corollary 7. From (4) and (49), we know $\lambda = -\frac{\theta}{2+\theta}$. And Corollary 7 follows from Proposition 1.

Proof of Corollary 8. This follows directly from (50) and Proposition 4.

Proof of Proposition 7. As defined in (16), $V_0\left(\bar{w}_0+\Delta;\{\lambda_s\}_{s=0}^{T-1}\right)$ captures the consumer's utility as a function of her mistakes given by (32) or (32) and the realization of the shock Δ. I use $V_0^{\text{Frcitionless}}\left(\bar{w}_0+\Delta\right) \equiv V_0\left(\bar{w}_0+\Delta-L\left(\{\lambda_t\}_{t=0}^{T-1},\Delta\right);\{0\}_{t=0}^{T-1}\right)$ to capture its frictionless counterpart. And $U^{\text{Frictionless}}\left(\bar{w}_0\right) \equiv \mathbb{E}\left[V_0^{\text{Frcitionless}}\left(\bar{w}_0+\Delta\right)\right]$ to capture its expected value over Δ.

I can then calculate the equivalent monetary loss L, i.e., the amount of money the consumer is willing to pay to avoid mistakes in response to saving changes, defined as:

$$\mathbb{E}\left[V_0\left(\bar{w}_0+\Delta;\{\lambda_t\}_{t=0}^{T-1}\right)\right] \equiv \mathbb{E}\left[V_0^{\text{Frcitionless}}\left(\bar{w}_0+\Delta-L\left(\{\lambda_t\}_{t=0}^{T-1}\right)\right)\right] = U^{\text{Frictionless}}\left(\bar{w}_0-L\left(\{\lambda_t\}_{t=0}^{T-1}\right)\right),$$

where $\mathbb{E}[\cdot]$ averages over the realization of Δ.

Note that, in this Proposition, I consider the case that consumption mistakes only come from response to saving changes, so (33) holds and

$$V_0\left(\bar{w}_0;\{\lambda_t\}_{t=0}^{T-1}\right) = V_0^{\text{Frictionless}}\left(\bar{w}_0\right).$$

Let us first consider the quadratic case. In this case, from (56), we know

$$\frac{\partial V_0\left(\bar{w}_0; \{\lambda_t\}_{t=0}^{T-1}\right)}{\partial w_0} = \frac{\partial V_0^{\text{Frictionless}}\left(\bar{w}_0\right)}{\partial w_0} = u'\left(\bar{c}_0\right). \tag{72}$$

As a result,

$$V_0\left(\bar{w}_0 + \Delta; \{\lambda_t\}_{t=0}^{T-1}\right) - V_0^{\text{Frcitionless}}\left(\bar{w}_0 + \Delta\right) = \frac{1}{2}u''\left(\Gamma_0 - \Gamma_0^{\text{Frictionles}}\right)\Delta^2,$$

where Γ_0 and $\Gamma_0^{\text{Frictionles}}$ are defined in (57). From (57), we know

$$\Gamma_0 - \Gamma_0^{\text{Frictionles}} \sim \mathbb{O}^2\left(\{\lambda_t\}_{t=0}^{T-1}\right).$$

As a result,

$$L\left(\{\lambda_t\}_{t=0}^{T-1}\right) = \frac{1}{2}\frac{u''\left(\bar{c}_0\right)}{u'\left(\bar{c}_0\right)}\left(\Gamma_0 - \Gamma_0^{\text{Frictionles}}\right)\mathbb{E}\left[\Delta^2\right]$$

Proposition 7 follows.

Now, we turn to general concave utility case. From (62), we know (62) still holds. So

$$V_0\left(\bar{w}_0 + \Delta; \{\lambda_t\}_{t=0}^{T-1}\right) - V_0^{\text{Frcitionless}}\left(\bar{w}_0 + \Delta\right) = \frac{1}{2}u''\left(\bar{c}_0\right)\left(\Gamma_0 - \Gamma_0^{\text{Frictionles}}\right)\Delta^2 + \mathbb{O}^3\left(\{\lambda_t\}_{t=0}^{T-1}, \Delta\right),$$

where Γ_0 and $\Gamma_0^{\text{Frictionles}}$ are defined in (66). As a result

$$L\left(\{\lambda_t\}_{t=0}^{T-1}\right) = \frac{1}{2}\frac{u''\left(\bar{c}_0\right)}{u'\left(\bar{c}_0\right)}\left(\Gamma_0 - \Gamma_0^{\text{Frictionles}}\right)\mathbb{E}\left[\Delta^2\right] + \mathbb{O}^3\left(\{\lambda_t\}_{t=0}^{T-1}\right) \tag{73}$$

Proposition 7 follows.

Proof of Corollary 9. This case is not directly nested in Propositions 3 or 4, as the actual consumption rule is stochastic. But the proof is essentially unchanged.

The value function in (18) is now given by

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

where $E_t[\cdot]$ averages over the potential realizations of λs. The deliberate consumption in (17) is unchanged.

In the proof of Proposition 3, the deliberate MPC is still given by (54), but (57) becomes

$$\begin{aligned}\Gamma_t &= \int \left[\left(\phi_t^{\text{Deliberate}}(1-\lambda_t)\right)^2 + \left(1-\phi_t^{\text{Deliberate}}(1-\lambda_t)\right)^2 \Gamma_{t+1}\delta R^2\right] d\lambda_t \\ &= \frac{\delta R^2 \Gamma_{t+1}}{1+\delta R^2\Gamma_{t+1}} + \sigma_t^2 \frac{\left(\delta R^2 \Gamma_{t+1}\right)^2}{1+\delta R^2 \Gamma_{t+1}}.\end{aligned}$$

As a result, Γ_t increases with $\left\{\sigma_{t+k}^2\right\}_{k=0}^{T-t-1}$. Corollary 9 then follows directly from (54).

In the proof of Proposition 4, the deliberate MPC is still given by (65), but (66) becomes

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

As a result, Γ_t increases with $\left\{\sigma_{t+k}^2\right\}_{k=0}^{T-t-1}$. Corollary 9 then follows directly from (65).

Proof of Proposition 8. This directly follows from (73), where $\frac{u''(\bar{c}_0)}{u'(\bar{c}_0)} = \frac{-\gamma \bar{c}_0^{-\gamma-1}}{\bar{c}_0^{-\gamma}}$ and $\bar{c}_0$ decreases with $\bar{w}_0$.

Proof of Proposition 9. This is a special case of Corollary 12 in Appendix C.

Appendix B: Additional Results

The Noisy Signal Approach to Inattention.

In the inattention case studied in Corollary 6, each self's perceived permanent income (or wealth) is given by a deterministic weighted average between the actual permanent income (or wealth) and the default. This follows the sparsity approach in Gabaix (2014). An alternative way to model inattention is through noisy signals (Sims, 2003). In fact, with linear consumption rules and Normally distributed fundamentals, the two approaches will lead to similar predictions on MPCs.

Here, I use the quadratic case in Section 4 as an example to illustrate. I assume a Normally distributed exogenous shock, i.e., $\Delta \sim \mathcal{N}(0, \sigma^2)$.[45] Unlike in 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 $\epsilon_t \sim \mathcal{N}\left(0, \sigma_{\epsilon_t}^2\right)$

[45] This together with the linear actual consumption rule in (77) guarantees that each w_t is Normally distributed too.

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.

$$E\left[w_t \mid x_t\right] = (1-\lambda_t)x_t + \lambda_t \bar{w}_t, \tag{74}$$

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

$$c_t\left(x_t\right) = \arg\max_{c_t}\ u\left(c_t\right) + \delta E\left[V_{t+1}\left(R\left(w_t - c_t\right)\right)|x_t\right], \tag{75}$$

where the continuation value function V_{t+1} is defined similarly to the benchmark case, based on future selves' actual consumption rules and potential signals. The deliberate consumption is defined based on the correct permanent income taking future selves' inattention to permanent income as given. We have

Corollary 10. *Each self t's deliberate MPC $\phi_t^{Deliberate}$ is the same as that in Corollary 6, based on $\{\lambda_{t+k}\}$ defined above.*

Proof of Corollary 10. The value in (18) is now given by

$$V_t\left(w_t\right) = \int \left[u\left(c_t\left(w_t+\epsilon_t\right)\right) + \delta V_{t+1}\left(R\left(w_t - c_t\left(w_t+\epsilon_t\right)\right)\right)\right] f_t\left(\epsilon_t\right) d\epsilon_t, \tag{76}$$

where $f_t\left(\cdot\right)$ is the p.d.f. given $\epsilon_t \sim \mathcal{N}\left(0, \sigma_{\epsilon_t}^2\right)$. Similar to (23), I use $\Gamma_t \equiv \frac{\partial^2 V_t(w_t)}{\partial w_t^2}/u'' > 0$ to define the "concavity" of the continuation value function.

The deliberate consumption and MPC is still given by (24) and (54). For the actual consumption in (75), we have

$$\begin{aligned} c_t\left(x_t\right) &= (1-\lambda_t)\phi_t^{\text{Deliberate}}\left(x_t - \bar{w}_t\right) + \bar{c}_t^{\text{Deliberate}}, \\ &= \phi_t\left(x_t - \bar{w}_t\right) + \bar{c}_t^{\text{Deliberate}}. \end{aligned} \tag{77}$$

From (76), we have

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

where $w_{t+1} = R\left(w_t - c_t\left(w_t + \epsilon_t\right)\right)$. The recursive formulation of Γ_t in (57) is then still given by

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

Corollary 10 then follows.

Combined Multiplicative Mistakes.

In some popular behavioral foundations, mistakes in response to saving changes ($\{\lambda_{t+k}\}_{k=1}^{T-t-1}$ in (42)) come together with mistakes in the overall consumption level ($\{\bar{\lambda}_{t+k}\}_{k=1}^{T-t-1}$ in (41)). 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 $k \in \{0, \cdots, T-t\}$,

$$c_{t+k}\left(w_{t+k}\right) = \Phi_{t+k} w_{t+k} \quad \text{and} \quad c_{t+k}^{\text{Deliberate}}\left(w_{t+k}\right) = \Phi_{t+k}^{\text{Deliberate}} w_{t+k}, \tag{78}$$

where, similar to (21), self $t+k$'s mistake Λ_{t+k} is given by

$$\Phi_{t+k} = \left(1-\Lambda_{t+k}\right)\Phi_{t+k}^{\text{Deliberate}}. \tag{79}$$

In the homothetic environment here, future mistake Λ_{t+k} takes a multiplicative as in (44) and plays a dual role. When $\Lambda_{t+k} > 0$, self $t+k$ both under-consumes overall and under-reacts to changes in w_{t+k}. When $\Lambda_{t+k} < 0$, self $t+k$ both over-consumes overall and over-reacts to changes in w_{t+k}. In other words, Λ_{t+k} combines the role of $\{\lambda_{t+k}\}_{k=1}^{T-t-1}$ in (42) and the role of $\{\bar{\lambda}_{t+k}\}_{k=1}^{T-t-1}$ in (41).

I can now study how these "combined" future mistakes $\{\Lambda_{t+k}\}_{k=1}^{T-t-1}$ impact the current consumption. I define $c_t^{\text{Deliberate}}(w_t)$ based on Definition 1 as usual, given actual consumption rules $\{c_{t+k}(w_{t+k})\}_{k=1}^{T-1-t}$ in (78). I can write $c_t^{\text{Deliberate}}(w_t)$ as

$$c_t^{\text{Deliberate}}\left(w_t\right) = \Phi_t^{\text{Deliberate}} w_t,$$

where, in the homothetic environment here, $\Phi_t^{\text{Deliberate}}$ also plays a dual role. It determines both the current MPC and the overall current consumption level. Future mistakes' impact on $\Phi_t^{\text{Deliberate}}$

then combines the high-MPC effect in Proposition 4 and the low-consumption-level effect in 6.

Proposition 10. *(1) When $\gamma < 1$, $\Phi_t^{Deliberate}$ increases with the future mistake $|\Lambda_{t+k}|$ in a neighborhood of $\Lambda_{t+k} = 0$ for each $k \in \{1, \cdots, T-t-1\}$.*

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

When the utility function is not that concave ($\gamma < 1$), the high-MPC channel in Proposition 4, which pushes $\Phi_t^{\text{Deliberate}}$ higher, dominates the precautionary saving channel in Proposition 6, which pushes $\Phi_t^{\text{Deliberate}}$ lower. When the utility function is very concave ($\gamma > 1$), the precautionary saving channel in Proposition 6, which pushes $\Phi_t^{\text{Deliberate}}$ lower, dominates.[46]

Now, let us turn to the plain-vanilla version of hyperbolic discounting without liquidity constraints and commitment devices, e.g., Barro (1999). Specifically, the actual consumption rule is given by

$$c_t(w_t) = \arg\max_{c_t} u(c_t) + \delta\beta_t V_{t+1}(R(w_t - c_t)) \qquad \forall t \in \{0, \cdots, T-1\}, \tag{80}$$

where $\delta \in [0,1]$ is the standard discount factor, $\beta_t \in [0,1]$ captures self t's present bias (a smaller β_t means a larger present bias), and $V_{t+1}(\cdot)$ is the continuation value function defined as in (16). The consumer is not subject to the borrowing constraint, as in the result of the paper.

The actual consumption rule in (80) combines the direct effect of present bias on current consumption with the effect of anticipated future present bias (sophistication). To isolate the impact of future present bias on current consumption, I define the deliberate consumption rule given future selves' present bias $\{\beta_{t+k}\}_{k=1}^{T-t-1}$ as usual:

$$c_t^{\text{Deliberate}}(w_t) = \arg\max_{c_t} u(c_t) + \delta V_{t+1}(R(w_t - c_t)) \qquad \forall t \in \{0, \cdots, T-1\}. \tag{81}$$

This case can be mapped to the homothetic environment of Proposition 10. As its Corollary:

Corollary 11. *When $u(x) = v(x) = \frac{x^{1-\gamma}}{1-\gamma}$, the hyperbolic discounting case in (80) and (81) is nested by Proposition 10. When $\gamma < 1$, the current MPC $\phi_t^{Deliberate}$ increases with future selves' present bias, i.e., decreases with each $\{\beta_{t+k}\}_{k=1}^{T-t-1}$.*

This is consistent with the result in Maxted (2020) about how sophistication impacts MPCs with presently-biased agents.

[46] One may wonder how to reconcile Proposition 10 with Figure 3, where the precautionary saving motive does not matter much for the MPC. Note that, in Figure 3, as the rest of the paper, mistakes in overall consumption level take the form of an "additive" deviation from the deliberate counterpart, similar to (41). Figure 3 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 (79) take a multiplicative form. It leads to large deviations from the deliberation counterpart and large precautionary saving motives in Proposition 10.

Proof of Proposition 10. I guess and verify the continuation value function defined in (18) takes the form of

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

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 (17), the deliberate consumption is given by

$$\left(c_t^{\text{Deliberate}}(w_t)\right)^{-\gamma} = \delta R \kappa_{t+1} \left(R\left(w_t - c_t^{\text{Deliberate}}(w_t)\right)\right)^{-\gamma}$$

$$\Phi_t^{\text{Deliberate}} = \frac{(\delta\kappa_{t+1})^{-\frac{1}{\gamma}} (R)^{1-\frac{1}{\gamma}}}{1 + (\delta\kappa_{t+1})^{-\frac{1}{\gamma}} (R)^{1-\frac{1}{\gamma}}} \tag{82}$$

From (79), 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 (17), 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

$$\phi^{\text{Deliberate}} = \frac{(\delta\kappa)^{-\frac{1}{\gamma}} (R)^{1-\frac{1}{\gamma}}}{1 + (\delta\kappa)^{-\frac{1}{\gamma}} (R)^{1-\frac{1}{\gamma}}}.$$

Moreover,

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

We have

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

So

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

Moreover,

$$\frac{\partial f(0,\kappa)}{\partial\kappa} = \delta R^{1-\gamma}\left(1-\phi^{\text{Deliberate}}\right)^{1-\gamma} > 0.$$

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 (82), we arrive at Proposition 10.

Proof of Corollary 11. From (80) and (81), we have

$$u'(c_t(w_t)) = \delta\beta_t R V'_{t+1}(R(w_t - c_t(w_t))), \tag{83}$$

and

$$u'\left(c_t^{\text{Deliberate}}(w_t)\right) = \delta R V'_{t+1}\left(R\left(w_t - c_t^{\text{Deliberate}}(w_t)\right)\right). \tag{84}$$

Comparing (83) and (84), we have:

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

Corollary 11 then follows directly from 10.

Appendix C: The General Case Allowing Non-fungibility

I now turn to the general, non-fungible case, in which mistakes in future consumption may also include inefficiently differential responses to different components of permanent income. In this general case, I first show that the main high-MPC result remains true: as long as future consumption responds inefficiently to saving changes, current MPCs are higher. Then, I show that the non-fungibility of future consumption by itself suffices to generate non-fungibility of current consumption. That is, even if the current self understands how to calculate permanent income correctly, as long as she anticipates future consumption mistakes in the form of future non-fungibility, she will also respond differentially to changes in different components of permanent income. In this sense, mistakes in future consumption beget current non-fungibility. A empirically useful non-fungibility is that future mistakes in response to saving changes generate an excess discounting of future income. This is consistent with empirical studies which find limited consumption responses to news about future income.

The Environment

In Section 3, I restrict the actual consumption to be a function of permanent income: $w_t = a_t + y_t + \sum_{k=1}^{T-t} R^{-k} y_{t+k}$. Here, I allow actual consumption to respond to different components of permanent income differently. In other words, mistakes in actual consumption rules may include inefficiently differential responses to different components of permanent income. For simplicity I consider the quadratic-linear case studied in Section 4. But the result in this section can be easily extended to the general concave utility case studied in Proposition 4.

Specifically, I explicitly write the consumer's budget (13) as (31), where y_s is her income at period s and a_s is her saving/borrowing at the start of period s. I use s_t to capture the exogenous income state at period t summarizing information about current income y_t and future incomes $\{y_{t+k}\}_{k\geq 1}$. For illustration purposes and follow the main analysis, I assume that all income uncertainty in the economy is resolved in period 0, so $s_t = (y_t, \cdots, y_T)$.

The actual consumption rule of each self $t \in \{0, \cdots, T-1\}$ is given by:

$$c_t(a_t, s_t) = \phi_t^a a_t + \phi_t^y \left(y_t + \sum_{k=1}^{T-t} \omega_{t,k} R^{-k} y_{t+k} \right) + \widehat{c}_t, \tag{85}$$

where ϕ_t^a captures the actual MPC out of wealth (i.e. saving/borrowing), ϕ_t^y captures the actual MPC out of current income, $\phi_t^y \omega_{t,k}$ captures the actual MPC out of future income k periods later, and $\omega_{t,k}$ captures how this MPC violates the fungibility principle. For example, when $\omega_{t,k} < 1$, the consumer excessively discounts future income k periods later. Finally, $\widehat{c}_t$ in (85) is an exogenous constant capturing the level of self t's actual consumption, whose value does not influence the deliberate MPCs calculated below.

The actual consumption rule in (85) allows differential mistakes in response to different components of permanent income. I use $\lambda_t = \left(\lambda_t^a,\ \left\{\lambda_{t,k}^y\right\}_{k=0}^{T-t}\right)$ to capture self t's mistakes, i.e., how the actual MPCs in (85) deviate from the deliberate MPCs $\phi_t^{\text{Deliberate}}$ and $\left\{\phi_t^{\text{Deliberate}}\omega_{t,k}^{\text{Deliberate}}\right\}_{k=0}^{T-t}$ introduced below in (87). Specifically, for all $t \in \{0, \cdots, T-1\}$ and $k \in \{0, \cdots, T-t\}$,

$$\phi_t^a = (1-\lambda_t^a)\,\phi_t^{\text{Deliberate}}, \quad \phi_t^y = \left(1-\lambda_{t,0}^y\right)\phi_t^{\text{Deliberate}}, \quad \text{and} \quad \phi_t^y\omega_{t,k} = \left(1-\lambda_{t,k}^y\right)\phi_t^{\text{Deliberate}}\omega_{t,k}^{\text{Deliberate}}, \tag{86}$$

where λ_t^a captures the mistake in self t's actual MPC out of wealth (i.e. saving/borrowing), $\lambda_{t,0}^y$ captures the mistake in self t's actual MPC out of current income, and $\lambda_{t,k}^y$ captures the mistake in self t's actual MPC out of future income $k \geq 1$ periods later. Similar to (21), a positive λ means under-reaction and a negative λ means over-reaction. As in Section 4, the mistakes λ_t^a and $\left\{\lambda_{t,k}^y\right\}_{k=0}^{T-t}$ are treated as exogenous now but I will connect them to the exact underlying behavioral biases below.

The fungibility case analyzed in Section 4 is nested here by $\lambda_t = \lambda_t^a = \lambda_{t,k}^y$, for all t and $k \in \{0, \cdots, T-t\}$. That is, the fungibility case analyzed above is a special case in which mistakes in response to different components of permanent income are the same.

Similar to Definition 1 and based on future selves' actual consumption rules $\{c_{t+k}(a_{t+k}, s_{t+k})\}_{k=0}^{T-t-1}$ above, each self t's deliberate consumption rule will take the following form.

Lemma 2. *For $t \in \{0, \cdots, T-1\}$, each self t's deliberate consumption rule is given by:*

$$c_t^{\textit{Deliberate}}(a_t, s_t) = \phi_t^{\textit{Deliberate}}\left(a_t + y_t + \sum_{k=1}^{T-t}\omega_{t,k}^{\textit{Deliberate}}R^{-k}y_{t+k}\right) + \hat{c}_t^{\textit{Deliberate}}, \tag{87}$$

where $\hat{c}_t^{\textit{Deliberate}}$ is a scalar, $\phi_t^{\textit{Deliberate}}$ is a function of $\left(\left\{\lambda_{t+l}^a\right\}_{l=1}^{T-t-1},\ \delta,\ R\right)$, and $\omega_{t,k}$ is a function of $\left(\left\{\lambda_{t+l}^a\right\}_{l=1}^{T-t-1},\ \left\{\lambda_{t+l,k-l}^y\right\}_{l=1}^{\min\{k,\ T-t-1\}},\ \delta,\ R\right)$.

In (87), $\phi_t^{\text{Deliberate}}$ captures the MPC of deliberate consumption out of current income and wealth, $\phi_t^{\text{Deliberate}}\omega_{t,k}^{\text{Deliberate}}$ captures the deliberate MPC out of future income k periods later, $\omega_{t,k}^{\text{Deliberate}}$ captures how this MPC violates the fungibility principle, and $\bar{c}_t^{\text{Deliberate}}$ captures the overall level of self t's deliberate consumption. It is worth noting that $\omega_{t,k}$ is a function of $\left\{\lambda_{t+l,k-l}^y\right\}$ (but not other λ^ss) because $\omega_{t,k}$ is about self t's response to future income y_{t+k} and the relevant future mistakes are $\left\{\lambda_{t+l,k-l}^y\right\}$, i.e., how the future self $t+l$ responds to income y_{t+k}.

In this Section, I establish two general results about how future consumption mistakes impact current MPCs. First, the above high MPCs result still holds: as long as future consumption responds inefficiently to changes in saving/borrowing $\left(\lambda^a_{t+l} \neq 0\right)$, current deliberate MPCs, i.e., $\phi^{\text{Deliberate}}_t$ in (87), will be higher. Second, non-fungibility of future consumption ($\lambda^a_{t+l} \neq \lambda^y_{t+l,k-l}$) suffices to generate the non-fungibility of current deliberate consumption ($\omega^{\text{Deliberate}}_{t,k} \neq 1$). In other words, even if the current self knows how to calculate permanent income correctly, as long as she anticipates future consumption mistakes in the form of future non-fungibility, she will violate the fungibility principle and respond differentially to changes in different components of permanent income.

High Current MPCs

Here, I show that the main results in Section 4, i.e., how future consumption mistakes lead to excess concavity of the continuation value function and high current MPCs, still hold. I further emphasize that the key behind this result is the inefficient responses of future consumption to changes in saving/borrowing.

Similar to Lemma 1, I use $\Gamma_{t+1} > 0$ to denote the "concavity" of the consumer's continuation value function in (16): $\frac{\partial^2 V_{t+1}(a_{t+1}, s_{t+1})}{\partial a^2_{t+1}} \equiv u'' \cdot \Gamma_{t+1}$.

Proposition 11.

i. Excess concavity of the continuation value function: for $t \in \{0, \cdots, T-2\}$, Γ_{t+1} *strictly increases with* $\left\{\left|\lambda^a_{t+l}\right|\right\}^{T-t-1}_{l=1}$.

ii. High current MPCs: for $t \in \{0, \cdots, T-2\}$, $\phi^{\text{Deliberate}}_t \geq \phi^{\text{Frictionless}}_t$ *and* $\phi^{\text{Deliberate}}_t$ *increases with each of* $\left\{\left|\lambda^a_{t+l}\right|\right\}^{T-t-1}_{l=1}$.

The intuition behind part (i) of Proposition 11 is similar to Lemma 1. Larger $\left\{\left|\lambda^a_{t+l}\right|\right\}^{T-t-1}_{l=0}$ means more inefficient future consumption responses to changes in saving/borrowing. As a result, the marginal value of saving $\frac{\partial V_{t+1}(a_{t+1}, s_{t+1})}{\partial a_{t+1}}$ decreases faster with a_{t+1} and the continuation value function V_{t+1} becomes more concave. It is worth noting that, here, the relevant mistakes $\left\{\lambda^a_{t+l}\right\}^{T-t-1}_{l=0}$ are inefficient responses of future consumption to changes in saving/borrowing. This is because these responses directly determine the marginal value of saving $\frac{\partial V_{t+1}(a_{t+1}, s_{t+1})}{\partial a_{t+1}}$ and hence the concavity Γ_t. On the other hand, Γ_{t+1} is independent of $\lambda^y_{t+l,k-l}$ for all l and k.

The intuition behind part (ii) of Proposition 11 is similar to Proposition 3. From part (i), with future consumption mistakes (larger$\left\{\left|\lambda^a_{t+l}\right|\right\}^{T-t-1}_{l=1}$), the continuation value function becomes more concave. As a result, in response to changes in current income, the current self is more willing to adjust her current consumption instead of her saving. She hence displays a higher MPC.

Similar to part (i), the relevant mistakes $\left\{\lambda^a_{t+l}\right\}^{T-t-1}_{l=1}$ for the high current MPCs result are future selves' inefficient responses to changes in saving/borrowing. On the other hand, $\phi^{\text{Deliberate}}_t$ is independent of $\lambda^y_{t+l,k-l}$ for all l and k. This result has an independent use: for a behavioral bias causing inefficiently differential responses of future consumption to different components of permanent income, it helps predict whether anticipation of this behavioral bias contributes to high current MPCs.

Future Non-fungibility Begets Current Non-fungibility

Now, I turn to a new prediction.

Proposition 12. *Generically, the deliberate consumption in (87) violates the fungibility principle. That is, for* $t \in \{0, \cdots, T-2\}$ *and* $k \in \{0, \cdots, T-t\}$, *generically,* $\omega_{t,k}^{Deliberate} \neq 1$. *Here, generically is in the sense of the Euclidean measure of the product space generated by future selves' mistakes* $\left(\left\{\lambda_{t+l}^{a}\right\}_{l=1}^{T-t-1}, \left\{\lambda_{t+l,k-l}^{y}\right\}_{l=1}^{\min\{k,\ T-t-1\}}\right)$.

This result means that the inefficient differential responses of future consumption to different components of permanent income, by themselves, suffice to generate the non-fungibility of the current consumption. Even if the current self is not subject to any behavioral mistakes, her consumption endogenously responds differentially to changes in different components of permanent income.

In other words, the fungibility case studied in Section 4 is rather special. There, future actual consumption exhibits the same degree of mistakes in responses to changes in different components of permanent income,

$$\lambda_{t+l} = \lambda_{t+l}^{a} = \lambda_{t+l,k-l}^{y} \quad \forall l, k. \tag{88}$$

In this case, the current deliberate consumption remains to follow the fungibility principle. Away from (88), generically, current deliberate consumption will violate the fungibility principle.

Excess discounting.

To better understand the intuition behind Proposition 12, here I study an empirically relevant case of how future selves violate fungibility: mistakes in future actual consumption take the form of an smaller MPC out of wealth than out of income, i.e., $\lambda_{t+l}^{a} \geq \lambda_{t+l,k-l}^{y}$ for all $l \in \{1, \cdots, T-t-1\}$ and $k \in \{l, \cdots, T-t+l\}$ (recall a larger λ means a smaller MPC). This case is consistent with the empirical evidence on smaller MPCs out of financial wealth in Thaler (1990), Baker, Nagel and Wurgler (2007), Paiella and Pistaferri (2017), Di Maggio, Kermani and Majlesi (2018), and Fagereng et al. (2019).

Proposition 13. *Consider the case that* $\lambda_{t+l}^{a} \geq \lambda_{t+l,k-l}^{y}$ *and* $\lambda_{t+l}^{a} \geq 0$ *for all* $l \in \{1, \cdots, T-t-1\}$ *and* $k \in \{l, \cdots, T-t+l\}$.

The current deliberate consumption in (87) has the following properties: for $k \in \{0, \cdots, T-t\}$, *(i)* $\omega_{t,k}^{Deliberate} \leq 1$. *That is, the current self excessively discounts future income.*

(ii) $\omega_{t,k}^{Deliberate}$ *decreases with each* $\left\{\lambda_{t+l}^{a}\right\}_{l=1}^{T-t-1}$ *(i.e., increases with future selves' actual MPCs out of wealth) and increases with each* $\left\{\lambda_{t+l,k-l}^{y}\right\}_{l=1}^{\min\{k,\ T-t-1\}}$ *(i.e., decreases with future selves' actual MPCs out of income).*

(iii) $\omega_{t,k}^{Deliberate} \leq \omega_{t+1,k-1}^{Deliberate} \leq \cdots \leq \omega_{t+k-1,1}^{Deliberate} \leq 1$.

Proposition 13 means that, if the non-fungibility of future actual consumption takes the form of inefficiently small MPCs out of wealth, the current self exhibits excess discounting of future income.

To understand the intuition behind Proposition 13, note that, when future selves mistakenly respond too little to changes in saving/wealth (a larger $\left\{\lambda_{t+l}^{a}\right\}_{l=1}^{T-t-1}$), the excess concavity in Proposition 12 means that the current self will be less willing to change her saving. As a result, the current self is less willing to adjust her current consumption in response to changes in future income, since the response of current consumption to future income requires changes in saving. Hence, there is excess discounting ($\omega_{t,k}^{\text{Deliberate}} < 1$) and $\omega_{t,k}^{\text{Deliberate}}$ decreases with each $\left\{\lambda_{t+l}^{a}\right\}_{l=1}^{T-t-1}$.

On the other hand, $\omega_{t,k}^{\text{Deliberate}}$ increases with each $\left\{\lambda_{t+k,l-k}^{y}\right\}_{k=1}^{\min\{l,\ T-t-1\}}$. That is, if future selves' mistakenly respond too little to changes in future income y_{t+k} (a larger $\left\{\lambda_{t+k,l-k}^{y}\right\}_{k=1}^{\min\{l,\ T-t-1\}}$), the current self will be more willing to respond to y_{t+k}. In other words, there is essentially some "substitution" across different selves in response to future income.

In the empirically relevant case here that future consumption responds less to wealth than to income, the first channel dominates and the current self exhibits excess discounting of future income.

Part (iii) of Proposition 13 further establishes a "distance effect." The consumer's response to changes in future income, y_{t+k}, exhibits more discounting when the period $t + k$ is further away. This is because the mechanism behind excess discounting accumulates over the distance between current consumption and future income.

Consistent with excess discounting of future income, empirical studies find limited consumption responses to news about future income, i.e., a very limited "announcement effect." Papers documenting this pattern away from liquidity constraints include Stephens and Unayama (2011), Parker (2017), Olafsson and Pagel (2018), and Kueng (2018).[47]

The $T \to \infty$ and hand-to-mouth limit. Similar to Corollary 1, I can establish a simple limit for the deliberate consumption rule in (87) when the consumer's horizon T goes to infinity.

[47] For the potentially empirically irrelevant case that mistakes in future consumption take the form of inefficiently large MPCs out of wealth, the main lesson in Proposition 12 remains true: the non-fungibility of future consumption leads to non-fungibility of current consumption. In this case, $\omega_{t,k}^{\text{Deliberate}}$ can be larger than 1. In fact, this is consistent with the intuition behind the comparative statics in part (ii) of Proposition 13.

Corollary 12. *Let* $\lambda^a_{t+l} = \lambda^a$ *with* $|\lambda^a| < \left(\delta^{-1/2}R^{-1}\right)$ *and* $\lambda^y_{t+l,k-l} = \lambda^y$ *for all* k *and* l*. When* $T \to +\infty$,

$$\phi_t^{Deliberate} \to \phi^{Deliberate} \equiv \frac{\delta R^2 - 1}{\delta R^2 \left(1 - (\lambda^a)^2\right)}, \tag{89}$$

$$\omega_{t,k}^{Deliberate} \to \left(\omega^{Deliberate}\right)^k \equiv \left(1 - \frac{\left(\delta R^2 - 1\right)\lambda^a\left(\lambda^a - \lambda^y\right)}{1 - (\lambda^a)^2}\right)^k.$$

Furthermore, when $\lambda^a \to \left(\delta^{-1/2}R^{-1}\right)^-$ *and* $\lambda^y \to 0$*,*

$$\phi^{Deliberate} \to 1 \quad \textit{and} \quad \omega^{Deliberate} \to 0. \tag{90}$$

The limit in (90) is effectively a "hand-to-mouth" limit. When the current self is very worried about the mistaken responses of future consumption to changes in savings, she becomes unwilling to change her savings. As a result, she does not respond to changes in future income and absorbs all changes in current income. In other words, she is effectively "hand-to-mouth" with respect to *changes* in income, even though her consumption level does not need to track the current income level ($c_t \neq y_t$).

This simple "hand-to-mouth" limit also illustrates how my mechanism can explain the empirical evidence on excess sensitivity to *anticipated* income shocks away from liquidity constraint (e.g. Kueng, 2018). In this limit, consumption does not respond to future income until it arrives. At that point, consumption fully absorbs the anticipated income shock.

A smaller effect of interest rate changes.

Another famous puzzle in intertemporal consumption is the empirical evidence on the weak intertemporal substitution motive and the small response of consumption to interest rate changes (Hall, 1988; Campbell and Mankiw, 1989; Havránek, 2015). My proposed channel, i.e., the impact of future consumption mistakes in response to saving changes, can also help resolve this puzzle.

The intuition is similar: the response of current consumption to interest rate changes leads to saving changes; with future consumption mistakes in response to saving changes, the current self is less willing to respond to interest rate changes.

To formalize this, I study responses to changes in the interest rate between period t and $t+1$, R_t. To isolate the intertemporal substitution motive, I study deviations away from a frictionless path with zero net saving at the end of period t.[48]

[48] The zero net saving condition guarantees that the response to interest rate changes is driven by the intertemporal saving motive. Away from this restriction, interest rate changes may also have income effects on consumption. Future consumption mistakes may amplify the income effect of interest rates on consumption, similar to the main high

Proposition 14. *The response of deliberate consumption to interest rate changes,* $\left|\frac{\partial c_t^{Deliberate}}{\partial R_t}\right|$, *decreases with each future self's mistake* $\{|\lambda_{t+k}|\}_{k=1}^{T-t-1}$.

Proof of Lemma 2. Similar to (23), we define $\left\{\Gamma_t, \Gamma_{t,k}^y\right\}_{t\in\{0,\cdots,T\},k\in\{0,\cdots,T-t\}}$ based on

$$\frac{\partial V_t}{\partial a_t} \equiv u'' \cdot \left(\Gamma_t a_t + \sum_{k=0}^{T-t} \Gamma_{t,k}^y R^{-k} y_{t+k} + \hat{\Gamma}_t\right). \tag{91}$$

To prove Lemma 2, we work with backward induction. At T, we have:

$$\Gamma_T = \Gamma_{T,0}^y = \frac{v''}{u''} > 0.$$

For each $t \leq T-1$, similar to (26), the deliberate consumption is given by

$$u'\left(c_t^{\text{Deliberate}}\left(a_t, s_t\right)\right) = R\delta \frac{\partial V_{t+1}}{\partial a_{t+1}} \left(R\left(a_t + y_t - c_t^{\text{Deliberate}}\left(a_t, s_t\right)\right), s_{t+1}\right).$$

Together (91) at $t+1$, we have

$$c_t^{\text{Deliberate}}\left(a_t, s_t\right) = \phi_t^{\text{Deliberate}} \left(a_t + y_t + \sum_{k=1}^{T-t} \omega_{t,k}^{\text{Deliberate}} R^{-k} y_{t+k}\right) + \hat{c}_t^{\text{Deliberate}},$$

with

$$\phi_t^{\text{Deliberate}} = \frac{\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}} \tag{92}$$

and for $\forall k \in \{1, \cdots, T-t\}$,

$$\omega_{t,k}^{\text{Deliberate}} = \frac{\delta R R^{-(k-1)} \Gamma_{t+1,k-1}^y}{1 + \Gamma_{t+1} \delta R^2} / \left(\phi_t^{\text{Deliberate}} R^{-k}\right) = \frac{\Gamma_{t+1,k-1}^y}{\Gamma_{t+1}}. \tag{93}$$

Now, from the recursive formulation of the value function similar to (17), we have:

$$\frac{\partial V_t\left(a_t, s_t\right)}{\partial a_t} = \phi_t^a u'\left(c_t\left(a_t, s_t\right)\right) + \left(1 - \phi_t^a\right) \delta R \frac{\partial V_{t+1}\left(a_{t+1}, s_{t+1}\right)}{\partial a_{t+1}}. \tag{94}$$

MPCs result in response to income changes.

Together with the budget constraint $a_{t+1} = R\left(a_t + y_t - c_t\right)$, we have:

$$\Gamma_t a_t + \sum_{k=0}^{T-t} \Gamma^y_{t,k} R^{-k} y_{t+k} + \hat{\Gamma}_t = \left(\phi^a_t - \left(1-\phi^a_t\right)\delta R^2 \Gamma_{t+1}\right)\left(\phi^a_t a_t + \phi^y_t \left(y_t + \sum_{k=1}^{T-t} \omega_{t,k} R^{-k} y_{t+k}\right) + \hat{c}_t\right)$$
$$+ \left(1-\phi^a_t\right)\delta R \left(\Gamma_{t+1} R\left(a_t + y_t\right) + \sum_{k=0}^{T-t-1} \Gamma^y_{t+1,k} R^{-k} y_{t+1+k} + \hat{\Gamma}_{t+1}\right).$$

Together with (86), we have, for all $t \in \{0, \cdots, T-1\}$:

$$\begin{aligned}\Gamma_t &= \phi^a_t \left(\phi^a_t - \left(1-\phi^a_t\right)\delta R^2 \Gamma_{t+1}\right) + \left(1-\phi^a_t\right)\delta R^2 \Gamma_{t+1} \\ &= \left(1+\delta R^2 \Gamma_{t+1}\right)\left(\phi^a_t - \frac{\delta R^2 \Gamma_{t+1}}{1+\delta R^2 \Gamma_{t+1}}\right)^2 + \frac{\delta R^2 \Gamma_{t+1}}{1+\delta R^2 \Gamma_{t+1}} \\ &= \frac{\left(\delta R^2 \Gamma_{t+1}\right)^2}{1+\delta R^2 \Gamma_{t+1}}\left(\lambda^a_t\right)^2 + \frac{\delta R^2 \Gamma_{t+1}}{1+\delta R^2 \Gamma_{t+1}},\end{aligned} \tag{95}$$

and

$$\begin{aligned}\Gamma^y_{t,0} &= \phi^y_t \left(\phi^a_t - \left(1-\phi^a_t\right)\delta R^2 \Gamma_{t+1}\right) + \left(1-\phi^a_t\right)\delta R^2 \Gamma_{t+1} \\ &= \left(1+\delta R^2 \Gamma_{t+1}\right)\left(\phi^a_t - \frac{\delta R^2 \Gamma_{t+1}}{1+\delta R^2 \Gamma_{t+1}}\right)\left(\phi^y_t - \frac{\beta R^2 \Gamma_{t+1}}{1+\beta R^2 \Gamma_{t+1}}\right) + \frac{\delta R^2 \Gamma_{t+1}}{1+\delta R^2 \Gamma_{t+1}} \\ &= \frac{\left(\delta R^2 \Gamma_{t+1}\right)^2}{1+\delta R^2 \Gamma_{t+1}}\lambda^a_t \lambda^y_{t,0} + \frac{\delta R^2 \Gamma_{t+1}}{1+\delta R^2 \Gamma_{t+1}},\end{aligned} \tag{96}$$

and for $k \in \{1, \cdots, T-t\}$:

$$\begin{aligned}\Gamma^y_{t,k} &= \phi^y_t \omega_{t,k} \left(\phi^a_t - \left(1-\phi^a_t\right)\delta R^2 \Gamma_{t+1}\right) + \left(1-\phi^a_t\right)\delta R^2 \Gamma^y_{t+1,k-1} \\ &= \left(1+\delta R^2 \Gamma_{t+1}\right)\left(\phi^a_t - \frac{\delta R^2 \Gamma_{t+1}}{1+\delta R^2 \Gamma_{t+1}}\right)\left(\phi^y_t \omega_{t,k} - \frac{\delta R^2 \Gamma^y_{t+1,k-1}}{1+\delta R^2 \Gamma_{t+1}}\right) + \frac{\delta R^2 \Gamma^y_{t+1,k-1}}{1+\delta R^2 \Gamma_{t+1}} \\ &= \frac{\left(\delta R^2\right)^2 \Gamma_{t+1}\Gamma^y_{t+1,k-1}}{1+\delta R^2 \Gamma_{t+1}}\lambda^a_t \lambda^y_{t,k} + \frac{\delta R^2 \Gamma^y_{t+1,k-1}}{1+\delta R^2 \Gamma_{t+1}}.\end{aligned} \tag{97}$$

Lemma 2 follows from (92) – (97).

Proof of Proposition 11. From Lemma 2, we know the expressions for ϕ^a_t, $\phi^{\text{Deliberate}}_t$, and Γ_t here are identical to those in Lemma 1 and Proposition 3, with $\{\phi^a_t\}_{t=0}^{T-1}$ replacing the role of $\{\phi_t\}_{t=0}^{T-1}$ and $\{\lambda^a_t\}_{t=0}^{T-1}$ replacing the role of $\{\lambda_t\}_{t=0}^{T-1}$. Proposition 11 then follows directly from Lemma 1 and Proposition 3.

Proof of Proposition 12. From (93), (95), and (96), for $t \in \{0, \cdots, T-2\}$, we have

$$\omega_{t,1}^{\text{Deliberate}} = \frac{\delta R^2 \Gamma_{t+2} \lambda_{t+1}^a \lambda_{t+1,0}^y + 1}{\delta R^2 \Gamma_{t+2} \left(\lambda_{t+1}^a\right)^2 + 1} = 1 - \frac{\delta R^2 \Gamma_{t+2} \lambda_{t+1}^a \left(\lambda_{t+1}^a - \lambda_{t+1,0}^y\right)}{\delta R^2 \Gamma_{t+2} \left(\lambda_{t+1}^a\right)^2 + 1}, \tag{98}$$

and $\omega_{T-1,1}^{\text{Deliberate}} = 1$.

From (93), (95), and (97), for $t \in \{0, \cdots, T-2\}$ and $k \in \{2, \cdots, T-t\}$, we have

$$\begin{aligned} \omega_{t,k}^{\text{Deliberate}} &= \frac{\Gamma_{t+1,k-1}^y}{\Gamma_{t+1}} = \frac{\delta R^2 \Gamma_{t+2} \lambda_{t+1}^a \lambda_{t+1,k-1}^y + 1}{\delta R^2 \Gamma_{t+2} \left(\lambda_{t+1}^a\right)^2 + 1} \frac{\Gamma_{t+2,k-2}^y}{\Gamma_{t+2}} \\ &= \left[1 - \frac{\delta R^2 \Gamma_{t+2} \lambda_{t+1}^a \left(\lambda_{t+1}^a - \lambda_{t+1,k-1}^y\right)}{\delta R^2 \Gamma_{t+2} \left(\lambda_{t+1}^a\right)^2 + 1}\right] \omega_{t+1,k-1}^{\text{Deliberate}}. \end{aligned} \tag{99}$$

Together, we know, generically, $\omega_{t,k}^{\text{Deliberate}} \neq 1$. Here, generically is in the sense of the Euclidean measure of the product space generated by future selves' mistakes $\left(\left\{\lambda_{t+l}^a\right\}_{l=1}^{T-t-1}, \left\{\lambda_{t+l,k-l}^y\right\}_{l=1}^{\min\{k,\ T-t-1\}}\right)$.

Proof of Proposition 13. Consider the case that $\lambda_{t+l}^a \geq \lambda_{t+l,k-l}^y$ and $\lambda_{t+l}^a \geq 0$ for all $l \in \{1, \cdots, T-t-1\}$ and $k \in \{l, \cdots, T-t+l\}$.

(i) This comes directly from (98) and (99).

(ii) The comparative statics with respect to $\left\{\lambda_{t+l,k-l}^y\right\}_{l=1}^{\min\{k,\ T-t-1\}}$ come directly from (98) and (99). To prove comparative statics with respect to $\left\{\lambda_{t+l}^a\right\}_{l=1}^{T-t-1}$, define:

$$f\left(\Gamma, \lambda^y, \lambda^a\right) \equiv \frac{\delta R^2 \Gamma \lambda^y \lambda^a + 1}{\delta R^2 \Gamma \left(\lambda^a\right)^2 + 1}.$$

We have

$$\begin{aligned} \frac{\partial f}{\partial \lambda^a}\left(\Gamma, \lambda^y, \lambda^a\right) &= \frac{\delta R^2 \Gamma \lambda^y}{\delta R^2 \Gamma \left(\lambda^a\right)^2 + 1} - \frac{2\delta R^2 \Gamma \lambda^a \left(\delta R^2 \Gamma \lambda^y \lambda^a + 1\right)}{\left(\delta R^2 \Gamma \left(\lambda^a\right)^2 + 1\right)^2} \\ &= \frac{\delta R^2 \Gamma}{\delta R^2 \Gamma \left(\lambda^a\right)^2 + 1} \left(\lambda^y - \frac{2\lambda^a \left(\delta R^2 \Gamma \lambda^y \lambda^a + 1\right)}{\delta R^2 \Gamma \left(\lambda^a\right)^2 + 1}\right) \\ &= \frac{\delta R^2 \Gamma \lambda^y}{\delta R^2 \Gamma \left(\lambda^a\right)^2 + 1} \left(\frac{\lambda^y - \lambda^y \delta R^2 \Gamma \left(\lambda^a\right)^2 - 2\lambda^a}{\delta R^2 \Gamma \left(\lambda^a\right)^2 + 1}\right). \end{aligned}$$

As a result, $\frac{\partial f}{\partial \lambda^a}\left(\Gamma, \lambda^y, \lambda^a\right) \leq 0$ if $\lambda^a \geq \lambda^y$ and $\lambda^a \geq 0$. Applying this result in (98) and (99), we know $\omega_{t,k}^{\text{Deliberate}}$ decreases with each $\left\{\lambda_{t+l}^a\right\}_{l=1}^{T-t-1}$.

(iii) This comes directly (99).

Proof of Proposition 14. Consider the non-fungible environment studied here. As mentioned in the main text, I fixed a t and study responses to changes in the interest rate between period t and $t+1$, R_t. To isolate the intertemporal substitution motive, I study deviations away from a frictionless pre-shock path $\{\bar{a}_h, \bar{c}_h, \bar{y}_h\}_{h=0}^{T-1}$, with zero net saving at the end of period t, i.e., $\bar{a}_{t+1} = 0$. On this path, similar to (33), actual consumption coincides with the deliberate consumption $\bar{c}_t = c_t(\bar{a}_t, \bar{s}_t) = c_t^{\text{Delibrate}}(\bar{a}_t, \bar{s}_t)$.

Since interest rates are fixed from $t+1$, the continuation value function is still given by $V_{t+1}(a_{t+1}, s_{t+1})$ defined in (16). Self t's deliberate consumption is given by

$$u'\left(c_t^{\text{Deliberate}}(a_t, s_t, R_t)\right) = \delta R_t \frac{\partial V_{t+1}(a_{t+1}, s_{t+1})}{\partial a_{t+1}},$$

where $a_{t+1} = R_t\left(a_t + y_t - c_t^{\text{Deliberate}}(a_t, s_t, R_t)\right)$. Take a derivative with respect to R_t and evaluated at $(\bar{a}_t, \bar{s}_t, R)$, we have

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

where I use $\bar{a}_{t+1} = R(\bar{a}_t + \bar{y}_t - \bar{c}_t) = 0$. As a result,

$$\frac{\partial c_t^{\text{Deliberate}}(\bar{a}_t, \bar{s}_t, R)}{\partial R_t} = \frac{\delta u'(\bar{c}_{t+1})}{u''(1 + \delta R^2 \Gamma_{t+1})},$$

where I use $\frac{\partial V_{t+1}(\bar{a}_{t+1}, \bar{s}_{t+1})}{\partial a_{t+1}} = u'(\bar{c}_{t+1})$ on the frictionless path[49] and $\Gamma_{t+1} \equiv \frac{\partial^2 V_{t+1}(\bar{a}_{t+1}, \bar{s}_{t+1})}{\partial a_{t+1}^2}/u''$ is given by Proposition 11. Proposition 14 then follows from Proposition 11.

Proof of Corollary 12. Similar to Corollary 1, we have, when $T \to +\infty$,

$$\phi_t^{\text{Deliberate}} \to \phi^{\text{Deliberate}} \equiv \frac{\delta R^2 - 1}{\delta R^2\left(1 - (\lambda^a)^2\right)}$$

$$\Gamma_t \to \Gamma \equiv \frac{\delta R^2 - 1}{\delta R^2\left(1 - \delta R^2 (\lambda^a)^2\right)}$$

From (98) and (99), we know

$$\omega_{t,k}^{\text{Deliberate}} \to \left(\omega^{\text{Deliberate}}\right)^k,$$

where $\omega^{\text{Deliberate}} = 1 - \frac{\delta R^2 \Gamma \lambda^a (\lambda^a - \lambda^y)}{\delta R^2 \Gamma (\lambda^a)^2 + 1} = 1 - \frac{\left(\delta R^2 - 1\right)\lambda^a(\lambda^a - \lambda^y)}{1 - (\lambda^a)^2}$.

[49] This comes from (94) and the fact that $u'(\bar{c}_{t+1}) = \frac{\partial V_{t+2}(\bar{a}_{t+2}, \bar{s}_{t+2})}{\partial a_{t+2}}$ because $\bar{c}_{t+1} = c_{t+1}^{\text{Delibrate}}(\bar{a}_{t+1}, \bar{s}_{t+1})$.

References

Agarwal, Sumit, John Driscoll, Xavier Gabaix, and David Laibson. (2008) “Learning in the Credit Card Market.” *NBER Working Paper No. 13822.*

Allcott, Hunt, Joshua Kim, Dmitry Taubinsky, and Jonathan Zinman. (2020) “Are High-Interest Loans Predatory? Theory and Evidence from Payday Lending.” *UC Berkeley mimeo.*

Anderson, Siwan, and Jean-Marie Baland. (2002) “The economics of roscas and intrahousehold resource allocation.” *The Quarterly Journal of Economics*, 117(3): 963–995.

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.

Ashraf, Nava. (2009) “Spousal control and intra-household decision making: An experimental study in the Philippines.” *American Economic Review*, 99(4): 1245–77.

Auclert, Adrien. (2019) “Monetary Policy and the Redistribution Channel.” *American Economic Review*, 109(6): 2333–67.

Auclert, Adrien, Matthew Rognlie, and Ludwig Straub. (2018) “The intertemporal keynesian cross.” *NBER Working Paper No. 25020.*

Azeredo da Silveira, Rava, and Michael Woodford. (2019) “Noisy Memory and Over-Reaction to News.” Vol. 109, 557–61.

Baicker, Katherine, Sendhil Mullainathan, and Joshua Schwartzstein. (2015) “Behavioral hazard in health insurance.” *The Quarterly Journal of Economics*, 130(4): 1623–1667.

Baker, Malcolm, Stefan Nagel, and Jeffrey Wurgler. (2007) “The Effect of Dividends on Consumption.” *Brookings Papers on Economic Activity*, 2007(1): 231–291.

Barro, Robert. (1999) “Ramsey Meets Laibson in the Neoclassical Growth Model.” *The Quarterly Journal of Economics*, 114(4): 1125–1152.

Bernheim, B Douglas, and Dmitry Taubinsky. (2018) “Behavioral public economics.” *Handbook of Behavioral Economics: Applications and Foundations 1*, 1: 381–516.

Bianchi, Francesco, Cosmin L Ilut, and Hikaru Saijo. (2021) “Implications of Diagnostic Expectations: Theory and Applications.” *NBER Working Paper No. 28604.*

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. (2018) "Diagnostic expectations and credit cycles." *The Journal of Finance*, 73(1): 199–227.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. (2020) "Overreaction in Macroeconomic Expectations." *American Economic Review.*

Browning, Martin, Francois Bourguignon, Pierre-Andre Chiappori, and Valerie Lechene. (1994) "Income and outcomes: A structural model of intrahousehold allocation." *Journal of political Economy*, 102(6): 1067–1096.

Campbell, John, and Gregory Mankiw. (1989) "Consumption, Income and Interest Rates: Reinterpreting the Time Series Evidence." In *NBER Macroeconomics Annual 1989, Volume 4.* 185–246. MIT Press.

Carlin, Bruce, Arna Olafsson, and Michaela Pagel. (2017) "Fintech Adoption across Generations: Financial Fitness in the Information Age." *NBER Working Paper No. 23798.*

Carrera, Mariana, Heather Royer, Mark Stehr, Justin Sydnor, and Dmitry Taubinsky. (2021) "Who Chooses Commitment? Evidence and Welfare Implications." *Review of Economic Studies.*

Carroll, Christopher. (1997) "Buffer-stock Saving and the Life Cycle/Permanent Income Hypothesis." *The Quarterly journal of economics*, 112(1): 1–55.

Carroll, Christopher D, Edmund Crawley, Jiri Slacalek, Kiichi Tokuoka, and Matthew N White. (2020) "Sticky expectations and consumption dynamics." *American economic journal: macroeconomics*, 12(3): 40–76.

Carroll, Christopher D, Martin B Holm, and Miles S Kimball. (2021) "Liquidity constraints and precautionary saving." *Journal of Economic Theory*, 195: 105276.

Chari, Varadarajan V, Patrick J Kehoe, and Ellen R McGrattan. (2007) "Business cycle accounting." *Econometrica*, 75(3): 781–836.

Chetty, Raj, and Adam Szeidl. (2007) "Consumption commitments and risk preferences." *The Quarterly Journal of Economics*, 122(2): 831–877.

Chodorow-Reich, Gabriel, Plamen Nenov, and Alp Simsek. (2019) "Stock Market Wealth and the Real Economy: A Local Labor Market Approach." *NBER Working Paper No. 25959.*

Cochrane, John. (1989) "The Sensitivity of Tests of the Intertemporal Allocation of Consumption to Near-Rational Alternatives." *The American Economic Review*, 319–337.

Di Maggio, Marco, Amir Kermani, and Kaveh Majlesi. (2018) "Stock Market Returns and Consumption." *NBER Working Paper No. 24262.*

Duflo, Esther. (2003) "Grandmothers and granddaughters: old-age pensions and intrahousehold allocation in South Africa." *The World Bank Economic Review*, 17(1): 1–25.

Duflo, Esther, and Christopher R Udry. (2004) "Intrahousehold resource allocation in Cote d'Ivoire: Social norms, separate accounts and consumption choices." *NBER Working Paper No. 10498.*

Fagereng, Andreas, Martin Holm, and Gisle Natvik. (2019) "MPC Heterogeneity and Household Balance Sheets." *University of Oslo mimeo.*

Fagereng, Andreas, Martin Holm, Benjamin Moll, and Gisle Natvik. (2019) "Saving Behavior Across the Wealth Distribution." *University of Oslo mimeo.*

Farhi, Emmanuel, and Xavier Gabaix. (2020) "Optimal Taxation with Behavioral Agents." *American Economic Review*, 110(1): 298–336.

Gabaix, Xavier. (2014) "A Sparsity-Based Model of Bounded Rationality." *Quarterly Journal of Economics*, 129(4).

Gabaix, Xavier. (2016) "Behavioral Macroeconomics via Sparse Dynamic Programming." *NBER Working Paper No. 21848.*

Gourinchas, Pierre-Olivier, and Jonathan Parker. (2002) "Consumption over the Life Cycle." *Econometrica*, 70(1): 47–89.

Hall, Robert E. (1988) "Intertemporal Substitution in Consumption." *Journal of political economy*, 96(2): 339–357.

Harris, Christopher, and David Laibson. (2001) "Dynamic Choices of Hyperbolic Consumers." *Econometrica*, 69(4): 935–957.

Havránek, Tomáš. (2015) "Measuring Intertemporal Substitution: The Importance of Method Choices and Selective Reporting." *Journal of the European Economic Association*, 13(6): 1180–1204.

Holm, Martin Blomhoff. (2018) "Consumption with liquidity constraints: An analytical characterization." *Economics Letters*, 167: 40–42.

Holm, Martin Blomhoff, Pascal Paul, and Andreas Tischbirek. (2020) "The Transmission of Monetary Policy under the Microscope."

Ilut, Cosmin, and Rosen Valchev. (2020) "Economic Agents as Imperfect Problem Solvers." *NBER Working Paper No. 27820.*

Kahneman, Daniel, and Amos Tversky. (1972) "Subjective probability: A judgment of representativeness." *Cognitive psychology*, 3(3): 430–454.

Kahneman, Daniel, Peter Wakker, and Rakesh Sarin. (1997) "Back to Bentham? Explorations of Experienced Utility." *The Quarterly Journal of Economics*, 112(2): 375–406.

Kimball, Miles. (1990) "Precautionary Saving in the Small and in the Large." *Econometrica*, 58(1): 53–73.

Kőszegi, Botond, and Matthew Rabin. (2009) "Reference-dependent Consumption Plans." *American Economic Review*, 99(3): 909–36.

Kueng, Lorenz. (2018) "Excess Sensitivity of High-Income Consumers." *The Quarterly Journal of Economics.*

Laibson, David. (1997) "Golden Eggs and Hyperbolic Discounting." *The Quarterly Journal of Economics*, 112(2): 443–478.

Laibson, David, Peter Maxted, and Benjamin Moll. (2020) "Present Bias Amplifies the Household Balance-Sheet Channels of Macroeconomic Policy." *Harvard University mimeo.*

Levi, Yaron. (2015) "Information architecture and intertemporal choice: A randomized field experiment in the United States." *UCLA mimeo.*

Le Yaouanq, Yves, and Peter Schwardmann. (2019) "Learning about one's self." *CEPR Discussion Paper No. 13510.*

Lian, Chen. (2019) "Consumption with Imperfect Perception of Wealth." *UC Berkeley mimeo.*

Lise, Jeremy, and Ken Yamada. (2019) "Household sharing and commitment: Evidence from panel data on individual expenditures and time use." *The Review of Economic Studies*, 86(5): 2184–2219.

Maćkowiak, Bartosz, and Mirko Wiederholt. (2015) "Business Cycle Dynamics under Rational Inattention." *The Review of Economic Studies*, 82(4): 1502–1532.

Mackowiak, Bartosz, Filip Matejka, and Mirko Wiederholt. (2021) "Rational Inattention: A Review." *Goethe University Frankfurt mimeo.*

Matejka, Filip. (2016) "Rationally inattentive seller: Sales and discrete pricing." *Review of Economic Studies*, 83(3): 1125–1155.

Maxted, Peter. (2020) "Present bias in consumption-saving models: A tractable continuous-time approach."

Mazzocco, Maurizio. (2007) "Household intertemporal behaviour: A collective characterization and a test of commitment." *The Review of Economic Studies*, 74(3): 857–895.

McDowall, Robert. (2020) "Consumption Behavior Across the Distribution of Liquid Assets." *NYU mimeo.*

Mehra, Rajnish, and Edward Prescott. (1985) "The Equity Premium: A Puzzle." *Journal of monetary Economics*, 15(2): 145–161.

Mullainathan, Sendhil. (2002) "A Memory-based Model of Bounded Rationality." *Quarterly Journal of Economics*, 117(3): 735–774.

Mullainathan, Sendhil, Joshua Schwartzstein, and William J Congdon. (2012) "A reduced-form approach to behavioral public finance." *Annu. Rev. Econ.*, 4(1): 511–540.

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.

Olafsson, Arna, and Michaela Pagel. (2018) "The Liquid Hand-to-mouth: Evidence from Personal Finance Management Software." *The Review of Financial Studies*, 31(11): 4398–4446.

Paiella, Monica, and Luigi Pistaferri. (2017) "Decomposing the Wealth Effect on Consumption." *Review of Economics and Statistics*, 99(4): 710–721.

Parker, Jonathan A. (2017) "Why Don't Households Smooth Consumption? Evidence from a 25 Million Dollar Experiment." *American Economic Journal: Macroeconomics*, 9(4): 153–83.

Piccione, Michele, and Ariel Rubinstein. (1997) "On the interpretation of decision problems with imperfect recall." *Games and Economic Behavior*, 20(1): 3–24.

Shefrin, Hersh, and Richard H Thaler. (1988) "The Behavioral Life-cycle Hypothesis." *Economic inquiry*, 26(4): 609–643.

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

Stango, Victor, and Jonathan Zinman. (2014) "Limited and Varying Consumer Attention: Evidence from Shocks to the Salience of Bank Overdraft Fees." *The Review of Financial Studies*, 27(4): 990–1030.

Stephens, Melvin, and Takashi Unayama. (2011) "The Consumption Response to Seasonal Income: Evidence from Japanese Public Pension Benefits." *American Economic Journal: Applied Economics*, 86–118.

Thakral, Neil, and Linh To. (2020) "Anticipation and Consumption." *Brown University mimeo.*

Thaler, Richard. (1990) "Anomalies: Saving, Fungibility, and Mental Accounts." *Journal of Economic Perspectives*, 4(1): 193–205.

Thomas, Duncan. (1990) "Intra-household resource allocation: An inferential approach." *Journal of human resources*, 635–664.