Mistakes in Future Consumption, High MPCs Now*

Chen Lian†

February 11, 2021

Abstract

In a canonical intertemporal consumption problem, I show how anticipation of future behavioral mistakes, by itself, can explain several key empirical puzzles of high-liquidity consumers' consumption behavior—for example, their high marginal propensities to consume (MPCs) and violations of the fungibility principle. This result is independent of the psychological causes of future mistakes, and my framework can accommodate many widely-studied behavioral biases, such as inattention, mental accounting, rules of thumb, and hyperbolic discounting. The approach developed in the paper can be used to study predictions of sophistication independent of the underlying behavioral mistakes in other settings.

*I am extremely grateful to Marios Angeletos, Alp Simsek and Ricardo Caballero for continuous guidance through the project. I am grateful to Nick Barberis, Stefano Dellavigna, Yuriy Gorodnichenko, Pierre-Olivier Gourinchas, Jonathon Hazell, Julian Kozlowski, 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, China Star Tour, Chicago Booth, Columbia, Duke, Duke Fuqua, MIT, Yale SOM, Stanford GSB for very helpful comments and discussions. 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 Predoctoral Fellowship in Behavioral Macroeconomics, awarded through the NBER. Bruno Smaniotto provides excellent research assistance.

†UC Berkeley; chen_lian@berkeley.edu.

1 Introduction

There is increasing evidence that consumers exhibit important deviations from the permanent income hypothesis away from liquidity constraints (Thaler, 1990). High-liquidity consumers exhibit high MPCs (Parker, 2017; Kueng, 2018; Olafsson and Pagel, 2018; Fagereng, Holm and Natvik, 2019; McDowall, 2020). They also violate the fungibility principle (Shefrin and Thaler, 1988), 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 (Maggio, Kermani and Majlesi, 2019). This evidence is hard to square with canonical liquidity-constraints-based models (Carroll, 1997; Gourinchas and Parker, 2002; Kaplan and Violante, 2010, 2014) and points toward behavioral explanations. Understanding consumption behavior away from liquidity constraints is important because it plays a key role in determining the macroeconomic impact of monetary and fiscal policies (Auclert, 2019; Holm, Paul and Tischbirek, 2020).

The behavioral approach can explain deviations from the permanent income hypothesis away from liquidity constraints. But the myriad of potential behavioral biases, e.g., mental accounting (Thaler, 1990), inattention (Sims, 2003; Reis, 2006; Maćkowiak and Wiederholt, 2015; Gabaix, 2016; Caplin, Dean and Leahy, 2019), present focus (Laibson, 1997), self control (Gul and Pesendorfer, 2004; Fudenberg and Levine, 2006), and distorted expectations (Azeredo da Silveira and Woodford, 2019; Bordalo, Gennaioli and Shleifer, 2018; Bordalo et al., 2020) can make it difficult to derive robust, consistent lessons about consumption behavior.

In this paper, I provide a new angle for studying how behavioral biases can influence consumption behavior. Unlike the majority of existing behavioral literature, I do not take an exact stand on the underlying behavioral biases. Instead, I use "wedges" (Chari, Kehoe and McGrattan 2007; Shimer 2009; Farhi and Gabaix, 2020) to capture how actual consumption rules deviate from their optimal counterparts. I can then study the implications of behavioral mistakes on consumption independent of their exact psychological causes. The innovation of this paper is to use this wedge-based approach to study robust, positive behavioral implications of knowledge about future mistakes, i.e., sophistication in the language of O'Donoghue and Rabin (1999, 2001).[1] A recent behavioral literature instead uses such wedge-based approaches to study normative welfare implications of behavioral mistakes (Mullainathan, Schwartzstein and Congdon, 2012; Baicker,

[1] My approach is different from O'Donoghue and Rabin (1999, 2001) in two ways. First, I do not take an exact stand on where future mistakes come from, while they focus on the impact of future present focus. Second, I focus on continuous decisions, e.g., standard intertemporal consumption and saving problems, while they focus on discrete choices.

Mullainathan and Schwartzstein, 2015; Bernheim and Taubinsky, 2018; Farhi and Gabaix, 2020).

Generally speaking, behavioral mistakes can impact consumption through two distinct channels. The first channel captures the direct impact of current behavioral mistakes on current decisions, e.g., how current inattention or current present focus impacts current consumption. Though the impact of this channel can be quantitatively very important, its sign depends on the exact underlying bias and this channel can lead to either high or low current MPCs. The second channel, instead, captures how anticipation of future mistakes, i.e., sophistication, impacts current consumption. My contribution is to show that, once we isolate this channel, anticipation of future consumption mistakes (in response to changes in savings), no matter the behavioral cause of these mistakes, robustly leads to high current MPCs.

Mistakes in future consumption, high MPCs now. I study a canonical intertemporal consumption and saving problem. The key result is that future consumption mistakes in response to changes in savings lead to higher current MPCs. With these future mistakes, the consumer is less willing to adjust her savings and more willing to adjust her current consumption. Hence she displays higher current MPCs.

Let me use responses to a positive current income shock as an example. In the standard frictionless model, if the consumer increases her savings, her future selves need to perfectly coordinate by increasing each of their consumptions roughly equally. This coordinated response may be hard in practice. If her future selves respond to changes in savings imperfectly, the consumer will instead increase her current consumption more.[2]

The high current MPCs result does not depend on the behavioral causes of future consumption mistakes. No matter whether future consumption mistakes lead to over-reaction or under-reaction to changes in savings, these mistakes robustly increase current MPCs. I then illustrate how my framework can accommodate many widely-studied behavioral biases, such as inattention, rules of thumb, hyperbolic discounting, and mental accounting. My result also accommodates an alternative inter-personal interpretation. Since household consumption is decided jointly by different members of the household, the current self (e.g., the wife) may display a higher MPC because she is worried about the inefficient spending of future selves (e.g., the husband or the children).

I further clarify that it is the anticipation of future mistakes that leads to the high current MPCs. In other words, current MPCs increase with perceived future mistakes. Using the language of O'Donoghue and Rabin (1999, 2001), partial sophistication, i.e., partial understanding of

[2]By the same token, after a negative shock, if the consumer decreases her savings, her future selves will not be able to perfectly coordinate their consumption decreases. The consumer instead decreases her current consumption more.

future mistakes, suffices for all qualitative results. An important comparative statics result is that current MPCs increase with the degree of sophistication. I also clarify that the reason why future "mistakes" matter is that future selves may behave differently from what the current self deems optimal.

The key intermediate step to prove the high current MPC results is to notice that future consumption mistakes lead to the excess concavity of the continuation value function. This same mechanism also naturally helps explain two other well-known puzzles in intertemporal decisions: high risk aversion and the equity premium puzzle (Mehra and Prescott, 1985); and the small elasticity of intertemporal substitution, i.e., the empirical evidence on the small consumption responses to interest rate changes (Hall, 1988; Campbell and Mankiw, 1989; Havránek, 2015).

Future non-fungibility begets current non-fungibility. The anticipation of future consumption mistakes, by itself, can also explain the violation of the fungibility principle (Shefrin and Thaler, 1988) away from liquidity constraints. By the fungibility principle, I mean 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.

In a general non-fungible case, I apply the same methodology to use behavioral wedges to capture how future consumption deviates from its frictionless counterpart. I allow inefficiently differential responses of future consumption to different components of permanent income. I show how this non-fungibility of future consumption, by itself, suffices to generate the non-fungibility of the current consumption. In other words, even if the current self fully understands the permanent income hypothesis, as long as she anticipates future consumption mistakes in the form of future non-fungibility, she will respond differentially to different components of permanent income.

For example, if future consumption responds inefficiently more to income than to savings, current consumption will respond less to future income and exhibit excess discounting of future income. In this sense, mistakes in future consumption beget current non-fungibility. Such excess discounting of future income away from liquidity constraints is also consistent with the empirical evidence in Kueng (2018).

Contributions to the literature. This paper builds upon the behavioral literature on intertemporal consumption problems, e.g. , inattention (e.g. Gabaix and Laibson, 2002; Sims, 2003; Reis, 2006; Luo, 2008; Abel, Eberly and Panageas, 2007, 2013; Luo and Young, 2010; Alvarez, Guiso and Lippi, 2012; Maćkowiak and Wiederholt, 2015; Gabaix, 2016; Caplin, Dean and Leahy, 2019), present focus (e.g. Laibson, 1997; Barro, 1999; Angeletos et al., 2001), mental accounting (Shefrin and Thaler, 1988; Thaler, 1990), distorted expectations (e.g. Azeredo da Silveira and

Woodford, 2019; Bordalo, Gennaioli and Shleifer, 2018; Bordalo et al., 2020), news utility (e.g. Kőszegi and Rabin, 2009; Pagel, 2017), and anticipatory utility (Thakral and To, 2020).

Compared to this large literature, this paper takes a new route. Instead of studying a specific behavioral bias, I apply the wedge approach, widely used to study macroeconomic frictions (Chari, Kehoe and McGrattan, 2007; Shimer, 2009), to develop robust predictions independent of the exact behavioral mistakes. A separate strand of behavioral literature (e.g., Mullainathan, Schwartzstein and Congdon, 2012; Baicker, Mullainathan and Schwartzstein, 2015; Bernheim and Taubinsky, 2018; Farhi and Gabaix, 2020) uses the wedge approach to study welfare implications and optimal policies with behavioral agents. This literature does not touch upon the robust positive implications of sophistication on behavior that I focus on.

In terms of applications, this paper provides a potential explanation for the empirical evidence on deviations from the permanent income hypothesis away from liquidity constraints. This evidence includes high liquidity consumers' high MPCs (Parker, 2017; Kueng, 2018; Olafsson and Pagel, 2018; Fagereng, Holm and Natvik, 2019; McDowall, 2020) and their deviations from the fungibility principle (Thaler, 1990; Baker, Nagel and Wurgler, 2007; Di Maggio, Kermani and Majlesi, 2018; Fagereng et al., 2019).

Ilut and Valchev (2020) provide another behavioral theory of high-liquidity consumers' high MPCs. That theory is based on the consumer's current difficulty in figuring out the optimal consumption rules and does not depend on the anticipation of future mistakes.

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, and quadratic utility function, and the discount factor is set to be 1 for simplicity.

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. Her intertemporal budget is given by

$$a_{t+1} = R\left(a_t + y_t - c_t\right) \quad \forall t \in \{0, 1\}, \tag{2}$$

where y_t is her exogenous income at period t and a_t is her wealth (i.e. savings/borrowings) at the start of period t.

The question of interest is how current consumption c_0 responds to changes in current income y_0, that is, the current MPC. For illustration purposes, I focus on the impact of y_0 and set initial wealth $a_0 = 0$ and future income $y_1 = y_2 = 0$.

At period $t = 2$, the consumer consumes out of her remaining savings,[3]

$$c_2\left(a_2\right) = a_2. \tag{3}$$

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

$$c_1\left(a_1\right) = \frac{1}{2}\left(1 - \lambda_1\right) a_1 + \bar{c}_1, \tag{4}$$

which may deviate from the frictionless consumption rule $c_1\left(a_1\right) = \frac{1}{2} a_1$. λ_1 captures the mistake in the response of c_1 to changes in savings a_1. When $\lambda_1 > 0$, c_1 under-reacts to changes in savings a_1. When $\lambda_1 < 0$, c_1 over-reacts to changes in savings a_1. In (4), the overall consumption level can also deviate from its frictionless counterpart. This "level mistake" is captured by $\bar{c}_1$.

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). Later, in Section 4.2, I discuss in detail the interpretation of "mistakes."

This paper focuses on how these future mistakes impact current MPCs at $t = 0$. To isolate the impact of this channel, I define the notion of "deliberate consumption," i.e., the consumption that the consumer would have chosen if she were not subject to any current mistake at $t = 0$ but took her future mistakes as given:

$$\begin{aligned} c_0^{\text{Deliberate}}\left(y_0\right) &= \arg\max_{c_t} u\left(c_0\right) + u\left(c_1\left(a_1\right)\right) + u\left(c_2\left(a_2\right)\right) \\ &= \phi_0^{\text{Deliberate}} y_0 + \bar{c}_0^{\text{Deliberate}}, \end{aligned} \tag{5}$$

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

subject to the budget (2) and future consumptions rules in (3) and (4), where $\phi_0^{\text{Deliberate}}$ captures the current MPC.

Proposition 1. *The current MPC $\phi_0^{Deliberate}$, a function of the future mistake λ_1, strictly increases with $|\lambda_1|$.*

Proposition 1 shows that a larger future consumption mistake $|\lambda_1|$ increases the current MPC. It is always true that $\phi_0^{\text{Deliberate}} \geq \phi_0^{\text{Frictionless}}$, where $\phi_0^{\text{Frictionless}}$ is the frictionless MPC if the consumer is not subject to any current or future mistakes ($\lambda_1 = 0$). When future consumption responds inefficiently to changes in savings (a larger $|\lambda_1|$), the consumer is less willing to adjust her savings. In response to changes in current income, she is more willing to adjust her current consumption and displays a higher MPC.

For example, consider a positive income shock to y_0. In the standard frictionless model, if the consumer saves this additional income, her future selves will perfectly coordinate by increasing c_1 and c_2 by exactly the same amount. In practice, however, there may be some imperfection in future consumption responses. Because of these inefficient responses of future consumption, the consumer instead increases her current consumption more. By the same token, after a negative shock, if the current self decreases her savings, her future selves will not respond to the decrease in savings efficiently. She instead decreases her current consumption more.

Let me also provide an overview of the proof of Proposition 1, which is essentially the same as the more general cases studied below. For this purpose, let me define the continuation value function based on future consumption rules in (3) and (4):

$$V_1(a_1) \equiv u(c_1(a_1)) + u(c_2(a_1 - c_1(a_1))).$$

The proof has two steps. First, a larger future consumption mistake $|\lambda_1|$ (in response to changes in savings) leads to excess concavity of the continuation value function, i.e., a larger $\left|V_1''(a_1)\right|$. The concavity of the continuation value function captures how fast the marginal value of saving falls with additional saving. In the standard frictionless case, future selves can spread the consumption increase evenly across different periods, preventing the marginal value of saving from falling too fast. With inefficient responses of future consumption to changes in savings, the consumption increase will be more concentrated in some periods.[4] The marginal value of saving falls faster, and the continuation value function is more concave.

[4]Here, if $\lambda_1 < 0$ (period 1 consumption over-reacts), the consumption increase after an increase of a_1 will be more concentrated in $t = 1$. If $\lambda_1 > 0$ (period 1 consumption under-reacts), the consumption increase will be more concentrated in $t = 2$.

Second, the excess concavity of the continuation value function leads to a high current MPC $\phi_0^{\text{Deliberate}}$. To see this, first notice that the *level* of $t=0$ consumption is connected to the first derivative of the continuation value (marginal value of savings): $u'\left(c_0^{\text{Deliberate}}\left(y_0\right)\right) = V_1'\left(R\left(y_0 - c_0^{\text{Deliberate}}\left(y_0\right)\right)\right)$. The current *MPC* $\phi_0^{\text{Deliberate}} \equiv \frac{\partial c_0^{\text{Deliberate}}(y_0)}{\partial y_0}$ is then connected to the second derivative of the continuation value, i.e., the concavity: $\phi_0^{\text{Deliberate}} = \frac{V_1''}{u''+V_1''}$. The excess concavity of the continuation value then leads to a high current MPC.

The result in Proposition 1 does not depend on the exact behavioral causes of the future mistake λ_1. As illustrated in the rest of the paper, my framework can accommodate several widely-studied behavioral biases, such as inattention, rules of thumb, hyperbolic discounting, and mental accounting.

There are several additional points worth mentioning. First, the high MPC result holds regardless of whether the future consumption mistake takes the form of under-reaction ($\lambda_1 > 0$) or over-reaction ($\lambda_1 < 0$). In this sense, future mistakes robustly increase the current MPC.

Second, the key for the high current MPC $\phi_0^{\text{Deliberate}}$ is the inefficient response of future consumption to changes in savings, i.e., λ_1 in (4). On the other hand, the mistakes in overall future consumption level, $\bar{c}_1$ in (4), do not matter for the current MPC $\phi_0^{\text{Deliberate}}$. Some behavioral biases (e.g. hyperbolic discounting) lead to both inefficient responses and "level mistakes." The result here clarifies that, in terms of the impact from future mistakes, it is inefficient responses that generate the high current MPC.

Beyond the quadratic-linear environment here, one may naturally wonder what happens with a "prudent" utility ($u''' > 0$). The key result in the literature (Kimball, 1990) is that, facing future uncertainty, a prudent agent will display a lower consumption *level*, i.e., a lower $\bar{c}_0^{\text{Deliberate}}$ in (5). A corollary is that, with prudence, future mistakes (e.g., $\bar{c}_1$) lower the current consumption *level* ($\bar{c}_0^{\text{Deliberate}}$), or equivalently raise the current saving level ("precautionary saving"). I further explain this result in Appendix B. But such a lower consumption level can coexist with my key result of a higher current MPC $\phi_0^{\text{Deliberate}}$. In fact, with general concave utilities, the key result here that future inefficient responses lead to a higher current MPC remains to be true (see Proposition 4 below).

Third, it the anticipation of future mistakes, i.e., perceived future mistakes, leads to the consumer's high current MPC. In the analysis here, perceived future mistakes coincide with actual future mistakes. But the analysis accommodates a more general interpretation if we re-define the deliberate consumption in (5) based on perceived future consumption rules and perceived future mistakes. Proposition 1 can be re-stated as that the current MPC increases with perceived future mistakes. With this re-interpretation, the analysis here easily accommodates the case of partial sophistication in O'Donoghue and Rabin (1999, 2001). A natural corollary of Proposition 1 is that a larger degree of sophistication, i.e., more knowledge about future mistakes, leads to a higher current MPC.

This more general re-interpretation also helps clarify why future "mistakes" matter. The key is that the current self anticipates that her future selves' consumption rules deviate from what she deems optimal. These "mistakes" can come from either optimization errors or time-inconsistency in preferences. See Section 4.2 below for details.

3 Set up

This section introduces a standard intertemporal consumption and saving problem. Then, I introduce the notion of "deliberate consumption" to isolate the impact of future consumption mistakes on current consumption.

Utility and budget. I first introduce a canonical, single-agent, intertemporal consumption problem. 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.

The consumer's experienced utility is given by

$$U_0 \equiv \sum_{t=0}^{T-1} \delta^t u\left(c_t\right) + \delta^T v\left(a_T + y_T\right), \tag{6}$$

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 concave. In the main analysis, for analytical results and to ensure that the high MPCs results are not driven by precautionary saving motives, I let $u(\cdot)$ and $v(\cdot)$ be quadratic, similar to the example in Section 2. But the main high MPCs results hold with general concave utilities (see Proposition 4 below).

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

$$a_{t+1} = R\left(a_t + y_t - c_t\right) \quad \forall t \in \{0, \cdots, T-1\}, \tag{7}$$

where y_t is her exogenous income at period t, a_t is her wealth (i.e. savings/borrowings) at the start of period t, and R is the gross interest rate on the risk-free asset.

In each period t, the payoff relevant state for the consumer in each period t can be summarized by

$$(a_t, s_t),$$

where s_t is the exogenous income state at period t summarizing information about current income y_t and future incomes $\{y_{t+k}\}_{k\geq 1}$, and a_t is the endogenously determined current wealth level based on the consumer's past decisions (except the exogenous initial wealth a_0).

For illustration purposes, I follow Chetty and Szeidl (2007) and assume that all income uncertainty in the economy is resolved in period 0, so $s_t = (y_t, \cdots, y_T)$. It is worth noticing that, with a quadratic utility and a linear decision rule, the well-known certainty equivalence result implies that the consumer's MPC remains the same with gradual resolution of income uncertainty (see Corollary 8 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 $c_t(a_t, s_t)$ to denote each self t's *actual* consumption rule, subject to behavioral biases.

Isolating the impact of future mistakes. Behavioral biases can impact self t's actual consumption rule $c_t(a_t, s_t)$ 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 focus 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 $c_t^{\text{Deliberate}}(a_t, s_t)$, 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.

Definition 1. *For each $t \in \{0, \cdots, T-1\}$, self t's deliberate consumption rule optimizes the consumer's utility in (6), taking future selves' actual consumption rules $\{c_{t+k}(a_{t+k}, s_{t+k})\}_{k=1}^{T-k-1}$ as given:*

$$c_t^{\textit{Deliberate}}(a_t, s_t) \equiv \arg\max_{c_t} u(c_t) + \sum_{k=1}^{T-t-1} \delta^{k-1} u(c_{t+k}(a_{t+k}, s_{t+k})) + \delta^{T-t} v(a_T + y_T), \tag{8}$$

subject to the budget in (7).

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

$$c_t(a_t, s_t) = \mathcal{S}\left(c_t^{\text{Deliberate}}(a_t, s_t), \lambda_t\right). \tag{9}$$

Self t's own behavioral bias (parameterized by λ_t) impacts actual consumption by letting it deviate from deliberate consumption $c_t^{\text{Deliberate}}(a_t, s_t)$, captured by the function $\mathcal{S}$.[5] On the other hand, anticipation of future selves' mistakes $\{\lambda_{t+k}\}_{k=1}^{T-t-1}$ impacts current actual consumption through deliberate consumption $c_t^{\text{Deliberate}}(a_t, s_t)$.

The main theme for the rest of the paper is that, once we isolate the impact of future consumption mistakes through deliberate consumption (8), we can show that future mistakes robustly lead to high current MPCs, no matter the micro-foundations of these mistakes.

As in the example in Section 2, here the deliberate consumption in (8) is defined based on correct knowledge of future actual consumptions rules. This choice significantly simplifies the notation without changing the essence. The analysis can accommodate a more general interpretation if we define deliberate consumption (8) based on perceived future consumption rules. See Section 4.2 for details.

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

$$V_t(a_t, s_t) \equiv u(c_t(a_t, s_t)) + \sum_{k=1}^{T-t-1} \delta^k u(c_{t+k}(a_{t+k}, s_{t+k})) + \delta^{T-t} v(a_T + y_T), \tag{10}$$

subject to the budget in (7). For the last period T, we have $V_T(a_T, s_T) = v(a_T + y_T)$.

Based on (10), I can express the deliberate consumption rule in (8) recursively. This recursive formulation paves the ways for the analysis in the rest of the paper.

Proposition 2. *For $t \in \{0, \cdots, T-1\}$, each self t's deliberate consumption rule defined in (8) satisfies*

$$c_t^{\textit{Deliberate}}(a_t, s_t) = \max_{c_t}\ u(c_t) + \delta V_{t+1}(R(a_t + y_t - c_t), s_{t+1}). \tag{11}$$

[5] We have $\mathcal{S}\left(c_t^{\text{Deliberate}}(a_t, s_t), 0\right) = c_t^{\text{Deliberate}}(a_t, s_t)$. 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 (8).

Moreover, for $t \in \{0, \cdots, T-1\}$, the value function $V_t(a_t, s_t)$ defined in (10) satisfies

$$V_t(a_t, s_t) = u(c_t(a_t, s_t)) + \delta V_{t+1}(R(a_t + y_t - c_t(a_t, s_t)), s_{t+1}), \tag{12}$$

where the actual consumption rule $c_t(a_t, s_t)$ is given by (9).

Finally, if consumption rules and value functions $\left\{c_t^{Deliberate}(a_t, s_t), c_t(a_t, s_t)\right\}_{t=0}^{T-1}$ and $\{V_t(a_t, s_t)\}_{t=0}^{T}$ satisfy (9), (11), (12), and the boundary condition $V_T(a_T, s_T) = v(a_T + y_T)$, they coincide with the corresponding objects defined sequentially in (8)–(10).

A note on budget constraints. It is worth noting that the final wealth a_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 (7) 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 Fungibility Case

I first study a benchmark fungibility case in which actual future consumptions, as in the frictionless case, remain functions of the permanent income (i.e. total present value of all components of incomes and savings/borrowings). Future consumption mistakes come from their inefficient responses to changes in permanent income (which embed inefficient responses to changes in savings). The main result is that these mistakes lead to high current MPCs. This result does not depend on the exact micro-foundations of future consumption mistakes. I then illustrate how my framework can accommodate common behavioral biases such as inattention, rules of thumb, hyperbolic discounting, and stochastic mistakes.

4.1 Mistakes in Future Consumption, High MPCs Now

In this section, I consider a benchmark fungibility case where, as in the permanent income hypothesis, both c_t and $c_t^{\text{Deliberate}}$ are functions of permanent income: $w_t = a_t + y_t + \sum_{k=1}^{T-t} R^{-k} y_{t+k}$.

In the quadratic-linear environment considered here, I can write the actual and deliberate consumption rules as

$$c_t(w_t) = \phi_t w_t + \bar{c}_t \quad \text{and} \quad c_t^{\text{Deliberate}}(w_t) = \phi_t^{\text{Deliberate}} w_t + \bar{c}_t^{\text{Deliberate}}, \tag{13}$$

where ϕ_t captures self t's actual MPC, $\phi_t^{\text{Deliberate}}$ captures her deliberate MPC, $\bar{c}_t$ captures the level of her actual consumption, and $\bar{c}_t^{\text{Deliberate}}$ captures the level of her deliberate consumption. I can then define the value function $\{V_t(w_t)\}_{t=0}^{T}$ based on (10).

Here, the key mistakes in actual consumption rules come from their inefficient responses to changes in permanent income (which embed inefficient responses to changes in savings). That is, the actual MPC ϕ_t may deviate from the deliberate MPC $\phi_t^{\text{Deliberate}}$. As in (9), I use λ_t to capture this mistake,

$$\phi_t = (1 - \lambda_t)\, \phi_t^{\text{Deliberate}}. \tag{14}$$

In other words, λ_t in (14) can be viewed as a behavioral "wedge" between self t's actual MPC ϕ_t and her deliberate MPC $\phi_t^{\text{Deliberate}}$. When $\lambda_t > 0$, self t's actual consumption under-reacts to changes in w_t. When $\lambda_t < 0$, self t's actual consumption over-reacts to changes in w_t. Each self's mistake $\{\lambda_t\}_{t=0}^{T-1}$ is exogenous here, but I will connect these mistakes to the exact underlying behavioral biases in Section 4.3.

Mistakes in actual consumption rules may also involve "level" mistakes, i.e., $\bar{c}_t \neq \bar{c}_t^{\text{Deliberate}}$. But as in the example in Section 2, future "level" mistakes will not directly impact the current self's deliberate MPC $\phi_t^{\text{Deliberate}}$.

The main result of this section studies how future selves' consumption mistakes robustly impact current deliberate consumption. Based on Definition 1, I can calculate current self t's deliberate consumption rule $c_t^{\text{Deliberate}}(w_t)$ and her deliberate MPC $\phi_t^{\text{Deliberate}}$ from future selves' actual consumption rules $\{c_{t+k}(w_{t+k})\}_{k=1}^{T-1-t}$.

Proposition 3. *For $t \in \{0, \cdots, T-2\}$, each self t's deliberate MPC $\phi_t^{Deliberate}$ is a function of $\left(\{\lambda_{t+k}\}_{k=1}^{T-t-1}, \delta, R\right)$. Moreover, $\phi_t^{Deliberate}$ increases with each future self's mistake $\{|\lambda_{t+k}|\}_{k=1}^{T-t-1}$.*

Proposition 3 shows that future consumption mistakes increase the current MPC. It is always true 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. In other words, regardless of whether future consumption mistakes take the form of under-reaction ($\lambda_{t+k} > 0$) or over-reaction ($\lambda_{t+k} < 0$), these mistakes robustly increase current deliberate MPCs.

Excess concavity of the continuation value function. To understand Proposition 3, let me first introduce an intermediate step. From the recursive formulation in (11), we know that understanding the properties of the continuation value function is crucial for understanding MPCs today.

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 $t \in \{0, \cdots, T-1\}$,

$$\Gamma_{t+1} \equiv \frac{\partial^2 V_{t+1}(w_{t+1})}{\partial w_{t+1}^2} / u'' > 0, \tag{15}$$

where a larger Γ_{t+1} means a more concave value function $V_{t+1}(w_{t+1})$.[6][7]

Lemma 1. *Future consumption mistakes lead to excess concavity of the continuation value function. That is, for $t \in \{0, \cdots, T-2\}$, Γ_{t+1} strictly increases with each $\{|\lambda_{t+k}|\}_{k=1}^{T-t-1}$.*

To understand this result, note that the concavity of the continuation value function captures how fast the marginal value $V'_{t+1}(w_{t+1})$ decreases with additional w_{t+1}. In the standard frictionless case, future selves can spread the consumption increase evenly across different periods, preventing the marginal value of saving from falling too fast. With inefficient responses of future consumption to changes in w_{t+1}, the consumption increase will be more concentrated in some periods. The marginal value $V'_{t+1}(w_{t+1})$ decreases faster, and the continuation value function is more concave.

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 consumption mistakes robustly 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 (11), we know

$$c_t^{\text{Deliberate}}(w_t) = u(c_t) + \delta V_{t+1}(R(w_t - c_t)). \tag{16}$$

Because of the excess concavity of the continuation value function in Lemma 1, in response to changes in current permanent income, the current self is more willing to adjust her current consumption instead of her savings. She hence displays a higher MPC. 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.

Consider a positive shock to w_t. If the current self saves this additional money, in the standard frictionless model, her future selves can perfectly coordinate to increase each of their consumptions

[6] $u'' < 0$, a constant, is the second derivative of the utility function. Moreover, the definition in (15) can be extended to $\Gamma_0 \equiv \frac{\partial^2 V_0(w_0)}{\partial w_0^2} / u''$.

[7] 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).

accordingly. This coordinated response may be hard in practice. Here, future selves may respond inefficiently to the increase in savings. The current self then increases her current consumption more. By the same token, after a negative shock, if the current self decreases her savings, her future selves will not respond to the decrease in savings efficiently. She instead decreases her current consumption more.

Mathematically, first notice that the level of consumption at t is connected to the first derivative of the continuation value:

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

The current *MPC* $\phi_t^{\text{Deliberate}} \equiv \frac{\partial c_t^{\text{Deliberate}}(w_t)}{\partial w_t}$ is then connected to 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 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 always raise the current MPC, regardless of whether future selves over-react ($\lambda_{t+k} < 0$) or under-react ($\lambda_{t+k} > 0$) to changes in permanent income. This result is in contrast with the impact of current behavioral biases (λ_t) on the current MPC, which can go either way.

4.2 Perceived Future Mistakes and the Degree of Sophistication

A more general interpretation of the main analysis. In essence, it is the anticipation of future mistakes that leads to high current MPCs. In the main analysis, for notation simplicity, I define the deliberate consumption (8) based on correct anticipation of future actual consumption rules and future mistakes. But the analysis can accommodate a more general interpretation if we I re-define deliberate consumption (8) based on perceived future consumption rules and perceived future mistakes.

Specifically, let $\{\tilde{c}_{t,t+k}\left(w_{t+k}\right)\}_{k=1}^{T-t-1}$ capture self t's perceived future consumption rules. We can redefine the deliberate consumption based on these perceived future consumption rules:

$$c_t^{\text{Deliberate}}\left(w_t\right) \equiv \arg\max_{c_t} u\left(c_t\right) + \sum_{k=1}^{T-t-1} \delta^{k-1} u\left(\tilde{c}_{t,t+k}\left(w_{t+k}\right)\right) + \delta^{T-t} v\left(w_T\right), \tag{17}$$

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

We can then re-state Proposition 3 as how perceived future mistakes $\left\{\tilde{\lambda}_{t,t+k}\right\}_{k=1}^{T-t-1}$, defined based on perceived future consumption rules $\left\{\tilde{c}_{t,t+k}\left(w_{t+k}\right)\right\}_{k=1}^{T-t-1}$ similar to (14), increase current MPCs, $\phi_t^{\text{Deliberate}} \equiv \frac{\partial c_t^{\text{Deliberate}}}{\partial w_t}$.[8]

Corollary 1. *Based on the definition in (17), for* $t \in \{0, \cdots, T-2\}$, *each self* t*'s deliberate MPC* $\phi_t^{Deliberate}$ *is given by the same function as in Proposition 3, with perceived future mistake* $\tilde{\lambda}_{t,t+k}$ *replacing the actual future mistake* λ_{t+k}. *As a result,* $\phi_t^{Deliberate} \geq \phi_t^{Frictionless}$ *and increases with perceived future mistake* $\left\{\left|\tilde{\lambda}_{t,t+k}\right|\right\}_{k=1}^{T-t-1}$.

In other words, in this re-interpretation based on perceived future mistakes, the main analysis is exactly the same. As a result, for the rest of the paper, I maintain the original definition of deliberate consumption in (8) based on future actual consumption rules. This tremendously simplifies the notation without changing the essence of the results.

Partial sophistication and comparative statics. 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 has a partial understanding of future mistakes, and her perceived future mistakes are given by:

$$\tilde{\lambda}_{t,t+k} = s_t \lambda_{t+k}, \tag{18}$$

where $s_t \in [0, 1]$ captures the degree of self t's sophistication. From Corollary 1, 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 2. *Each self* t*'s deliberate MPC* $\phi_t^{Deliberate}$, *defined based on (17) and (18), increases with the degree of sophistication* s_t.

The essence behind future "mistakes." The above more general re-interpretation of my results also helps clarify why future "mistakes" matter. In the main analysis, I follow Kahneman, Wakker and Sarin (1997), define the experienced utility in (6), and use "mistakes" to capture

[8] Based on perceived future consumption rules $\left\{\tilde{c}_{t,t+k}\left(w_{t+k}\right)\right\}_{k=1}^{T-t-1}$, we can define perceived future mistakes $\left\{\tilde{\lambda}_{t,t+k}\right\}_{k=1}^{T-t-1}$ similar to (14). 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 $\left\{\tilde{c}_{t,t+k+l}\left(w_{t+k+l}\right)\right\}_{l=1}^{T-t-k-1}$: $c_{t,t+k}^{\text{Deliberate}}\left(w_{t+k}\right) \equiv \arg\max_{c_t} u\left(c_t\right) + \sum_{l=1}^{T-t-k-1} \delta^{k-1} u\left(\tilde{c}_{t,t+k+l}\left(w_{t+k+l}\right)\right) + \delta^{T-t} v\left(w_T\right)$, subject to the budget. We can then define self t's perceived future mistake $\tilde{\lambda}_{t,t+k}$ at $t+k$ similar to (14): $\frac{\partial \tilde{c}_{t,t+k}}{\partial w_{t+k}} = \left(1 - \tilde{\lambda}_{t,t+k}\right) \frac{\partial c_{t,t+k}^{\text{Deliberate}}}{\partial w_{t+k}}$.

deviations from the deliberate decision rule derived based on the experienced utility. With the more general re-interpretation here, perceived future mistakes matter because the current self anticipates that her future selves will deviate from what she deems optimal. In fact, one can view (17) as self t optimizing her decision utility. Perceived future mistakes $\left\{\tilde{\lambda}_{t,t+k}\right\}_{k=1}^{T-t-1}$ capture how self t's perceived future consumption rules $\{\tilde{c}_{t,t+k}\left(w_{t+k}\right)\}_{k=1}^{T-t-1}$ deviate from what self t deems optimal based on the decision utility. Corollary 1 applies verbatim.

4.3 Different Micro-Foundations, Same Results

The high MPC result in Proposition 3 does not depend on the exact behavioral causes of mistakes in future consumption. But even the simple fungibility case here accommodates many widely-studied behavioral biases, such as inattention, rules of thumb, hyperbolic discounting, and stochastic mistakes.

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

$$w_t^p\left(w_t\right) = \left(1-\lambda_t\right)w_t + \lambda_t w_t^d, \tag{19}$$

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 (an exogenous constant whose value does not matter for the MPCs). It is worth noting that an alternative way to model inattention is through noisy signals (Sims, 2003). With linear consumption rules and Normally distributed incomes, the two approaches lead to the same predictions on MPCs (see Appendix B).

Based on the perceived permanent income $w_t^p\left(w_t\right)$ in (19), the actual consumption rule for each self t is given by

$$c_t\left(w_t\right) = \arg\max_{c_t}\; u\left(c_t\right) + \delta V_{t+1}\left(R\left(w_t^p\left(w_t\right) - c_t\right)\right), \tag{20}$$

where the continuation value function V_{t+1} is defined as in (10), based on future selves' actual consumption rules.

To isolate the impact of future inattention on current consumption, the deliberate consumption is defined as (11) taking future selves' inattention to permanent income as given:

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

As a corollary of Proposition 3, future consumption mistakes in the form of inattention lead to high current MPCs.

Corollary 3. *For $t\in\left\{0,\cdots,T-2\right\}$, $\phi_t^{Deliberate}\equiv\frac{\partial c_t^{Deliberate}}{\partial w_t}\geq\phi_t^{frictionless}$ and $\phi_t^{Deliberate}$ increases with future selves' degrees of inattention $\left\{\lambda_{t+k}\right\}_{k=1}^{T-t-1}$. Moreover, the degrees of of inattention $\left\{\lambda_{t+k}\right\}_{k=1}^{T-t-1}$ here coincide exactly with $\left\{\lambda_{t+k}\right\}_{k=1}^{T-t-1}$ in the general framework in Proposition 3.*

This result means that, once we isolate the impact of future inattention on current MPCs, it raises current MPCs. When the current self is attentive ($\lambda_t=0$), this result then unambiguously translates into a high current actual MPC. Even if the current self is inattentive, the above result translates into a high current actual MPC out of *perceived* permanent income.

Heuristics and rules of thumb. Another commonly studied behavioral bias is heuristics and rules of thumb (e.g., Kahneman, 2011). To capture it in the environment here, I let the actual consumption rule for each self $t\in\left\{0,\cdots,T-1\right\}$ be given by

$$c_t\left(w_t\right)=\begin{cases}c_t^R\left(w_t\right) & \text{with probability}\ \ p_t\\ c_t^{\text{Deliberate}}\left(w_t\right) & \text{with probability}\ \ 1-p_t,\end{cases}$$

where $c_t^R\left(w_t\right)\equiv\phi_t^R w_t+\bar{c}_t^R$ captures a rule of thumb. That is, with probability p_t, the current self makes her consumption decision based on "system 1," following a simple rule of thumb captured by $c_t^R\left(w_t\right)$. With probability $1-p_t$, the current self makes her consumption decision based on "system 2:" actual consumption coincides with deliberate consumption.[9]

The deliberate consumption rule is defined as usual:

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

not subject to any current behavioral bias and taking future selves' mistakes as given. As a corollary of Proposition 3, future consumption mistakes, in the form of rules of thumb, lead to high current MPCs.

[9]This case is not directly nested in Proposition 3, since the actual consumption rule is stochastic. But the key results in Proposition 3 can be easily extended. See the proof of Corollary 4.

Corollary 4. *For $t \in \{0, \cdots, T-2\}$, $\phi_t^{Deliberate} \equiv \frac{\partial c_t^{Deliberate}}{\partial w_t} \geq \phi_t^{Frictionless}$ and $\phi_t^{Deliberate}$ increases with future selves' probabilities of following the rules of thumb $\{p_{t+k}\}_{k=1}^{T-t-1}$.*

This result means that, even when the current self is not subject to any behavioral bias on her own, the fact that future selves may follow rules of thumb raises current MPCs.

Hyperbolic discounting. My framework can also accommodate hyperbolic discounting (e.g. Laibson, 1997; Barro, 1999; Angeletos et al., 2001; Harris and Laibson, 2001). Consider a standard beta-delta model with sophistication and without borrowing constraints. In this case, the actual consumption rule is given by

$$c_t(w_t) = \arg\max_{c_t} u(c_t) + \delta\beta_t V_{t+1}(R(w_t - c_t)) \qquad \forall t \in \{0, \cdots, T-1\}, \tag{21}$$

where $\delta \in [0,1]$ is the standard discount factor, $\beta_t \in \left[\frac{1}{2}, 1\right]$ captures self t's present focus (a smaller β_t means a larger present focus), and $V_{t+1}(\cdot)$ is the continuation value function defined as in (10).[10] This actual consumption rule is the focus of the hyperbolic discounting literature. It combines the direct effect of present focus on current consumption with the effect of anticipated future present focus.

To isolate the impact of future present focus on current consumption, I define the deliberate consumption rule as usual:

$$c_t^{\text{Deliberate}}(w_t) = \arg\max_{c_t} u(c_t) + \delta V_{t+1}(R(w_t - c_t)) \qquad \forall t \in \{0, \cdots, T-1\}. \tag{22}$$

As a corollary of Proposition 3, these future consumption mistakes lead to high current deliberate MPCs.[11]

Corollary 5. *For $t \in \{0, \cdots, T-2\}$, $\phi_t^{Deliberate} \equiv \frac{\partial c_t^{Deliberate}}{\partial w_t} \geq \phi_t^{Frictionless}$ and $\phi_t^{Deliberate}$ increases with future selves' present focus, i.e., decreases with each $\{\beta_{t+k}\}_{k=1}^{T-t-1}$.*

In fact, in the environment here, high current MPCs under hyperbolic discounting come solely from the impact of future consumption mistakes. The current present focus, although it increases the *level* of current actual consumption, decreases current MPCs. To see this, using the FOC of actual consumption in (21) and taking a partial derivative with respect to w_t, we have

[10] The restriction $\beta_t \geq \frac{1}{2}$ makes sure that the comparative statics with respect to λs in Proposition 3 translate into comparative statics with respect to βs in Corollary 5.

[11] One can also derive Corollary 5 from the hyperbolic Euler equation in Harris and Laibson (2001). See Appendix B for details.

$$\phi_t = \frac{\partial c_t}{\partial w_t} = \frac{\delta R^2 \beta_t V''_{t+1}}{u'' + \delta R^2 \beta_t V''_{t+1}}.$$

That is, holding constant the concavity of the continuation value V''_{t+1}, the current MPC decreases with the degree of current present focus (i.e., increases with β_t). The intuition is that, with current present focus, the current self cares less about changes in the marginal value of saving and prefers to use savings instead of current consumption to absorb changes in w_t.

In the hyperbolic discounting literature, e.g., Laibson (1997) and Angeletos et al. (2001), liquidity constraints provide a reason why current present focus can increase current MPCs: since the current present focus makes the consumer's liquidity constraints more likely to bind in the future, she may display a high MPC. However, this mechanism is orthogonal to the motivating evidence on *high-liquidity* consumers' high MPCs.

Stochastic mistakes. The key result here does not require mistakes in future consumption to bias the consumer's behavior in a particular way. To illustrate, I let the actual consumption rule deviate from the deliberate one in a stochastic fashion. The actual consumption rule for each self $t \in \{0, \cdots, T-1\}$ is given by

$$c_t(w_t) = \phi_t w_t + \bar{c}_t = \left(\phi_t^{\text{Deliberate}} + \varphi_t\right) w_t + \bar{c}_t^{\text{Deliberate}} + \epsilon_t,$$

where the random variable φ_t captures the stochastic mistake in self t's actual MPC, the random variable ϵ_t captures the stochastic mistake in self t's actual consumption level, and $E[\varphi_t] = E[\epsilon_t] = 0$. These random variables are i.i.d. and independent of each other.[12]

The deliberate consumption rule is defined as usual:

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

not subject to any current behavioral bias and taking future selves' stochastic mistakes as given. Similar to Proposition 3, future stochastic consumption mistakes lead to high current MPCs.

Corollary 6. *For $t \in \{0, \cdots, T-2\}$, $\phi_t^{Deliberate} \equiv \frac{\partial c_t^{Deliberate}}{\partial w_t} \geq \phi_t^{Frictionless}$ and $\phi_t^{Deliberate}$ increases with the variances in future selves' actual MPCs $\{Var(\varphi_{t+k})\}_{k=1}^{T-t-1}$.*

This result means that, even if future selves' actual consumption may be unbiased on average, their stochastic consumption mistakes increase current MPCs. Moreover, as in the previous dis-

[12] This case is not directly nested in Proposition 3, since the actual consumption rule is stochastic. But the key results in Proposition 3 can be easily extended. See the proof of Corollary 6.

cussion, the key is future selves' stochastic mistakes in response to changes in savings, not their stochastic consumption levels.

Endogenous mistakes. 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. The above analysis applies verbatim for such a "stage-1."

There are two related points worth mentioning. First, because the utility loss of deviating from the optimal consumption rule is second order (Cochrane, 1989), the decision maker will always endogenously choose non-zero λs if the cognitive cost of not making mistakes is first order. In other words, the mistakes behind the key channel here can be prevalent. Second, for given λs, richer consumers (with lower $|u''|$) suffer smaller utility losses for a given degree of mistakes (λs). As a result, richer consumers will endogenously choose a higher degree of mistakes. Hence, the channel studied here may be particularly useful for explaining high-liquidity consumers' high MPCs. This observation is consistent with the discussion in Kueng (2018).

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 savings. With this knowledge and even without knowledge of the exact mistakes of her future selves, the consumer will have a higher current MPC as a second-best response to future consumption mistakes.

An inter-personal interpretation. In fact, most empirical evidence on MPCs is about consumption at the household level (e.g., Kueng, 2018; Fagereng, Holm and Natvik, 2019). Since household consumption is decided jointly by different members of the household, an alternative interpretation of my result is that the current self (e.g., the wife) is worried that the spending of future selves (the husband or the children) may respond inefficiently to changes in savings. The current self then displays a higher MPC.

4.4 The $T \to \infty$ limit and Gauging the Magnitudes

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

$$\phi_t^{Deliberate} \to \phi^{Deliberate} = \frac{\delta R^2 - 1}{\delta R^2 \left(1 - \lambda^2\right)}, \tag{23}$$

where the condition $|\lambda| < \left(\delta^{-1/2}R^{-1}\right)$ *guarantees that the transversality condition* $\lim_{k\to+\infty} \delta^k u'\left(c_{t+k}\right) = 0$ *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.

Gauging the magnitudes. The limit result in Corollary 7 helps us gauge how much anticipation of future consumption mistakes can impact current MPCs. In particular, one can use the standard calibration of a particular friction to calibrate λ and use (23) to gauge how much anticipation of this friction can increase the current MPCs. This exercise helps disentangle the channel of interest from the direct impact of this friction on current MPCs.

Consider the inattention example in Corollary 3. Of course, there is a caveat that attention to different objects differs (in fact Corollary 10 below studies such differential attention), but let me use the mean of the estimated attention in the literature review in Gabaix (2019), 0.44, to calibrate $\lambda = 1 - 0.44 = 0.56$ in (19). From (23), this implies that anticipation of future inattention can increase current MPCs by around 45%. In Appendix B, I also use Corollary 5 to map the standard present focus estimate $\beta = 0.504$ in Laibson et al. (2018) to $\lambda \approx 0.49$. This implies that anticipation of future hyperbolic discounting can increase current MPC by around 32%.

It is worth clarifying that the purpose of the calibration exercise is not to argue that the impact of future mistakes quantitively dominates the impact of current mistakes. The goal is to show that this channel can be sizable, and that it can help explain empirical puzzles on high-liquidity consumers' high MPCs.

4.5 Extensions and Discussion

Gradual resolution of uncertainty. Above, for illustration purposes, I assume that all income uncertainty in the economy is resolved in period 0. In fact, with quadratic utility here, the well known certainty equivalence result implies that the consumer's MPCs remain the same with gradual resolution of income uncertainty.

In the fungibility case here, with a graduate resolution of the income uncertainty, the actual and deliberate consumption rule of each self $t \in \{0, \cdots, T-1\}$ can now be written as a function of the expected permanent income. That is, $c_t(w_t)$ and $c_t^{\text{Deliberate}}(w_t)$ are given by

$$c_t(w_t) = \phi_t w_t + \bar{c}_t \quad \text{and} \quad c_t^{\text{Deliberate}}(w_t) = \phi_t^{\text{Deliberate}} w_t + \bar{c}_t^{\text{Deliberate}},$$

where $w_t = E_t\left[a_t + y_t + \sum_{k=1}^{T-t} R^{-k} y_{t+k} | (a_t, s_t)\right]$ now captures the expected permanent income based on period t's state (a_t, s_t). I still use λ_t to capture how self t's actual MPC ϕ_t deviates from the deliberate MPC $\phi_t^{\text{Deliberate}}$:

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

From future selves' actual consumption rules $\{c_{t+k}(w_{t+k})\}_{k=1}^{T-1-t}$, one can calculate current self t's deliberate consumption rule $c_t^{\text{Deliberate}}(w_t)$ and find her deliberate MPC $\phi_t^{\text{Deliberate}}$ as usual.

Corollary 8. *For $t \in \{0, \cdots, T-2\}$, $\phi_t^{Deliberate}$ shares the exact same formula as $\phi_t^{Deliberate}$ in Proposition 3.*

General concave utilities. In general, future consumption mistakes can take two forms: inefficient responses to changes in w_{t+k} and mistakes in the overall consumption level. In the quadratic case studied above, future selves' inefficient responses $\{\lambda_{t+k}\}_{k=1}^{T-t-1}$ robustly increase current MPCs. On the other hand, future selves' overall consumption levels $\{\bar{c}_{t+k}\}_{k=1}^{T-t-1}$ do not impact current MPCs.

With general concave utilities $u(\cdot)$ and $v(\cdot)$, the impact of future inefficient responses on current MPCs remains to be the same. To illustrate, consider the case in which each self's actual consumption responds inefficiently to changes in permanent income as in (14) but each self does not make "level" mistakes. That is, there is a path $\{\tilde{w}_t, \tilde{c}_t\}_{t=0}^{T-1}$ where the actual consumption level coincides with the deliberate consumption $\tilde{c}_t = c_t(\tilde{w}_t) = c_t^{\text{Delibrate}}(\tilde{w}_t)$. But actual consumption responds inefficiently to changes in permanent income away from this path. Similar to (14), I use

λ_t to capture self t's inefficient response. That is,[13]

$$\lambda_t \equiv 1 - \frac{\partial c_t\left(\tilde{w}_t\right)}{\partial w_t} / \frac{\partial c_t^{\text{Deliberate}}\left(\tilde{w}_t\right)}{\partial w_t}. \tag{24}$$

I can now re-establish that Proposition 3 in the case of general concave utility.

Proposition 4. *For each $t \in \{0, \cdots, T-1\}$,*

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

increases with each $\left\{\left|\lambda_{t+k}\right|\right\}_{k=1}^{T-t-1}$.

As a result, the above analysis about how future selves' inefficient responses to changes in permanent income lead to high current MPCs remains to hold in this more general setting.

An independent question is, with general utilities, whether the "level" mistakes in future consumption can also impact current MPCs. For this channel to be meaningful, it needs to come from the interaction with liquidity constraints. See Appendix B for more discussion.

Empirical support. Proposition 3 provides a potential explanation for the emerging empirical evidence on excess sensitivity for consumers with high liquidity. 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. Stephens and Unayama (2011), Parker (2017), Olafsson and Pagel (2018), Ganong and Noel (2019), McDowall (2020) also find that high-liquidity consumers can display high MPCs.

In regards to the key mechanism, there is also ample empirical evidence that consumers have at least partial knowledge about their future selves' mistakes and adjust behavior accordingly. For example, in the context of hyperbolic discounting, Allcott et al. (2020) find that the perceived and the actual present focus parameters are, respectively, 0.75 and 0.72. These values imply a degree of sophistication (s_t in (18)) close to 1.

[13] I assume u, v, and c_t are twice continuously differentiable.

5 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 above high MPCs result remains true: as long as future consumption responds inefficiently to changes in savings, 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. Finally, I illustrate how my framework can accommodate several behavioral biases causing inefficiently differential responses to different components of permanent income.

5.1 The Environment

In Section 4, 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.

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

$$c_t\left(a_t, s_t\right) = \phi_t^a a_t + \phi_t^y \left(y_t + \sum_{k=1}^{T-t} \omega_{t,k} R^{-k} y_{t+k}\right) + \bar{c}_t, \tag{25}$$

where ϕ_t^a captures the actual MPC out of wealth (i.e. savings/borrowings), 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, $\bar{c}_t$ in (25) 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 (25) 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 (25) 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 (27). Specifically, for all $t \in \{0, \cdots, T-1\}$ and $k \in \{0, \cdots, T-t\}$,

$$\phi_t^a = (1-\lambda_t^a)\,\phi_t^{\text{Deliberate}}, \quad \phi_t^y = \left(1-\lambda_{t,0}^y\right)\phi_t^{\text{Deliberate}}, \quad \text{and} \quad \phi_t^y\omega_{t,k} = \left(1-\lambda_{t,k}^y\right)\phi_t^{\text{Deliberate}}\omega_{t,k}^{\text{Deliberate}}, \tag{26}$$

where λ_t^a captures the mistake in self t's actual MPC out of wealth (i.e. savings/borrowings), $\lambda_{t,0}^y$ 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 (14), 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.

Based on Definition 1 and 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:*

$$c_t^{Deliberate}(a_t, s_t) = \phi_t^{Deliberate}\left(a_t + y_t + \sum_{k=1}^{T-t}\omega_{t,k}^{Deliberate}R^{-k}y_{t+k}\right) + \bar{c}_t^{Deliberate}, \tag{27}$$

where $\phi_t^{Deliberate}$ is a function of $\left\{\lambda_{t+l}^a\right\}_{l=1}^{T-t-1}$, δ, R, and $\omega_{t,k}$ is a function of $\left\{\lambda_{t+l}^a\right\}_{l=1}^{T-t-1}$, $\left\{\lambda_{t+l,k-l}^y\right\}_{l=1}^{\min\{k,\ T-t-1\}}$, δ, R.

In (27), $\phi_t^{\text{Deliberate}}$ 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 savings/borrowings $\left(\lambda_{t+l}^a \neq 0\right)$, current deliberate MPCs, i.e., $\phi_t^{\text{Deliberate}}$ in (27), will be higher. Second, non-fungibility of future consumption ($\lambda_{t+l}^a \neq \lambda_{t+l,k-l}^y$) suffices to generate the non-fungibility of current deliberate consumption ($\omega_{t,k}^{\text{Deliberate}} \neq 1$). 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.

5.2 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 savings/borrowings.

Similar to Lemma 1, I use $\Gamma_{t+1} > 0$ to denote the "concavity" of the consumer's continuation value function in (10): $\frac{\partial^2 V_{t+1}(a_{t+1}, s_{t+1})}{\partial a_{t+1}^2} \equiv u'' \cdot \Gamma_{t+1}$.

Proposition 5.

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

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

The intuition behind part (i) of Proposition 5 is similar to Lemma 1. Larger $\left\{\left|\lambda_{t+l}^a\right|\right\}_{l=0}^{T-t-1}$ means more inefficient future consumption responses to changes in savings/borrowings. As a result, the marginal value of savings $\frac{\partial V_{t+1}(a_{t+1}, s_{t+1})}{\partial a_{t+1}}$ 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 $\left\{\lambda_{t+l}^a\right\}_{l=0}^{T-t-1}$ are inefficient responses of future consumption to changes in savings/borrowings. This is because these responses directly determine the marginal value of savings $\frac{\partial V_{t+1}(a_{t+1}, s_{t+1})}{\partial a_{t+1}}$ and hence the concavity Γ_t. On the other hand, Γ_{t+1} is independent of $\lambda_{t+l,k-l}^y$ for all l and k.

The intuition behind part (ii) of Proposition 5 is similar to Proposition 3. From part (i), with future consumption mistakes (larger$\left\{\left|\lambda_{t+l}^a\right|\right\}_{l=1}^{T-t-1}$), 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 savings. She hence displays a higher MPC.

Similar to part (i), the relevant mistakes $\left\{\lambda_{t+l}^a\right\}_{l=1}^{T-t-1}$ for the high current MPCs result are future selves' inefficient responses to changes in savings/borrowings. On the other hand, $\phi_t^{\text{Deliberate}}$ is independent of $\lambda_{t+l,k-l}^y$ 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. For example, in the context of inattention, Corollary 10 below shows how future imperfect perception of wealth (i.e., savings/borrowings) increases current MPCs. On the other hand, if future selves are *only* inattentive to income, the current MPCs will not be influenced. In Appendix B, I also use this result to study when future selves' distorted expectations (Azeredo da Silveira and Woodford, 2019; Bordalo, Gennaioli and Shleifer, 2018; Bordalo et al., 2020) will lead to high current MPCs. For example, if future selves over-extrapolate based on their wealth, current MPCs will be higher.

5.3 Future Non-fungibility Begets Current Non-fungibility

Now, I turn to a new prediction.

Proposition 6. *Generically, the deliberate consumption in (27) 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,

$$\lambda_{t+l} = \lambda_{t+l}^{a} = \lambda_{t+l,k-l}^{y} \quad \forall l, k. \tag{28}$$

In this case, the current deliberate consumption remains to follow the fungibility principle. Away from (28), generically, current deliberate consumption will violate the fungibility principle.

Excess discounting. To better understand the intuition behind Proposition 6, 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 7. *Consider the case that* $\lambda^a_{t+l} \geq \lambda^y_{t+l,k-l}$ *and* $\lambda^a_{t+l} \geq 0$ *for all* $l \in \{1, \cdots, T-t-1\}$ *and* $k \in \{l, \cdots, T-t+l\}$.

The current deliberate consumption in (27) has the following properties: for $k \in \{0, \cdots, T-t\}$,

(i) $\omega^{Deliberate}_{t,k} \leq 1$. *That is, the current self excessively discounts future income.*

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

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

Proposition 7 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 7, note that, when future selves mistakenly respond too little to changes in savings/wealth (a larger $\left\{\lambda^a_{t+l}\right\}_{l=1}^{T-t-1}$), the excess concavity in Proposition 6 means that the current self will be less willing to change her savings. 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 savings. Hence, there is excess discounting ($\omega^{\text{Deliberate}}_{t,k} < 1$) and $\omega^{\text{Deliberate}}_{t,k}$ decreases with each $\left\{\lambda^a_{t+l}\right\}_{l=1}^{T-t-1}$.

On the other hand, $\omega^{\text{Deliberate}}_{t,k}$ increases with each $\left\{\lambda^y_{t+k,l-k}\right\}_{k=1}^{\min\{l,\ T-t-1\}}$. That is, if future selves' mistakenly respond too little to changes in future income y_{t+k} (a larger $\left\{\lambda^y_{t+k,l-k}\right\}_{k=1}^{\min\{l,\ T-t-1\}}$), 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 7 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).[14]

5.4 Extensions

Gradual resolution of uncertainty. Above, for illustration purposes, I assume that all income uncertainty in the economy is resolved in period 0. In fact, similar to Corollary 8, the consumer's MPC remains the same with gradual resolution of income uncertainty. See Corollary 16 in Appendix B.

The $T \to \infty$ and hand-to-mouth limit. Similar to Corollary 7, I can establish a simple limit for the deliberate consumption rule in (27) when the consumer's horizon T goes to infinity.

Corollary 9. *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 $T \to +\infty$,*

$$\phi_t^{Deliberate} \to \phi^{Deliberate} \equiv \frac{\delta R^2 - 1}{\delta R^2 \left(1 - (\lambda^a)^2\right)}, \tag{29}$$

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

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

$$\phi^{Deliberate} \to 1 \quad and \quad \omega^{Deliberate} \to 0. \tag{30}$$

The limit in (30) 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.

[14] 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 6 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 7.

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.

5.5 Micro-Foundations of Non-fungibility

The results in Propositions 5—7 do not depend on the exact behavioral causes of future consumption mistakes. Here, I illustrate how my framework can easily accommodate several widely studied behavioral biases that naturally cause inefficiently differential responses of future consumption to different components of permanent income. The biases studied in the previous fungible section can also be extended to the non-fungible case here.

Mental accounting. Thaler (1990) provides evidence that consumers systematically violate the fungibility principle. He proposes that consumers have separate mental accounts for current income, expected future income, and wealth. As a result, consumption exhibits different MPCs out of changes in these separate mental accounts. Mental accounting then provides a direct micro-foundation for different λs in (26). That is, why future consumption may exhibit differential responses to different components of income and wealth. The results in Propositions 5— 7 follow directly. In other words, the fact that future selves have separate mental accounts, by itself, suffices to generate the non-fungibility of current consumption.

Differential inattention to income and wealth. In Corollary 3, I accommodate inattention within the fungible framework in Section 4. There, actual consumption is decided based on the same degree of attention to all components of permanent income, as in (19). In the non-fungible framework here, I can accommodate different degrees of attentions to different components of income and wealth. Below I study the rather "overlooked" case in which the consumer is inattentive to her endogenous wealth a_t but attentive to her income state s_t. This is the focus of the job market version of Lian (2019), which the current, more general, paper replaces. In Appendix B, I study the more "familiar" case in the literature (e.g. Luo, 2008; Gabaix, 2016, 2019), where the consumer is inattentive to her income state s_t but attentive to her endogenous wealth a_t.

Imperfect perception of wealth. Here I study the case in which actual consumption is determined with inattention to wealth/savings a_t, but full attention to the income state s_t. Specifically, similar to Corollary 3 above, I follow the sparsity approach in Gabaix (2014) and let each self's perceived wealth be given by a weighted average of her actual wealth and a default. To isolate the friction of interest, I let each self perfectly perceive her current income state s_t. That is, for $t \in \{0, \cdots, T-1\}$,

$$a_t^p(a_t) = (1-\lambda_t^a) a_t + \lambda_t^a a_t^d \quad \text{and} \quad s_t^p(s_t) = s_t, \tag{31}$$

where $\lambda_t^a \in [0,1]$ captures self t's degree of imperfect perception of wealth (a larger λ_t^a means more inattention) and a_t^d captures the default (an exogenous constant of whose value does not matter for the MPCs). Also similar to Corollary 3, an alternative way to model inattention is through noisy signals (Sims, 2003). With linear consumption rules and Normally distributed incomes, the two approaches still lead to the same predictions on MPCs.

There is ample empirical support for imperfect perception of wealth 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. Brunnermeier and Nagel (2008) and Alvarez, Guiso and Lippi (2012) find that consumers often have imperfect knowledge of changes in their financial wealth and fail to adjust accordingly. Moreover, the recent literature on Fintechs shows that providing information about a consumer's total wealth 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 wealth. Participants significantly change their consumption and saving after seeing their wealth, implying that they have an imperfect perception of wealth without the tool. Likewise, Carlin, Olafsson and Pagel (2017) study a financial app that consolidates all of its users' bank account information and transaction histories. They show that the app significantly reduces its users' interest expenses on consumer debt and other bank fees.

Based on the perceived wealth in (31), the actual consumption rule of each self t is given by

$$c_t(a_t^p(a_t), s_t) = \arg\max_{c_t} \; u(c_t) + \delta V_{t+1}(R(a_t^p(a_t) + y_t - c_t), s_{t+1}), \tag{32}$$

where the continuation value function V_{t+1} is defined as in (10), based on future selves' actual consumption rules.

Deliberate consumption is decided as usual, isolating the impact from future selves' imperfect perception of wealth:

$$c_t(a_t, s_t) = \arg\max_{c_t} \; u(c_t) + \delta V_{t+1}(R(a_t + y_t - c_t), s_{t+1}).$$

Here, future selves' imperfect perception of wealth leads to inefficient responses of future consumption to changes in wealth. As discussed in Propositions 5—7, these mistakes lead to high current MPCs and excess discounting of future income.

Corollary 10. *For each $t \in \{0, \cdots, T-1\}$, self t's deliberate consumption rule is given by (27), where $\phi_t^{Deliberate}$ and $\left\{\omega_{t,k}^{Deliberate}\right\}_{l=0}^{T-t}$ are given by the formula in Lemma 2 with λ_{t+l}^a given by (31) and all λ^ys are zero. Moreover,*

(i) $\phi_t^{Deliberate} \geq \phi_t^{Frictionless}$ and $\phi_t^{Deliberate}$ increases with future selves' degrees of imperfect perception of wealth $\left\{\lambda_{t+l}^a\right\}_{l=1}^{T-t-1}$.

(ii) For $k \in \{0, \cdots, T-t\}$, $\omega_{t,k}^{Deliberate} \leq 1$ and $\omega_{t,k}^{Deliberate}$ decreases with future selves' degrees of imperfect perception of wealth $\left\{\lambda_{t+l}^a\right\}_{l=1}^{T-t-1}$.

Furthermore, because the consumer here is fully attentive to her current income state, the above properties of the deliberate MPCs out of current and future income naturally translate to properties of the actual MPCs.

Corollary 11. *For each $t \in \{0, \cdots, T-1\}$, self t's actual consumption rule in (25) has the following properties:*

(i) The MPC out of current income $\phi_t^y \geq \phi_t^{Frictionless}$ and ϕ_t^y increases with future selves' degrees of imperfect perception of wealth $\left\{\lambda_{t+l}^a\right\}_{l=1}^{T-t-1}$.

(ii) The MPC out of wealth is given by $\phi_t^a = (1-\lambda_t^a)\,\phi_t^y$.

(iii) For $k \in \{0, \cdots, T-t\}$, there is extra discounting of future income $\omega_{t,k} \leq 1$ and $\omega_{t,k}$ decreases with future selves' degrees of imperfect perception of wealth $\left\{\lambda_{t+l}^a\right\}_{l=1}^{T-t-1}$.

In other words, Corollary 11 shows how current and anticipated future imperfect perception of wealth provide a unified explanation of Thaler (1990)'s three key observations about how consumption deviates from the prediction of the permanent income hypothesis: excess sensitivity to current income, a smaller MPC out of wealth than out of current income, and excess discounting of future income.

It is worth noting that the relevant mistake here is imperfect perception of *changes* in wealth. The consumer can have good knowledge of her overall wealth. For example, for the high MPC result in (i), the key is: if the current self increases savings, her future selves have difficulty perfectly perceiving the increase; future selves nevertheless can perfectly perceive overall saving levels.

In terms of magnitudes, I can use the limit result in Corollary 9 to gauge how much anticipation of future imperfect perception of wealth, by itself, can increase the MPC out of current income.

Here, because direct estimates of imperfect perception of wealth are not necessarily available, I instead back out λ^a from relevant moments of MPCs in the data.

Specifically, in the $T \to \infty$ studied in Corollary 11, $\phi^a/\phi^y = 1 - \lambda^a$.[15] This ratio ϕ^a/ϕ^y between the MPC out of wealth and the MPC out of current income is directly available from empirical studies. For example, Di Maggio, Kermani and Majlesi (2018) estimate the MPC out of wealth and the MPC out of current income for rich households away from liquidity constraints. In their estimates, for consumers in the top half of wealth distribution, the MPC out of wealth is \$0.05 per year, and the MPC out of current income for rich households is \$0.35 per year. Together, these values imply $\lambda^a = 1 - 1/7 = 6/7$. In fact, their estimates reflect a general theme in the recent empirical literature: the estimates of MPC out of wealth are typically much smaller than the estimates of the MPC out of current income.[16] Based on this estimated friction λ^a, the anticipation of future imperfect perception of wealth can increase the current MPC by as much as 2.77 times.

6 Other Applications

The main application in this paper is to show that future consumption mistakes can explain high-liquidity consumers' high MPCs and non-fungibility. The key mechanism behind these results, i.e., the excess concavity of the continuation value function driven by future consumption mistakes, can also speak to other well-known puzzles in intertemporal decisions. First, the large risk aversion and equity premium puzzle (Mehra and Prescott, 1985). Second, the small elasticity of intertemporal substitution and 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. Using the fungibility case in Section 4 as an example: 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 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

[15] This relationship assumes perfect attention to the income state s_t ($\lambda^y = 0$) as in (31). For a given ϕ^a/ϕ^y, if I allow inattention to income state, the implied degree of imperfect perception of wealth (λ^a) will be larger.

[16] Using other estimates of the MPC out of wealth and the MPC out of current income, I can get similar, if not larger, estimates of the ratio λ^a. 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 savings rate out of capital gains close to one hundred percent.

excess concavity of the value function. We then know that consumption mistakes will also lead to a larger risk aversion.

To gauge the magnitudes of how much consumption mistakes can increase risk aversion, let us again use the $T \to +\infty$ limit in Corollary 7. In this limit, we have $\Gamma_t \equiv \frac{\partial^2 V_t}{\partial w_t^2}/u'' \to \Gamma = \frac{\delta R^2 - 1}{\delta R^2 (1 - \delta R^2 \lambda^2)}$. With the calibration of λ used in Section 4 ($\lambda = 0.56$ for inattention or $\lambda = 0.49$ for hyperbolic discounting) and standard calibration of δ and R (closer to 1), consumption mistakes can increase the degree of risk aversion by $30\% - 50\%$.

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, can also help resolve this puzzle.

The intuition is similar: the response of current consumption to interest rate changes leads to changes in savings; with future consumption mistakes in response to changes in savings, 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.[17]

Proposition 8. *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* $\left\{\left|\lambda_{t+k}^a\right|\right\}_{k=1}^{T-t-1}$.

7 Conclusion

In this paper, I show how inefficient responses of future consumption to changes in savings leads 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.

[17] 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 MPCs result in response to income changes.

Appendix A: Proofs

Proof of Proposition 1. Based on (5), we have

$$u'\left(c_0^{\text{Deliberate}}(y_0)\right) = \frac{1}{2}(1-\lambda_1)\,u'\left(\frac{1}{2}(1-\lambda_1)\,a_1+\bar{c}_1\right)+\frac{1}{2}(1+\lambda_1)\,u'\left(\frac{1}{2}(1+\lambda_1)\,a_1-\bar{c}_1\right),$$

where $a_1 = y_0 - c_0^{\text{Deliberate}}(y_0)$. Since u is quadratic, we know $c_0^{\text{Deliberate}}(y_0)$ 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 (8) at t together with the definition of the value function in (10) at $t+1$ lead to (11). The recursive formulation for the value function in (12) follows directly from the definition of the value function in (10).

Now, consider consumption rules and value functions $\left\{c_t^{\text{Deliberate}}(a_t,s_t),\,c_t(a_t,s_t)\right\}_{t=0}^{T-1}$ and $\{V_t(a_t,s_t)\}_{t=0}^{T}$ satisfy (9), (11), (12), and the boundary condition $V_T(a_T,s_T) = v(a_T+y_T)$. Since I am working with a finite horizon problem, I can iterate those conditions through backward induction and arrive at the sequential form in (8) – (10).

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 (11), the deliberate MPC is given by

$$\phi_t^{\text{Deliberate}} = \frac{\delta R^2\Gamma_{t+1}}{1+\delta R^2\Gamma_{t+1}}. \tag{33}$$

From (14), the actual MPC is given by

$$\phi_t = \frac{(1-\lambda_t)\,\delta R^2\Gamma_{t+1}}{1+\delta R^2\Gamma_{t+1}}. \tag{34}$$

From the recursive formulation of the value function in (11), we have:

$$\frac{\partial V_t(w_t)}{\partial w_t} = \phi_t u'(c_t(w_t)) + (1-\phi_t)\,\delta R \frac{\partial V_{t+1}(w_{t+1})}{\partial w_{t+1}}. \tag{35}$$

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

$$\begin{aligned}\Gamma_t &= (\phi_t)^2 + (1-\phi_t)^2\,\Gamma_{t+1}\delta R^2 \\ &= \left(1+\Gamma_{t+1}\delta R^2\right)\left(\phi_t - \frac{\delta R^2\Gamma_{t+1}}{1+\delta R^2\Gamma_{t+1}}\right)^2 + \frac{\delta R^2\Gamma_{t+1}}{1+\delta R^2\Gamma_{t+1}} \\ &= \frac{\left(\delta R^2\Gamma_{t+1}\right)^2}{1+\delta R^2\Gamma_{t+1}}\lambda_t^2 + \frac{\delta R^2\Gamma_{t+1}}{1+\delta R^2\Gamma_{t+1}}.\end{aligned} \tag{36}$$

Lemma 1 and Proposition 3 then follow directly.

Proof of Corollary 1 and Corollary 2. Let

$$\left\{\tilde{c}_{t,t+k}(w_{t+k}) = \tilde{\phi}_{t,t+k} w_{t+k} + \tilde{\bar{c}}_{t,t+k}\right\}_{k=1}^{T-t-1} \tag{37}$$

capture self t's perceived future consumption rules. I redefine the deliberate consumption based on these perceived future consumption rules:

$$\begin{aligned}c_t^{\text{Deliberate}}(w_t) &\equiv \arg\max_{c_t} u(c_t) + \sum_{k=1}^{T-t-1} \delta^{k-1} u\left(\tilde{c}_{t,t+k}(w_{t+k})\right) + \delta^{T-t} v(w_T) \\ &= \phi_t^{\text{Deliberate}} w_t + \bar{c}_t^{\text{Deliberate}},\end{aligned} \tag{38}$$

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

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

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

subject to the budget. We can then define self t's perceived future mistake $\tilde{\lambda}_{t,t+k}$ (in response to changes in permanent income) similar to (14):

$$\tilde{\phi}_{t,t+k} = \left(1 - \tilde{\lambda}_{t,t+k}\right) \tilde{\phi}^{\text{Deliberate}}_{t,t+k}. \tag{39}$$

Based on (37) – (39), 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 1 then follows. Corollary 2 then follows directly from (18).

Proof of Corollary 3. From (20), we have

$$u'\left(c_t\left(w_t\right)\right) = \delta R V'_{t+1}\left(R\left(w^p_t\left(w_t\right) - c_t\left(w_t\right)\right)\right),$$

while

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

Because both u and V_{t+1} are quadratic, u' and V'_{t+1} are linear. Together with (19), we know, in this case, $\phi_t = \left(1 - \lambda_t\right) \phi^{\text{Deliberate}}_t$, where λ_t is the degree of inattention in (19). Corollary 3 then follows directly.

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

The value function in (12) is now given by

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

where $E_t\left[\cdot\right]$ averages over the potential realizations of actual consumption rule. The deliberate consumption in (11) is unchanged.

In the proof of Proposition 3, the deliberate MPC is still given by (11), but (12) becomes

$$\begin{aligned}
\Gamma_t &= p_t\left[\left(\phi^R_t\right)^2 + \left(1 - \phi^R_t\right)^2 \Gamma_{t+1}\delta R^2\right] + \left(1 - p_t\right)\left[\left(\phi^{\text{Deliberate}}_t\right)^2 + \left(1 - \phi^{\text{Deliberate}}_t\right)^2 \Gamma_{t+1}\delta R^2\right] \\
&= p_t\left[\left(\phi^R_t\right)^2 + \left(1 - \phi^R_t\right)^2 \Gamma_{t+1}\delta R^2\right] + \left(1 - p_t\right)\frac{\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}}, \\
&= \frac{\left(\delta R^2 \Gamma_{t+1}\right)^2}{1 + \delta R^2 \Gamma_{t+1}} p_t \lambda^2_t + \frac{\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}}.
\end{aligned}$$

where $\lambda_t = 1 - \frac{\phi_t^R}{\phi_t^{\text{Deliberate}}}$. As a result, Γ_t increases with p_t and Γ_{t+1} (and thus $\{p_{t+k}\}_{k=1}^{T-t-1}$). Corollary 4 then follows directly from (11).

Proof of Corollary 5. From (21) and (22), we have

$$u'\left(c_t\left(w_t\right)\right) = \delta\beta_t R V_{t+1}'\left(R\left(w_t - c_t\left(w_t\right)\right)\right),$$

and

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

Because both u and V_{t+1} are quadratic, u' and V_{t+1}' are linear. We then have

$$\phi_t = \frac{\delta\beta_t R^2 \Gamma_{t+1}}{1 + \delta\beta_t R^2 \Gamma_{t+1}} \quad \text{and} \quad \phi_t^{\text{Deliberate}} = \frac{\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}}, \tag{40}$$

where, as in Lemma 1, $\Gamma_{t+1} = \frac{\partial^2 V_{t+1}(w_{t+1})}{\partial w_{t+1}^2} / u''$. As a result,

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

where

$$\lambda_t = \frac{1 - \beta_t}{1 + \delta\beta_t R^2 \Gamma_{t+1}}. \tag{41}$$

Substitute into (36), we have

$$\Gamma_t = \frac{\left(\delta R^2 \Gamma_{t+1}\right)^2}{1 + \delta R^2 \Gamma_{t+1}} \left(\frac{1 - \beta_t}{1 + \delta\beta_t R^2 \Gamma_{t+1}}\right)^2 + \frac{\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}}. \tag{42}$$

Define $f\left(x, \beta\right) = \frac{x^2}{1+x}\left(\frac{1-\beta}{1+\beta x}\right)^2 + \frac{x}{1+x}$. We have $f_\beta\left(x, \beta\right) = -\frac{2(1-\beta)x^2}{(1+x\beta)^3}$ and $f_x\left(x, \beta\right) = \frac{1+x\beta(-1+2\beta)}{(1+x\beta)^3}$. As a result, for $x \geq 0$ and $\beta \in [\frac{1}{2}, 1]$, we have $f_\beta\left(x, \beta\right) \leq 0$ and $f_x\left(x, \beta\right) \geq 0$. Using these properties in (42), we know Γ_t decreases with $\{\beta_{t+k}\}_{k=0}^{T-t-1}$. Corollary 5 then follows from (3).

Proof of Corollary 6. This case is not directly nested in Proposition 3, as the actual consumption rule is stochastic. But the proof is essentially unchanged.

The value function in (12) is now given by

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

where $E_t\left[\cdot\right]$ averages over the potential realizations of actual consumption rule. The deliberate consumption in (11) is unchanged.

In the proof of Proposition 3, the deliberate MPC is still given by (33), but (36) becomes

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

As a result, Γ_t increases with $\{Var(\varphi_{t+k})\}_{k=0}^{T-t-1}$. Corollary 6 then follows directly from (33).

Proof of Corollary 7. From (36), we know that $\Gamma_t = \frac{\left(\delta R^2 \Gamma_{t+1}\right)^2}{1+\delta R^2 \Gamma_{t+1}}\lambda^2 + \frac{\delta R^2 \Gamma_{t+1}}{1+\delta R^2 \Gamma_{t+1}} \equiv f\left(\Gamma_{t+1}\right)$, 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 $0 \leq \lambda < \delta^{-1/2} R^{-1}$, 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 $\Gamma > \Gamma_t = f^{(T-t)}\left(\Gamma_T\right) > f^{(T-t-1)}\left(\Gamma_T\right) > \cdots > \frac{v''}{u''} = \Gamma_T$. As a result, $\Gamma_t = f^{(T-t)}\left(\Gamma_T\right)$ converges to the fix point Γ with $T \to +\infty$.

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

Together, one way or another, as long as $0 \leq \lambda < \delta^{-1/2} R^{-1}$, $\Gamma_t \to \Gamma$ with $T \to +\infty$. From (33), 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 8. With graduate resolution of uncertainty, the optimal deliberate consumption in (11) becomes

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

while the recursive formulation for the value function in (12) becomes

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

where $E_t[\cdot] = E_t\left[\cdot \mid (a_t, s_t)\right]$ captures rational expectations based on period t's state (a_t, s_t).

The proof of Proposition 2 remains unchanged, except (35) becomes

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

In particular, the formula (33), (35), and (36) remain unchanged. So Corollary 8 follows directly.

Proof of Proposition 4. The recursive formulation in Proposition 2 remains to hold. Because I assume u, v, and c_t are twice continuously differentiable, V_t is twice continuously differentially too.

The optimal deliberate consumption now is given by[18]

$$u'\left(c_t^{\text{Deliberate}}(w_t)\right) = R\delta V_{t+1}'\left(R\left(w_t - c_t^{\text{Deliberate}}(w_t)\right)\right). \tag{43}$$

We henceforth have:

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

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

$$\frac{\partial c_t^{\text{Deliberate}}(\tilde{w}_t)}{\partial w_t} = \frac{R^2\delta\frac{\partial^2 V_{t+1}(\tilde{w}_{t+1})}{\partial w_{t+1}^2}}{u''\left(c_t^{\text{Deliberate}}(\tilde{w}_t)\right) + R^2\delta\frac{\partial^2 V_{t+1}(\tilde{w}_{t+1})}{\partial w_{t+1}^2}}. \tag{44}$$

From (11):

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

As a result,

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

and

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

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

At $\tilde{w}_t$, because $c_t(\tilde{w}_t) = c_t^{\text{Deliberate}}(\tilde{w}_t) = \tilde{c}_t$, from (43), we have $u'(c_t(\tilde{w}_t)) = \delta R \frac{\partial V_{t+1}(\tilde{w}_{t+1})}{\partial w_{t+1}}$. As a result,

$$\frac{\partial^2 V_t(\tilde{w}_t)}{\partial w_t^2} = \left(\frac{\partial c_t(\tilde{w}_t)}{\partial w_t}\right)^2 u''(c_t(\tilde{w}_t)) + \left(1 - \frac{\partial c_t(\tilde{w}_t)}{\partial w_t}\right)^2 \delta R^2 \frac{\partial^2 V_{t+1}(\tilde{w}_{t+1})}{\partial w_{t+1}^2}. \tag{45}$$

Define $\Gamma_t \equiv \frac{\partial^2 V_t(\tilde{w}_t)}{\partial w_t^2} / u''(c_t(\tilde{w}_t))$, $\phi_t^{\text{Deliberate}} \equiv \frac{\partial c_t^{\text{Deliberate}}(\tilde{w}_t)}{\partial w_t}$, and $\phi_t \equiv \frac{\partial c_t(\tilde{w}_t)}{\partial w_t} \equiv (1-\lambda_t)\frac{\partial c_t^{\text{Deliberate}}(\tilde{w}_t)}{\partial w_t}$. From (44) and (45), we have

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

and

$$\Gamma_t = \phi_t^2 + (1-\phi_t)^2 \Gamma_{t+1} \delta R^2 \frac{u''(\tilde{c}_{t+1})}{u''(\tilde{c}_t)}.$$
$$= \frac{\left(\delta R^2 \Gamma_{t+1} \frac{u''(\tilde{c}_{t+1})}{u''(\tilde{c}_t)}\right)^2}{1 + \delta R^2 \Gamma_{t+1} \frac{u''(\tilde{c}_{t+1})}{u''(\tilde{c}_t)}} \lambda_t^2 + \frac{\delta R^2 \Gamma_{t+1} \frac{u''(\tilde{c}_{t+1})}{u''(\tilde{c}_t)}}{1 + \delta R^2 \Gamma_{t+1} \frac{u''(\tilde{c}_{t+1})}{u''(\tilde{c}_t)}}.$$

Proposition 4 then follows.

Proof of Lemma 2. Similar to (15), we define $\left\{\Gamma_t, \Gamma_{t,k}^y\right\}_{t\in\{0,\cdots,T\},k\in\{0,\cdots,T-t\}}$ based on

$$\frac{\partial V_t}{\partial a_t} \equiv u'' \cdot \left(\Gamma_t a_t + \sum_{k=0}^{T-t} \Gamma_{t,k}^y R^{-k} y_{t+k} + \bar{\Gamma}_t\right). \tag{46}$$

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 $t \leq T-1$, from (11), 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 (46) at $t+1$, we have

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

with

$$\phi_t^{\text{Deliberate}} = \frac{\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}} \tag{47}$$

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

$$\omega_{t,k}^{\text{Deliberate}} = \frac{\delta R R^{-(k-1)} \Gamma_{t+1,k-1}^y}{1 + \Gamma_{t+1} \delta R^2} / \left(\phi_t^{\text{Deliberate}} R^{-k}\right) = \frac{\Gamma_{t+1,k-1}^y}{\Gamma_{t+1}}. \tag{48}$$

Now, from the recursive formulation of the value function in (11), we have:

$$\frac{\partial V_t\left(a_t, s_t\right)}{\partial a_t} = \phi_t^a u'\left(c_t\left(a_t, s_t\right)\right) + \left(1 - \phi_t^a\right) \delta R \frac{\partial V_{t+1}\left(a_{t+1}, s_{t+1}\right)}{\partial a_{t+1}}. \tag{49}$$

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

$$\begin{aligned}\Gamma_t a_t + \sum_{k=0}^{T-t} \Gamma_{t,k}^y R^{-k} y_{t+k} + \bar{\Gamma}_t = {}& \left(\phi_t^a - \left(1 - \phi_t^a\right) \delta R^2 \Gamma_{t+1}\right) \left(\phi_t^a a_t + \phi_t^y \left(y_t + \sum_{k=1}^{T-t} \omega_{t,k} R^{-k} y_{t+k}\right) + \bar{c}_t\right) \\ &+ \left(1 - \phi_t^a\right) \delta R \left(\Gamma_{t+1} R\left(a_t + y_t\right) + \sum_{k=0}^{T-t-1} \Gamma_{t+1,k}^y R^{-k} y_{t+1+k} + \bar{\Gamma}_{t+1}\right).\end{aligned}$$

Together with (26), we have, for all $t \in \{0, \cdots, T-1\}$:

$$\begin{aligned}\Gamma_t &= \phi_t^a \left(\phi_t^a - \left(1 - \phi_t^a\right) \delta R^2 \Gamma_{t+1}\right) + \left(1 - \phi_t^a\right) \delta R^2 \Gamma_{t+1} \\ &= \left(1 + \delta R^2 \Gamma_{t+1}\right) \left(\phi_t^a - \frac{\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}}\right)^2 + \frac{\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}} \\ &= \frac{\left(\delta R^2 \Gamma_{t+1}\right)^2}{1 + \delta R^2 \Gamma_{t+1}} \left(\lambda_t^a\right)^2 + \frac{\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}},\end{aligned} \tag{50}$$

and

$$\begin{aligned}
\Gamma^y_{t,0} &= \phi^y_t \left(\phi^a_t - (1-\phi^a_t)\,\delta R^2 \Gamma_{t+1}\right) + (1-\phi^a_t)\,\delta R^2 \Gamma_{t+1} \\
&= \left(1+\delta R^2\Gamma_{t+1}\right)\left(\phi^a_t - \frac{\delta R^2 \Gamma_{t+1}}{1+\delta R^2\Gamma_{t+1}}\right)\left(\phi^y_t - \frac{\beta R^2\Gamma_{t+1}}{1+\beta R^2\Gamma_{t+1}}\right) + \frac{\delta R^2\Gamma_{t+1}}{1+\delta R^2\Gamma_{t+1}} \\
&= \frac{\left(\delta R^2\Gamma_{t+1}\right)^2}{1+\delta R^2\Gamma_{t+1}}\lambda^a_t\lambda^y_{t,0} + \frac{\delta R^2\Gamma_{t+1}}{1+\delta R^2\Gamma_{t+1}}, \qquad (51)
\end{aligned}$$

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

$$\begin{aligned}
\Gamma^y_{t,k} &= \phi^y_t \omega_{t,k}\left(\phi^a_t - (1-\phi^a_t)\,\delta R^2\Gamma_{t+1}\right) + (1-\phi^a_t)\,\delta R^2\Gamma^y_{t+1,k-1} \\
&= \left(1+\delta R^2\Gamma_{t+1}\right)\left(\phi^a_t - \frac{\delta R^2\Gamma_{t+1}}{1+\delta R^2\Gamma_{t+1}}\right)\left(\phi^y_t\omega_{t,k} - \frac{\delta R^2\Gamma^y_{t+1,k-1}}{1+\delta R^2\Gamma_{t+1}}\right) + \frac{\delta R^2\Gamma^y_{t+1,k-1}}{1+\delta R^2\Gamma_{t+1}} \\
&= \frac{\left(\delta R^2\right)^2\Gamma_{t+1}\Gamma^y_{t+1,k-1}}{1+\delta R^2\Gamma_{t+1}}\lambda^a_t\lambda^y_{t,k} + \frac{\delta R^2\Gamma^y_{t+1,k-1}}{1+\delta R^2\Gamma_{t+1}}. \qquad (52)
\end{aligned}$$

Lemma 2 follows from (47) – (52).

Proof of Proposition 5. 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 5 then follows directly from Lemma 1 and Proposition 3.

Proof of Proposition 6. From (48), (50), and (51), for $t \in \{0,\cdots,T-2\}$, we have

$$\omega^{\text{Deliberate}}_{t,1} = \frac{\delta R^2\Gamma_{t+2}\lambda^a_{t+1}\lambda^y_{t+1,0}+1}{\delta R^2\Gamma_{t+2}\left(\lambda^a_{t+1}\right)^2+1} = 1 - \frac{\delta R^2\Gamma_{t+2}\lambda^a_{t+1}\left(\lambda^a_{t+1}-\lambda^y_{t+1,0}\right)}{\delta R^2\Gamma_{t+2}\left(\lambda^a_{t+1}\right)^2+1}, \qquad (53)$$

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

From (48), (50), and (52), for $t \in \{0,\cdots,T-2\}$ and $k \in \{2,\cdots,T-t\}$, we have

$$\begin{aligned}
\omega^{\text{Deliberate}}_{t,k} &= \frac{\Gamma^y_{t+1,k-1}}{\Gamma_{t+1}} = \frac{\delta R^2\Gamma_{t+2}\lambda^a_{t+1}\lambda^y_{t+1,k-1}+1}{\delta R^2\Gamma_{t+2}\left(\lambda^a_{t+1}\right)^2+1}\frac{\Gamma^y_{t+2,k-2}}{\Gamma_{t+2}} \\
&= \left[1 - \frac{\delta R^2\Gamma_{t+2}\lambda^a_{t+1}\left(\lambda^a_{t+1}-\lambda^y_{t+1,k-1}\right)}{\delta R^2\Gamma_{t+2}\left(\lambda^a_{t+1}\right)^2+1}\right]\omega^{\text{Deliberate}}_{t+1,k-1}. \qquad (54)
\end{aligned}$$

Together, we know, generically, $\omega^{\text{Deliberate}}_{t,k} \neq 1$. 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 7. 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 (53) and (54).

(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 (53) and (54). 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 (53) and (54), 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 (54).

Proof of Corollary 9. Similar to Corollary 7, we have, when $T \to +\infty$,

$$\begin{aligned}
\phi_{t}^{\text{Deliberate}} \to \phi^{\text{Deliberate}} &\equiv \frac{\delta R^{2}-1}{\delta R^{2}\left(1-\left(\lambda^{a}\right)^{2}\right)} \\
\Gamma_{t} \to \Gamma &\equiv \frac{\delta R^{2}-1}{\delta R^{2}\left(1-\delta R^{2}\left(\lambda^{a}\right)^{2}\right)}
\end{aligned}$$

From (53) and (54), we know

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

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

Proof of Corollary 10 and Corollary 11. From (31) and (32), we know the case of imperfect perception of wealth is nested by the general case studied in Lemma 1 with λ_t^a given by (31) and $\lambda_{t,k}^y = 0$ for all t, k. Corollary 10 and Corollary 11 then follow from (32), Lemma 1, and Propositions 5 – 7.

Proof of Proposition 8. Consider the environment in Section 5. 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 path $\{\tilde{a}_h, \tilde{c}_h, \tilde{y}_h\}_{h=0}^{T-1}$, with zero net saving at the end of period t, i.e., $\tilde{a}_{t+1} = 0$.[19]

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 (10). 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 $(\tilde{a}_t, \tilde{s}_t, R)$, we have

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

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

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

where I use $\frac{\partial V_{t+1}(\tilde{a}_{t+1}, \tilde{s}_{t+1})}{\partial a_{t+1}} = u'(\tilde{c}_{t+1})$ on the frictionless path[20] and $\Gamma_{t+1} \equiv \frac{\partial^2 V_{t+1}(\tilde{a}_{t+1}, \tilde{s}_{t+1})}{\partial a_{t+1}^2}/u''$ is given by Proposition 5. Proposition 8 then follows from Proposition 5.

Appendix B: Additional Results

Prudence

One may naturally wonder how a "prudent" utility ($u''' > 0$) will change my result. The key result in the literature, e.g., Kimball (1990), is that a prudent agent will display a lower consumption level

[19]On this path, actual consumption coincides with the deliberate consumption $\tilde{c}_t = c_t(\tilde{a}_t, \tilde{s}_t) = c_t^{\text{Delibrate}}(\tilde{a}_t, \tilde{s}_t)$.

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

facing future uncertainty, e.g., a lower $\bar{c}_0^{\text{Deliberate}}$ in (5). A corollary in the current environment is that, with prudence, future "level" mistakes ($|\bar{c}_1|$) will lower the current consumption level $\bar{c}_0^{\text{Deliberate}}$.

Specifically, consider the environment in Section 2 with a prudent cubic utility ($u''' > 0$). The consumption rule at $t = 2$ and $t = 1$ are still given by (3) and (4). The deliberate consumption at $t = 0$ is still given by (5). As I focus on the impact of future "level mistakes" ($\bar{c}_1$) on current consumption, I shut down λ_1 in (4) and write $c_0^{\text{Deliberate}}$ as $c_0^{\text{Deliberate}}(y_0, \bar{c}_1)$.

Proposition 9. *With prudence ($u''' > 0$),*

$$\frac{\partial c_0^{Deliberate}(y_0, 0)}{\partial \bar{c}_1} = 0 \quad \text{and} \quad \frac{\partial^2 c_0^{Deliberate}(y_0, 0)}{\partial \bar{c}_1^2} < 0.$$

Similar to the main analysis based on the quadratic utility, future "level mistakes" ($\bar{c}_1$) do not have a first order impact on current consumption level. However, with prudence, there mistakes will have a negative second order impact on current consumption level (no matter $\bar{c}_1 < 0$ or $\bar{c}_1 > 0$). The intuition is: these future mistakes will generate dispersion of marginal utilities across future periods; with prudence, this dispersion lowers current consumption level. This is similar to the intuition why a prudent agent will display a lower consumption level facing future uncertainty: uncertainty will generate dispersion of marginal utilities across future states; with prudence, this dispersion lowers current consumption level.

But such a lower consumption level can coexist with my key result of a higher current MPC $\phi_0^{\text{Deliberate}}$. In fact, with general concave utilities, the key result here about how future inefficient responses lead to a higher current MPC remains to be true. See Proposition 4 below.

Proof of Proposition 9. Based on the actual consumption rules in (3) and (4), let me define the continuation value function:

$$V_1(a_1, \bar{c}_1) \equiv u\left(\frac{1}{2}a_1 + \bar{c}_1\right) + u\left(\frac{1}{2}a_1 - \bar{c}_1\right).$$

We have:

$$\begin{aligned}\frac{\partial V_1(a_1, \bar{c}_1)}{\partial a_1} &= \frac{1}{2}u'\left(\frac{1}{2}a_1 + \bar{c}_1\right) + \frac{1}{2}u'\left(\frac{1}{2}a_1 - \bar{c}_1\right) \\ &= \frac{\partial V_1(a_1, 0)}{\partial a_1} + \frac{1}{2}u'''(\bar{c}_1)^2,\end{aligned}$$

where the last equation uses the fact that u is a cubic prudent utility function.

At period 0, we have

$$u'\left(c_0^{\text{Deliberate}}(y_0,\bar{c}_1)\right) = \frac{\partial V_1\left(y_0 - c_0^{\text{Deliberate}}(y_0,\bar{c}_1),\bar{c}_1\right)}{\partial a_1}$$
$$= \frac{\partial V_1\left(y_0 - c_0^{\text{Deliberate}}(y_0,\bar{c}_1),0\right)}{\partial a_1} + \frac{1}{2}u'''(\bar{c}_1)^2 .$$

inAs a result, we have (I omit some arguments of functions for notation simplicity)

$$\left[u'' + \frac{\partial^2 V_1}{\partial a_1^2}\right]\frac{\partial c_0^{\text{Deliberate}}(y_0,\bar{c}_1)}{\partial \bar{c}_1} = u''' \cdot \bar{c}_1$$
$$\left[u''' - \frac{\partial^3 V_1}{\partial a_1^3}\right]\left(\frac{\partial c_0^{\text{Deliberate}}(y_0,\bar{c}_1)}{\partial \bar{c}_1}\right)^2 + \left[u'' + \frac{\partial^2 V_1}{\partial a_1^2}\right]\frac{\partial^2 c_0^{\text{Deliberate}}(y_0,\bar{c}_1)}{\partial \bar{c}_1^2} = u'''.$$

Together, we have

$$\frac{\partial c_0^{\text{Deliberate}}(y_0,0)}{\partial \bar{c}_1} = 0 \quad \text{and} \quad \frac{\partial^2 c_0^{\text{Deliberate}}(y_0,0)}{\partial \bar{c}_1^2} = \frac{u'''}{u'' + \frac{\partial^2 V_1}{\partial a_1^2}} < 0.$$

The noisy signal approach to inattention.

In the inattention cases studied in Corollaries 3 and 10, each self's perceived permanent income (or wealth) is given by 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 the same predictions on MPCs.

Here, I use the fungibility case in Corollary 3 as an example to illustrate. The non-fungibility case in Corollary 10 follows similarly. I assume Normally distributed exogenous fundamentals, i.e., $w_0 \sim \mathcal{N}\left(0,\sigma_{w_0}^2\right)$.[21]

Unlike in the main analysis, each self t's knowledge of the current permanent income is now summarized by a noisy signal $x_t = w_t + \epsilon_t$, while $\epsilon_t \sim \mathcal{N}\left(0,\sigma_{\epsilon_t}^2\right)$ and is independent of w_0 and other ϵ_t. In this case, each self understands that her signal is noisy and tries to infer her actual permanent income from the signal.

$$E\left[w_t \mid x_t\right] = (1-\lambda_t)x_t, \tag{55}$$

[21] This together with the linear actual consumption rule in (58) guarantees that each w_t is Normally distributed too.

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

$$c_t(x_t) = \arg\max_{c_t} \ u(c_t) + \delta E\left[V_{t+1}(R(w_t - c_t))\,|x_t\right], \tag{56}$$

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 12. *Each self t's deliberate MPC $\phi_t^{Deliberate}$ is the same as that in Corollary 3, based on $\{\lambda_{t+k}\}$ defined above.*

Proof of Corollary 12. The value in (12) is now given by

$$V_t(w_t) = \int \left[u(c_t(w_t+\epsilon_t)) + \delta V_{t+1}(R(w_t - c_t(w_t+\epsilon_t)))\right] f_t(\epsilon_t)\, d\epsilon_t, \tag{57}$$

where $f_t(\cdot)$ is the p.d.f. given $\epsilon_t \sim \mathcal{N}\left(0, \sigma_{\epsilon_t}^2\right)$. Similar to (15), 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 (16) and (33). For the actual consumption in (56), we have

$$\begin{aligned} c_t(x_t) &= (1-\lambda_t)\phi_t^{\text{Deliberate}}(w_t+\epsilon_t) + \bar{c}_t^{\text{Deliberate}}, \\ &= \phi_t w_t + \phi_t \epsilon_t + \bar{c}_t^{\text{Deliberate}} \end{aligned} \tag{58}$$

where I use (55) and $\phi_t \equiv (1-\lambda_t)\phi_t^{\text{Deliberate}}$ as in the main text.

From (57), we have

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

where $w_{t+1} = R(w_t - c_t(w_t + \epsilon_t))$. The recursive formulation of Γ_t in (36) is then still given by

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

Corollary 12 then follows.

Hyperbolic discounting.

Here I establish some additional results regarding the hyperbolic discounting in Corollary 5.

First, I let $\beta_t = \beta$ for all t and consider the $T \to \infty$ limit. From (42), we know $\Gamma_t \to \Gamma$ where Γ solves

$$\Gamma = \frac{(\delta R^2\Gamma)^2}{1+\delta R^2\Gamma}\left(\frac{1-\beta}{1+\delta\beta R^2\Gamma}\right)^2 + \frac{\delta R^2\Gamma}{1+\delta R^2\Gamma}.$$

From (41), we know

$$\lambda = \frac{1-\beta}{1+\delta\beta R^2\Gamma}.$$

Using $\beta = 0.504$ in Laibson et al. (2018) and standard calibration for $\delta = 0.96$ and $R = 1.03$, we have $\lambda \approx 0.49$, used in Section 4.4.

Second, let me derive the hyperbolic Euler equation in Harris and Laibson (2001) based our framework here. From (21) have

$$\begin{aligned}u'(c_t(w_t)) &= \delta\beta_t R V_{t+1}'(w_{t+1}) \\ u'(c_{t+1}(w_{t+1})) &= \delta\beta_{t+1} R V_{t+2}'(w_{t+2}),\end{aligned}$$

where $w_{t+1} = R(w_t - c_t(w_t))$ and $w_{t+2} = R(w_{t+1} - c_{t+1}(w_{t+1}))$.

From (35), we have:

$$\begin{aligned}V_{t+1}'(w_{t+1}) &= \phi_{t+1}u'(c_{t+1}(w_{t+1})) + (1-\phi_{t+1})\delta R V'(w_{t+2}), \\ &= \phi_{t+1}u'(c_{t+1}(w_{t+1})) + \frac{1-\phi_{t+1}}{\beta_{t+1}}u'(c_{t+1}(w_{t+1})).\end{aligned}$$

Together, we have

$$u'(c_t(w_t)) = R\left[\delta\beta_t\phi_{t+1}u'(c_{t+1}(w_{t+1})) + \delta\beta_t\frac{1-\phi_{t+1}}{\beta_{t+1}}u'(c_{t+1}(w_{t+1}))\right].$$

When $\beta_t = \beta_{t+1} = \beta$, the above expression becomes

$$u'(c_t(w_t)) = R\left[\delta\beta\phi_{t+1}u'(c_{t+1}(w_{t+1})) + \delta(1-\phi_{t+1})u'(c_{t+1}(w_{t+1}))\right],$$

which is the hyperbolic Euler equation in Harris and Laibson (2001).

Inattention to the income state.

In the literature on intertemporal consumption problems with inattention, the focus is inattention to the exogenous income state (e.g. Sims, 2003; Gabaix, 2016, 2019; Luo, 2008).[22] In this literature, the consumer is nevertheless perfectly attentive to her endogenous wealth a_t and the actual consumption can respond frictionlessly to changes in wealth.

In the framework in Section 5, I capture inattention to income similar to (31). That is, for $t \in \{0, \cdots, T-1\}$, I let each self's perceived income state be given by a weighted average of the actual income state and a default. Each self nevertheless perfectly perceives her wealth a_t. That is, for $t \in \{0, \cdots, T-1\}$,

$$s_t^p(s_t) = (1-\lambda_t^y)s_t + \lambda_t^y s_t^d \quad \text{and} \quad a_t^p(a_t) = a_t, \tag{59}$$

where $\lambda_t^y \in [0,1]$ captures self t's degree of inattention to income (a larger λ_t^y means more inattention) and s_t^d captures the default (an exogenous constant whose value does not matter for the MPCs). Recall that, in the environment here, since all income uncertainty is resolved in period 0, $s_t = (y_t, \cdots, y_T)$.

Based on the perceived income state in (59), the actual consumption rule of each self t is given by

$$c_t(a_t, s_t^p(s_t)) = \arg\max_{c_t}\ u(c_t) + \delta V_{t+1}\left(R(a_t + y_t - c_t), s_{t+1}(s_t^p(s_t))\right),$$

where the continuation value function V_{t+1} is defined as usual and $s_{t+1}(s_t^p(s_t))$ captures the perceived future income state based on the perceived current income state $s_t^p(s_t)$. On the other hand, the deliberate consumption is decided as in (11), based on the correct income state and taking future selves' inattention to income as given.

Here, I recover the result in the literature (e.g. Sims, 2003; Gabaix, 2016, 2019; Luo, 2008)

[22] Sims (2003) also studies the inattention to exogenous initial wealth, which effectively plays the same role as exogenous income.

about each self's actual consumption. That is, one can start with the frictionless consumption rule and directly replace actual permanent income with perceived permanent income.

Corollary 13. *For each $t \in \{0, \cdots, T-1\}$, self t's actual consumption is given by*

$$c_t(a_t, s_t) = \phi_t^{Frictionless}\left(a_t + y_t^p + R^{-1}y_{t+1}^p + \cdots + R^{-(T-t)}y_T^p\right) + \bar{c}_t$$

where $\left\{y_{t+k}^p\right\}_{k=0}^{T-t}$ captures self t's perceived future income based on the perceived income state $s_t^p(s_t)$.

In other words, the deviation of the actual consumption from the frictionless one is driven by inattention to current income state. On the other hand, future selves' inattention to income, does not play a special role.

In fact, this result is consistent with the discussion after Proposition 5. The key behind the impact of future consumption mistakes on current MPCs rests upon their inefficient responses to changes in savings/wealth $\left\{\lambda_{t+k}^a\right\}_{k=0}^{T-t-1}$. Here, as future selves are perfectly attentive to their savings/wealth, their consumption responses to changes in savings/wealth are frictionless. This means different selves can frictionlessly coordinate their consumption decisions: if the current self changes her consumption hence her savings, her future selves can perfectly respond to this change. Inattention to income alone will not break this perfect coordination.

Proof of Corollary 13. The case here is nested by the general case studied in Lemma 1 with $\lambda_t^a = 0$ given by (31) and $\lambda_{t,k}^y = \lambda_t^y$ for all $t \in \{0, \cdots, T-t\}$ and $k \in \{0, \cdots, T-t\}$

In the proof of Propositions 5 and 6, if all $\lambda_t^a = 0$, we have $\phi_t^{\text{Delibrate}} = \phi_t^{\text{Frictionless}}$ and $\omega_{t,k}^{\text{Deliberate}} = 0$ for $k \in \{0, \cdots, T-t\}$. Corollary 13 follows.

Distorted expectations.

Another commonly studied behavioral bias in intertemporal consumption problems is distorted expectations (e.g. Mullainathan, 2002; Rozsypal and Schlafmann, 2017; Azeredo da Silveira and Woodford, 2019; Bordalo, Gennaioli and Shleifer, 2018; Bordalo et al., 2020). The general idea is the consumer over-extrapolates based on her current situation. The detailed psychological foundations may include bounded recall in Azeredo da Silveira and Woodford (2019), representativeness in Mullainathan (2002), and diagnostic expectations in Bordalo, Gennaioli and Shleifer, 2018 and Bordalo et al. (2020).

The fungibility case. Let me start from the simple fungibility case in Section 4, I summarize such a friction by letting each self t's perceived permanent income be given by

$$w_t^p(w_t) = w_t + \theta_t \left(w_t - w_t^d\right) \quad \forall t \in \{0, \cdots, T-1\}, \tag{60}$$

where θ_t captures self t's degree of distorted expectations and w_t^d captures the default (an exogenous constant whose value does not matter for the MPCs). $\theta_t > 0$ means that each self t's perceived permanent income $w_t^p(w_t)$ is based on an over-extrapolation from her current permanent income.[23]

In this case, the actual consumption rule is decided based on the perceived permanent income $w_t^p(w_t)$:

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

where the continuing value function V_{t+1} is defined similarly to above. On the other hand, the deliberate consumption rule is decided based on the correct permanent income

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

taking future selves' consumption mistakes as given, driven by future selves' distorted expectations. As a corollary of Proposition 3, these future consumption mistakes lead to a high MPC of the current deliberate consumption.

Corollary 14. *For $t \in \{0, \cdots, T-2\}$, $\phi_t^{Deliberate} \equiv \frac{\partial c_t^{Deliberate}}{\partial w_t} > \phi_t^{Frictionless}$ and increases with future selves' degrees of distorted expectations $\{|\theta_{t+k}|\}_{k=1}^{T-t-1}$.*

The general case allowing non-fungibility. Now we turn to the general case allowing non-fungibility, studied in Section 5. Similar to the discussion after 5 and the inattention case studied in Corollaries 10 and 13, the key about the impact of future consumption mistakes on current MPCs come from their inefficient responses to changes in savings/wealth.

For example, in the fungibility case in (60), future selves over-extrapolate from all components her permanent income equally. This means that future selves over-extrapolate from changes in savings/wealth. This leads to the high current MPCs in Corollary 14.

On the other hand, if future selves' distorted income expectations are fully driven by incomes and independent of savings/wealth (e.g. Mullainathan, 2002; Rozsypal and Schlafmann, 2017;

[23] In fact, when $\theta_t < 0$, the case here is the same as the inattention case studied above.

Azeredo da Silveira and Woodford, 2019; Bordalo, Gennaioli and Shleifer, 2018; Bordalo et al., 2020), these future mistakes will not directly impact current MPCs. This is the same as result in Corollary 13: future selves' inattention to their income will not impact current MPCs.

For example, consider the following variant of (60) regarding distorted expectations about future income:

$$y^p_{t+k}\left(y_{t+k}\right) = y_{t+k} + \theta_{t,k}\left(y_{t+k} - y^d_{t+k}\right) \quad \forall t \in \{0, \cdots, T-1\},$$

where $\theta_{t,k}$ captures self t's degree of distorted expectations with regard to y_{t+k} and y^d_{t+k} captures the exogenous default. $\theta_{t,k} > 0$ means that each self t's perceived future income over-reacts to changes in actual permanent income. Note that in this case, self t's distorted income expectations do not depend on current wealth a_t.

Corollary 15. *For $t \in \{0, \cdots, T-1\}$, each self's deliberate consumption in (27) coincides with the frictionless one. That is, $\phi_t^{Deliberate} = \phi_t^{Frictionless}$ and $w_{t,k}^{Delibrate} = 1$ for all $k \in \{0, \cdots, T-t\}$.*

Proof of Corollary 14. This case is nested in Proposition 3 with $\lambda_t = -\theta_t$.

Proof of Corollary 15. This case is nested in Lemma 2 with $\lambda^a_t = 0$ and $\lambda^y_{t,k} = -\theta_{t,k}$ for all $t \in \{0, \cdots, T-1\}$ and $k \in \{0, \cdots, T-t\}$. In the proof of Propositions 5 and 6 , if all $\lambda^a_t = 0$, we have $\phi_t^{\text{Delibrate}} = \phi_t^{\text{Frictionless}}$ and $\omega_{t,k}^{\text{Deliberate}} = 0$ for $k \in \{0, \cdots, T-t\}$. Corollary 15 follows.

Graduate Resolution of Uncertainty in the General Case

Consider the environment in Section 5. With a graduate resolution of the income uncertainty, the actual consumption rule of each self $t \in \{0, \cdots, T-1\}$ can now be written as

$$c_t\left(a_t, s_t\right) = \phi^a_t a_t + \phi^y_t\left(y_t + \sum_{k=1}^{T-t} \omega_{t,k} R^{-k} E_t\left[y_{t+k}\right]\right) + \bar{c}_t,$$

where $E_t\left[y_{t+k}\right] = E_t\left[y_{t+k}|s_t\right]$ captures the expected future income based on the current income state s_t. Self t's mistakes λ^a_t and $\left\{\lambda^y_{t,k}\right\}_{k=0}^{T-t}$ are still given by (26).

Based on future selves' actual consumption rules $\left\{c_{t+k}\left(a_{t+k}, s_{t+k}\right)\right\}_{k=0}^{T-1-t}$, each self t's deliberate consumption rule defined in (8) will take the following form.

Corollary 16. *For $t \in \{0, \cdots, T-1\}$, each self t's deliberate consumption rule is given by:*

$$c_t^{Deliberate}(a_t, s_t) = \phi_t^{Deliberate}\left(a_t + y_t + \sum_{k=1}^{T-t} \omega_{t,k}^{Deliberate} R^{-k} E_t[y_{t+k}]\right) + \bar{c}_t^{Deliberate}, \tag{61}$$

where $\phi_t^{Deliberate}$ and $\left\{\omega_{t,k}^{Deliberate}\right\}_{k=1}^{T-t}$ share the exact same formula as in Lemma 2.

Proof of Corollary 16. With graduate resolution of uncertainty, the optimal deliberate consumption in (11) becomes

$$c_t^{\text{Deliberate}}(a_t, s_t) = \max_{c_t} \; u(c_t) + \delta E_t[V_{t+1}(R(a_t + y_t - c_t), s_{t+1})],$$

while the recursive formulation for the value function in (12) becomes

$$V_t(a_t, s_t) = u(c_t(a_t, s_t)) + \delta E_t[V_{t+1}(R(a_t + y_t - c_t(a_t, s_t)), s_{t+1})],$$

where $E_t[\cdot] = E_t[\cdot|(a_t, s_t)]$ captures rational expectations based on period t's state (a_t, s_t).

The proof in Lemma 2 remains unchanged, except in all expressions y_{t+k} is replaced with $E_t[y_{t+k}] = E_t[y_{t+k}|s_t]$. In particular, the formulas (47) – (52) remain to be true. So Corollary 16 follows directly.

How Do "Level" Mistakes in Future Consumption Impact Current MPCs

In the main analysis, future selves' level mistakes, i.e., $\bar{c}_{t+k}$ in (13), do not impact current MPCs. Here let me talk about a situation where future level mistakes can also impact current MPCs.

If future selves' level mistakes (e.g., over-consumption) lead to future liquidity constraints binding, it can also increase current MPCs. To illustrate, consider the environment in Section 5. Consider the case that future self $t+1$'s level mistake binds her liquidity constraint: $c_{t+1} = a_{t+1} + y_{t+1} + \bar{b}_{t+1}$, where $\bar{b}_{t+1}$ captures self $t+1$'s borrowing limit. In this case, the concavity of the continuation value is given by

$$\Gamma_{t+1} \equiv \frac{\partial^2 V_{t+1}(a_{t+1}, s_{t+1})}{\partial a_{t+1}^2} / u'' = 1.$$

This is easily larger than $\Gamma_{t+1} = \frac{\delta R^2 \Gamma_{t+2}}{1+\delta R^2 \Gamma_{t+2}}$ in (36) when self $t+1$ does not have any bias. This excess concavity of the continuation value then generates a high current deliberate MPC $\phi_t^{\text{Deliberate}}$.

In fact, this channel is relevant for the high MPCs of hyperbolic discounting agents with liquidity constraints studied in the literature (Laibson, 1997; Angeletos et al., 2001).

References

Abel, Andrew, Janice Eberly, and Stavros Panageas. (2007) "Optimal Inattention to the Stock Market." *American Economic Review*, 97(2): 244–249.

Abel, Andrew, Janice Eberly, and Stavros Panageas. (2013) "Optimal Inattention to the Stock Market with Information Costs and Transactions Costs." *Econometrica*, 81(4): 1455–1481.

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.*

Alvarez, Fernando, Luigi Guiso, and Francesco Lippi. (2012) "Durable Consumption and Asset Management with Transaction and Observation Costs." *American Economic Review*, 102(5): 2272–2300.

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.

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

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.

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*.

Brunnermeier, Markus K, and Stefan Nagel. (2008) “Do wealth fluctuations generate time-varying risk aversion? Micro-evidence on individuals.” *American Economic Review*, 98(3): 713–36.

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.

Caplin, Andrew, Mark Dean, and John Leahy. (2019) “Rational Inattention, Optimal Consideration Sets, and Stochastic Choice.” *The Review of Economic Studies*, 86(3): 1061–1094.

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

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

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*.

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.

Fudenberg, Drew, and David Levine. (2006) "A dual-self Model of Impulse Control." *American Economic Review*, 96(5): 1449–1476.

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.*

Gabaix, Xavier. (2019) "Behavioral Inattention." *Handbook of Behavioral Economics.*

Gabaix, Xavier, and David Laibson. (2002) "The 6D Bias and the Equity-premium Puzzle." *NBER Macroeconomics Annual*, 16: 257–312.

Ganong, Peter, and Pascal Noel. (2019) "Consumer Spending during Unemployment: Positive and Normative Implications." *American Economic Review*, 109(7): 2383–2424.

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

Gul, Faruk, and Wolfgang Pesendorfer. (2004) "Self-control and the Theory of Consumption." *Econometrica*, 72(1): 119–158.

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, 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. (2011) *Thinking, Fast and Slow.* Macmillan.

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

Kaplan, Greg, and Giovanni L Violante. (2014) "A Model of the Consumption Response to Fiscal Stimulus Payments." *Econometrica*, 82(4): 1199–1239.

Kaplan, Greg, and Giovanni Violante. (2010) "How much Consumption Insurance beyond Self-insurance?" *American Economic Journal: Macroeconomics*, 2(4): 53–87.

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, Andrea Repetto, and Jeremy Tobacman. (2018) "Estimating Discount Functions with Consumption Choices over the Lifecycle." *Harvard University mimeo.*

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

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

Luo, Yulei. (2008) "Consumption Dynamics under Information Processing Constraints." *Review of Economic Dynamics*, 11(2): 366–385.

Luo, Yulei, and Eric Young. (2010) "Risk-sensitive Consumption and Savings under Rational Inattention." *American Economic Journal: Macroeconomics*, 2(4): 281–325.

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

Maggio, Marco Di, Amir Kermani, and Kaveh Majlesi. (2019) "Stock Market Returns and Consumption." *Journal of Finance, Forthcoming.*

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.

Pagel, Michaela. (2017) "Expectations-based Reference-dependent Life-cycle Consumption." *The Review of Economic Studies*, 84(2): 885–934.

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.

Reis, Ricardo. (2006) "Inattentive consumers." *Journal of monetary Economics*, 53(8): 1761–1800.

Rozsypal, Filip, and Kathrin Schlafmann. (2017) "Overpersistence Bias in Individual Income Expectations and its Aggregate Implications."

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

Shimer, Robert. (2009) "Convergence in Macroeconomics: The Labor Wedge." *American Economic Journal: Macroeconomics*, 1(1): 280–97.

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.