

Mistakes in Future Consumption, High MPCs Now^{*}

Chen Lian[†]

September 29, 2020

Preliminary and Incomplete.

Abstract

This paper develops an approach to study predictions independent from the exact psychological cause of behavioral mistakes. In a canonical intertemporal consumption problem, I show how anticipation of future mistakes, by itself, explains key deviations from the permanent income hypothesis. The result provides a potential explanation of the empirical puzzle on high liquidity consumers' high marginal propensities to consume (MPCs) and violations of the fungibility principle. I also illustrate how my framework can accommodate most widely-studied behavioral biases, such as inattention, mental accounting, rules of thumb, and hyperbolic discounting.

^{*}I am extremely grateful to Marios Angeletos, Alp Simsek and Ricardo Caballero for continuous guidance through the project. I am grateful to Nick Barberis, Jonathon Hazell, Julian Kozlowski, Jonathan Parker, Karthik Sastry, Frank Schilbach, Dmitry Taubinsky, Ludwig Straub, Ivan Werning and Muhamet Yildiz, and seminar participants at multiple institutions 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 Pre-doctoral Fellowship in Behavioral Macroeconomics, awarded through the NBER. Bruno Smaniotto provides excellent research assistance.

[†]UC Berkeley; chen_lian@berkeley.edu.

1 Introduction

There is mounting evidence that consumers exhibit important deviations from the permanent income hypothesis *away* from liquidity constraints (Thaler, 1990). Importantly, 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 on high-liquidity consumers’ deviations 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.

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 Penderfer, 2004; Fudenberg and Levine, 2006), and distorted expectations (Mullainathan, 2002; Azeredo da Silveira and Woodford, 2019) can make non-behavioral economists lost about what the actual take-home lessons are about the consumption behavior.

In this paper, I provide a new angle to study how behavioral biases can influence consumption behavior. Different from the existing behavioral literature, I do not take an exact stand on what the underlying behavioral biases are. 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 and study their robust implications.

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. Obviously, the impact of this channel depends on the exact underlying bias, and “anything goes.” The second channel, instead, captures how anticipation of future mistakes, i.e., sophistication in the language of (O’Donoghue and Rabin, 1999, 2001), impacts current consumption. My contribution is to develop a method to isolate the second channel and show how anticipation of future consumption mistakes (in response to changes in savings), no matter the exact behavioral cause of these mistakes, can robustly lead to high MPCs and violations of the fungibility principle. Moreover, consumers do not need to fully anticipate their future mistakes. Partial sophistication suffices for all results.

Mistakes in future consumption, high MPCs now. I study a canonical intertemporal consumption and saving problem. To illustrate how future consumption mistakes lead to high current MPCs, I first consider a benchmark, fungible case in which 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). Future consumption mistakes come from their inefficient responses to changes in permanent income. The main result is that these future consumption mistakes lead to high current MPCs.

To understand this result, I first show that the inefficient responses of future consumption lead to the excess concavity of the continuation value function. As a result, in response to changes in current permanent income, the consumer is more willing to adjust her current consumption instead of her savings. She hence displays high current MPCs.

Intuitively, using a positive shock as an example, if the consumer saves the additional money, her future selves will not respond to it efficiently. She instead increases her current consumption more.

The high current MPCs result does not depend on the exact behavioral causes of future consumption mistakes. No matter whether future consumption mistakes lead to over-reaction or under-reaction to changes in permanent income, these mistakes always increase current MPCs. But I also illustrate that my framework can accommodate many widely-studied behavioral biases, such as inattention, rules of thumb, and hyperbolic discounting.

Future non-fungibility begets current 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 to be true: as long as future consumption responds inefficiently to changes in savings, current MPCs are higher.

Interestingly, the non-fungibility of future consumption by itself also suffices to generate the non-fungibility of the current consumption. In other words, even if the current self fully understands permanent income hypothesis, 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.

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

the empirical evidence in Kueng (2018).

Applications. The main application of the proposed channel is to explain high-liquidity consumers' high MPCs. The key mechanism behind the high MPCs, i.e., the excess concavity of the continuation value function driven by future mistakes, can also speak to three other well-known puzzles in intertemporal decisions: excess discounting of future income mentioned above; large risk aversion and the equity premium puzzle (Mehra and Prescott, 1985); and 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).

Literature review. This paper builds upon the behavioral literature on intertemporal consumption problems. For example, 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. Mullainathan, 2002; Rozsypal and Schlafmann, 2017; Azeredo da Silveira and Woodford, 2019) and news utility (e.g. Kőszegi and Rabin, 2009; Pagel, 2017).

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 from the exact behavioral mistakes. Farhi and Gabaix (2020) use the wedge approach to study optimal taxation with behavioral agents but does not touch upon the robust predictions on behavior focused here.

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

Layout. Section 2 introduces a standard income fluctuations problem and shows how to isolate the impact of future consumption mistakes on current consumption. Section 3 studies the benchmark fungibility case and illustrates how future consumption mistakes can lead to high MPCs now. Section 4 studies the general non-fungibility case. Section 5 focuses on other applications. Section 6 concludes. The Appendix contains proofs and additional results.

2 Set up

This section introduces a standard income fluctuations 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 here is not subject to any borrowing constraints.

The consumer’s utility is given by

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

where c_t is her consumption at period $t \in \{0, 1, \dots, T-1\}$, δ is her discount factor, $u(\cdot)$ captures the consumption utility, and $v(\cdot) : \mathbb{R} \rightarrow \mathbb{R}$ 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. But the main high MPCs results remain true with general concave utility, as Proposition 4 below shows.

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

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

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 then 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 $s_t = (y_t, \dots, y_T)$. It is worth noticing that, with

quadratic utility and linear decision rule here, the well known certainty equivalence result implies that the consumer's MPC remains the same with gradual resolution of income uncertainty (see Corollary 6 below).

I use the widely adapted “multiple-selves” language as in Piccione and Rubinstein (1997) and Harris and Laibson (2001). That is, self $t \in \{0, \dots, T-1\}$ is in charge 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.

Isolate 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 (parametrized 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 latter channel, I introduce *deliberate* consumption rule $c_t^{\text{Deliberate}}(a_t, s_t)$. That is, the consumption self t would have chosen if she is not subject to any current behavioral bias but takes future selves' mistakes in their actual consumption rules as given.

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

$$c_t^{\text{Deliberate}}(a_t, s_t) \equiv \arg \max_{c_t} u(c_t) + \sum_{k=1}^{T-t-1} \delta^{k-1} u(c_{t+k}(a_{t+k}, s_{t+k})) + \delta^{T-t} v(a_T + y_T), \quad (3)$$

where $a_{t+k} = R(a_{t+k-1} + y_{t+k-1} - c_{t+k-1})$.

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

$$c_t(a_t, s_t) = \mathcal{S}(c_t^{\text{Deliberate}}(a_t, s_t), \lambda_t). \quad (4)$$

Self t 's own behavioral bias (parametrized by λ_t) impacts her actual consumption by letting it deviate from the deliberate consumption $c_t^{\text{Deliberate}}(a_t, s_t)$, captured by function \mathcal{S} .¹ On the other hand,

¹We have $\mathcal{S}(c_t^{\text{Deliberate}}(a_t, s_t), 0) = c_t^{\text{Deliberate}}(a_t, s_t)$. 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 (3).

anticipation of future selves' mistakes $\{\lambda_{t+k}\}_{k=1}^{T-t-1}$ impacts current actual consumption through the deliberate consumption $c_t^{\text{Deliberate}}(a_t, s_t)$.

The main theme for the rest of the paper is: once we isolate the impact of future consumption mistakes through deliberate consumption (3), we can show these future mistakes robustly lead to high current MPCs, no matter the micro-foundations of these mistakes.

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

$$V_t(a_t, s_t) \equiv u(c_t(a_t, s_t)) + \sum_{k=1}^{T-t-1} \delta^k u(c_{t+k}(a_{t+k}, s_{t+k})) + \delta^{T-t} v(a_T + y_T), \quad (5)$$

where $a_{t+k} = R(a_{t+k-1} + y_{t+k-1} - c_{t+k-1})$. For the last period T , we have $V_T(a_T, s_T) = v(a_T + y_T)$.

Based on (5), I can express the deliberate consumption rule in (3) recursively. This recursive formulation paves ways for the analysis in the rest of the paper.

Proposition 1. *For $t \in \{0, \dots, T-1\}$, each self t 's deliberate consumption rule defined in (3) satisfies*

$$c_t^{\text{Deliberate}}(a_t, s_t) = \max_{c_t} u(c_t) + \delta V_{t+1}(R(a_t + y_t - c_t), s_{t+1}). \quad (6)$$

Moreover, for $t \in \{0, \dots, T-1\}$, the value $V_t(a_t, s_t)$ defined in (5) satisfies

$$V_t(a_t, s_t) = u(c_t(a_t, s_t)) + \delta V_{t+1}(R(a_t + y_t - c_t(a_t, s_t)), s_{t+1}), \quad (7)$$

where the actual consumption rule $c_t(a_t, s_t)$ is given by (4).

Finally, if consumption rules and value functions $\{c_t^{\text{Deliberate}}(a_t, s_t), c_t(a_t, s_t)\}_{t=0}^{T-1}$ and $\{V_t(a_t, s_t)\}_{t=0}^T$ satisfy (4), (6), (7), and the boundary condition $V_T(a_T, s_T) = v(a_T + y_T)$, they coincide with the corresponding objects defined sequentially in (3) – (5).

A note on budget constraints. It is worth noting that the final wealth a_T is allowed to be negative, as 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 (2) 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 $T \rightarrow +\infty$.

3 The Benchmark Fungibility Case

To illustrate how future consumption mistakes lead to high current MPCs, I first consider 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. The main result is such mistakes lead to high current MPCs. This result does not depend on the exact micro-foundations of future consumption mistakes come from. But I illustrate how my framework can accommodate common behavioral biases such as inattention, rules of thumb, hyperbolic discounting, and stochastic mistakes.

3.1 Mistakes in Future Consumption, High MPCs Now

In this section, to illustrate the key result in the simplest manner, I consider a benchmark fungibility case. That is, both c_t and $c_t^{\text{Deliberate}}$ remain 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

$$c_t(w_t) = \phi_t w_t + \bar{c}_t \quad \text{and} \quad c_t^{\text{Deliberate}}(w_t) = \phi_t^{\text{Deliberate}} w_t + \bar{c}_t^{\text{Deliberate}}, \quad (8)$$

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 (5).

Here, the key mistakes in actual consumption rules come from their inefficient responses to changes in permanent income. That is, the actual MPC ϕ_t may deviate from the deliberate MPC $\phi_t^{\text{Deliberate}}$. Aligned with (4), I use λ_t to capture this mistake

$$\phi_t = (1 - \lambda_t) \phi_t^{\text{Deliberate}}. \quad (9)$$

In other words, λ_t in (9) 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 treated as exogenous here, but will be connected to the exact underlying behavioral biases below.

Mistakes in actual consumption rules may also involve “level mistakes,” i.e., $\bar{c}_t \neq \bar{c}_t^{\text{Deliberate}}$. But

as shown below, future consumption mistakes in these form 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 the current deliberate consumption. Specifically, based on Definition 1, 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 her deliberate MPC $\phi_t^{\text{Deliberate}}$.

Proposition 2. *For $t \in \{0, \dots, T-2\}$, each self t 's deliberate MPC $\phi_t^{\text{Deliberate}}$ is a function of $(\{\lambda_{t+k}\}_{k=1}^{T-t-1}, \delta, R)$. Moreover, $\phi_t^{\text{Deliberate}} \geq \phi_t^{\text{Frictionless}}$ and increases with each future self's mistake $\{|\lambda_{t+k}|\}_{k=1}^{T-t-1}$, where $\phi_t^{\text{Frictionless}}$ is the frictionless MPC of the actual consumption when all λ s are equal to 0.*

In other words, no matter 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 2, let me first introduce an intermediate step. From the recursive formulation in (6), 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 $t \in \{0, \dots, T-1\}$,

$$\Gamma_{t+1} \equiv \frac{\partial^2 V_{t+1}(w_{t+1})}{\partial w_{t+1}^2} / u'' > 0, \quad (10)$$

where a larger Γ_{t+1} means a more concave value function $V_{t+1}(w_{t+1})$.^{2'3}

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

The intuition behind the excess concavity is: larger $\{|\lambda_{t+k}|\}_{k=0}^{T-t-1}$ lead to more inefficient responses of future consumption to changes in w_{t+1} . As a result, the marginal value $\frac{\partial V_{t+1}}{\partial w_{t+1}}$ decreases faster with w_{t+1} and the continuation value function becomes more concave.

² $u'' < 0$, a constant, is the second derivative of the utility function. Moreover the definition in (10) can be extended to $\Gamma_0 \equiv \frac{\partial^2 V_0(w_0)}{\partial w_0^2} / u''$.

³Even with future consumption mistakes, the continuation value function here $V_{t+1}(w_{t+1})$ is always concave. This feature is guaranteed because the current paper 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).

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 2. From the recursive formulation in (6), we know

$$c_t^{\text{Deliberate}}(w_t) = \max_{c_t} u(c_t) + \delta V_{t+1}(R(w_t - c_t)). \quad (11)$$

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.

Intuitively, after a positive shock, if the current self saves the additional money, her future selves will not respond to the increase in saving efficiently. She 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.

In sum, Proposition 2 shows that, once we isolate the impact of future consumption mistakes on current MPCs, it always raises the current MPC, no matter 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 current MPCs, which can go either way and “anything goes.”

Partial sophistication. By definition, the deliberate consumption $c_t^{\text{Deliberate}}(w_t)$ defined above studies how full knowledge about future selves’ mistakes impacts current consumption. But the high MPC result in Proposition 2 can be easily translated to the case when the current self only has a partial understanding of her future selves’ mistakes, i.e, partial sophistication in O’Donoghue and Rabin (1999, 2001). In fact, one reason why I analyze the problem through the lens of excess concavity of the continuation value function in Lemma 1 is that it can easily translate to the case with partial sophistication.

Specifically, I now let the deliberate consumption $c_t^{\text{Deliberate}}$ be determined with a partial understanding of her future selves’ mistakes:

$$\tilde{\lambda}_{t,t+k} = s_t \lambda_{t+k}, \quad (12)$$

where $s_t \in [0, 1]$ captures the degree of self t 's sophistication. The optimality condition of the deliberate consumption in (6) can be re-written as

$$c_t^{\text{Deliberate}}(w_t) = \max_{c_t} u(c_t) + \delta \tilde{V}_{t,t+1}(R(w_t - c_t)).$$

where $\tilde{V}_{t,t+1}(\cdot)$ is self t 's perceived continuation value function based on her partial understanding about her future selves' mistakes.

In fact, $\tilde{V}_{t,t+1}(\cdot)$ coincides with the actual continuation value function in (5) if future selves' actual mistakes are given by $\left\{ \tilde{\lambda}_{t,t+k} \right\}_{k=1}^{T-t-1}$. As a result, Proposition 2 can be rewritten as:

Proposition 3. *The deliberate MPC $\phi_t^{\text{Deliberate}}$ is now a function of $\left\{ \tilde{\lambda}_{t,t+k} \right\}_{k=1}^{T-t-1}$ and increases with each $\left\{ \left| \tilde{\lambda}_{t,t+k} \right| \right\}_{k=1}^{T-t-1}$, i.e., self t 's perceived mistake of her future self $t+k$.*

3.2 Different Micro-Foundations, Same Results

The high MPC result in Proposition 2 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 will show that future selves' inattention to permanent income lead to high current deliberate MPCs.

Mathematically, I follow the sparsity approach in Gabaix (2014) and let each self t 's perceived permanent income be given by

$$w_t^p(w_t) = (1 - \lambda_t) w_t + \lambda_t w_t^d, \tag{13}$$

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 of which the exact 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 will lead to the same predictions on MPCs, as explained in the Appendix B.

Based on the perceived permanent income $w_t^p(w_t)$ in (13), the actual consumption rule of each self t is given by

$$c_t(w_t) = \arg \max_{c_t} u(c_t) + \delta V_{t+1}(R(w_t^p(w_t) - c_t)), \quad (14)$$

where the continuation value function V_{t+1} is defined similarly to above, based on future selves' actual consumption rules.

To isolate the impact of future inattention on current consumption, the deliberate consumption is defined based on the correct permanent income taking future selves' inattention to permanent income as given. As a corollary of Proposition 2, future consumption mistakes in the form of inattention lead to high current deliberate MPCs.

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

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 $t \in \{0, \dots, T-1\}$ be given by

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

where $c_t^R(w_t) \equiv \phi_t^R w_t + \bar{c}_t^R$ 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 $c_t^R(w_t)$. With probability $1 - p_t$, the current self makes her consumption decision based on “system 2:” the actual consumption is given by the deliberate consumption.⁴

The deliberate consumption rule is defined as usual:

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

not subject to any current behavioral bias and taking future selves' mistakes as given. As a

⁴This case is not directly nested in Proposition 2, as the actual consumption rule is stochastic. But the key results in Proposition 2 can be easily extended. See the proof of Corollary 2.

corollary of Proposition 2, future consumption mistakes, in the form of rules of thumb, lead to high current MPC of the deliberate consumption.

Corollary 2. *For $t \in \{0, \dots, T-2\}$, $\phi_t^{\text{Deliberate}} \equiv \frac{\partial c_t^{\text{Deliberate}}}{\partial w_t} \geq \phi_t^{\text{Frictionless}}$ and 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, future selves' potential rules of thumb behavior raise 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

$$c_t(w_t) = \arg \max_{c_t} u(c_t) + \delta \beta_t V_{t+1}(R(w_t - c_t)) \quad \forall t \in \{0, \dots, T-1\}, \quad (15)$$

where $\delta \in [0, 1]$ is the standard discount factor, $\beta_t \in [\frac{1}{2}, 1]$ captures self t 's present bias (a smaller β_t means a larger present bias), and $V_{t+1}(\cdot)$ is the continuation value function defined similarly to above.⁵ Such an actual consumption rule is the focus of the hyperbolic discounting literature. It combines the direct effect of present bias on current consumption with the effect of anticipated future mistakes.

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

$$c_t^{\text{Deliberate}}(w_t) = \arg \max_{c_t} u(c_t) + \delta V_{t+1}(R(w_t - c_t)) \quad \forall t \in \{0, \dots, T-1\}. \quad (16)$$

That is, the consumption that would have been chosen if the current self is not subject to the present bias but takes future selves' present biases into consideration. As a corollary of Proposition 2, these future consumption mistakes lead to high current deliberate MPCs.⁶

Corollary 3. *For $t \in \{0, \dots, T-2\}$, $\phi_t^{\text{Deliberate}} \equiv \frac{\partial c_t^{\text{Deliberate}}}{\partial w_t} \geq \phi_t^{\text{Frictionless}}$ and increases with future selves' present bias, i.e., decreases in each $\{\beta_{t+k}\}_{k=1}^{T-t-1}$.*

In fact, in the environment here, high actual current MPCs under hyperbolic discounting come solely from the impact of future consumption mistakes. The current present bias, though increases

⁵The restriction $\beta_t \geq \frac{1}{2}$ makes sure that there is an upper bound for the concavity Γ_{t+1} in (10) and the comparative statics in Corollary 3 is well behaved.

⁶One can also easily derive the hyperbolic Euler equation in Harris and Laibson (2001) based on our framework here. See Appendix B for details.

the *level* of current actual consumption, decreases current MPCs. To see this, using the FOC of the actual consumption in (15) and taking a partial derivative with respect to w_t , we have

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

That is, holding constant the concavity of the continuation value V''_{t+1} , the current actual MPC decreases with the degree of present bias (i.e., increases with β_t). The intuition is, with present bias, 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 permanent income.

An additional clarification regarding this hyperbolic discounting example is worth mentioning. Future selves' present biases can generate two types of future consumption mistakes, i.e., inefficient responses to changes in permanent income (focused above) and mistakes in the over-recall consumption level. As discussed above, with quadratic utility, the latter channel does not impact current MPCs here. With precautionary saving motives (i.e., when $u''' \neq 0$), these “level” mistakes in future consumption can also impact current MPCs: future selves' over-consumption due to present biases may further increase current deliberate MPCs because it increases the current self's precautionary saving motives. This channel in principle can also be studied based on the framework in Section 2 independent from the exact behavioral biases, but is beyond the scope of the current paper.

Near-rationality and stochastic mistakes. In the context of intertemporal consumption problems, another often discussed notion of behavioral mistakes is “near-rationality” (Cochrane, 1989 and Kueng, 2018). The idea is: because the utility loss of deviations from the optimal consumption rule is at most second order, mistakes in actual consumption can be prevalent. Different from the above micro-foundations, this “near-rationality” foundation does not bias the actual consumption rule in a particular way. As a result, in the environment here, I capture it by letting the actual consumption rule deviate from the deliberate one in a stochastic fashion. That is, the actual consumption rule for each self $t \in \{0, \dots, T-1\}$ is given by

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

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 $E[\varphi_t] = E[\epsilon_t] = 0$. They are i.i.d and independent from each other.⁷

⁷This case is not directly nested in Proposition 2, as the actual consumption rule is stochastic. But the key

The deliberate consumption rule is defined as usual:

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

not subject to any current behavioral bias and taking future selves' stochastic mistakes as given. As a corollary of Proposition 2, these future consumption mistakes lead to high current MPCs of the deliberate consumption.

Corollary 4. *For $t \in \{0, \dots, T-2\}$, $\phi_t^{\text{Deliberate}} \equiv \frac{\partial c_t^{\text{Deliberate}}}{\partial w_t} \geq \phi_t^{\text{Frictionless}}$ and increases with the variances in future selves' actual MPCs $\{Var(\varphi_{t+k})\}_{k=1}^{T-t-1}$.*

This result means, even if future selves' actual consumption may be unbiased on average, their stochastic consumption mistakes increase current MPCs. Moreover, aligned with the previous discussion, the key is future selves stochastic mistakes in response to changes in permanent income, not their stochastic consumption levels.

An interpretation independent of the specific biases. Beyond the specific biases studied above, let me provide another interpretation independent of the specific biases. Through her life experiences, the consumer knows that she has cognitive limitations and her future consumption may not respond efficiently to changes in permanent income. With this knowledge and even without knowledge of the exact mistakes of their future selves, the consumer will increase her current MPC as a second-best response to future consumption mistakes.

3.3 The $T \rightarrow \infty$ limit and Gauging the Magnitudes

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

Corollary 5. *Fix a self t . Let $\lambda_{t+k} = \lambda$ with $|\lambda| < (\delta^{-1/2} R^{-1})$ for all $k \geq 1$. We have, for $T \rightarrow +\infty$,*

$$\phi_t^{\text{Deliberate}} \rightarrow \phi^{\text{Deliberate}} = \frac{\delta R^2 - 1}{\delta R^2 (1 - \lambda^2)}, \quad (17)$$

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

results in Proposition 2 can be easily extended. See the proof of Corollary 4.

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

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

That is, when future selves' consumption mistakes are large enough, the current self is so worried about her future selves' mistakes that she follows a simple rule of thumb: she consumes all changes in her permanent income.

Gauging the magnitudes. Another use of the limit result in Corollary 5 is that it helps us gauge how much anticipation of future consumption mistakes can impact current MPCs. In particular, one can use standard calibration of a particular friction to calibrate λ and use (17) 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 1. Of course there is caveat that attention to different objects differ (in fact Corollary 9 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 (13). From (17), this implies that anticipation of future inattention can increase current MPCs by around 45%. In Appendix B, I also use Corollary 3 to map the standard present bias 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%.

3.4 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 gradual resolution of the income uncertainty, the actual and deliberate consumption rule of each self $t \in \{0, \dots, T-1\}$ 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 $w_t = E_t \left[a_t + y_t + \sum_{k=1}^{T-t} R^{-k} y_{t+k} \mid (a_t, s_t) \right]$ 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 express her deliberate MPC $\phi_t^{\text{Deliberate}}$ as functions of δ , R , and future selves' mistakes $\{\lambda_{t+k}\}_{k=1}^{T-t-1}$. We have:

Corollary 6. *For $t \in \{0, \dots, T-2\}$, $\phi_t^{\text{Deliberate}}$ shares the exact same formula as $\phi_t^{\text{Deliberate}}$ in Proposition 2.*

General concave utilities. As discussed in the hyperbolic discounting example, in general, future consumption mistakes can take in two forms: inefficient responses to changes in permanent income and mistakes in the over-recall consumption level. The above analysis uses the tractable quadratic utility case to illustrate how the former channel, i.e., future selves' inefficient responses, robustly increases current MPCs. In this case, the latter channel, i.e., future selves' "level" mistakes, does not impact current MPCs.

With general concave utilities $u(\cdot)$ and $v(\cdot)$, the impact of the former channel on current MPCs remains to be the same. To illustrate, consider the case that each self's actual consumption responds inefficiently to changes in permanent income 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 coincides with the deliberate consumption $\tilde{c}_t = c_t(\tilde{w}_t) = c_t^{\text{Deliberate}}(\tilde{w}_t)$. In other words, on this path, each self's actual consumption level coincides with that if she is not subject to any current behavioral bias. On the other hand, each self's actual consumption responds inefficiently to changes in permanent income away from the path. Similar to (9), I use λ_t to capture self t 's inefficient responses. That is,

$$\lambda_t \equiv 1 - \frac{\partial c_t(\tilde{w}_t) / \partial w_t}{\partial c_t^{\text{Deliberate}}(\tilde{w}_t) / \partial w_t}. \quad (18)$$

I can now re-establish that the main Proposition 2 above in the case of general concave utility.

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

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

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

As a result, the above analysis about how future selves' inefficient responses to change in permanent income lead to high current MPCs remains to hold in the general case here. In future

works, I also want to explore how future “level” mistakes, independent from the exact behavioral causes, impact current MPCs.

Empirical support. Proposition 2 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 the MPC remains high 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. Relatedly, Stephens and Unayama (2011), Parker (2017), Olafsson and Pagel (2018), Ganong and Noel (2019), McDowall (2020) also question whether liquidity constraints can explain their findings on 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 contest of hyperbolic discounting, Allcott et al. (2020) find that the perceived and actual present bias parameters are, respectively, 0.75 and 0.72. This implies a degree of sophistication (s_t in 12) close to 1. In fact, I am not aware of any empirical study which finds that consumers are fully naive about their future mistakes.

4 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/borrowings, current MPCs are higher. Then, I show that the non-fungibility of future consumption by itself also suffices to generate the non-fungibility of the current consumption. Even if the current self fully 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 begets current non-fungibility. Finally, I illustrate how the framework can accommodate several behavioral biases causing inefficiently differential responses to different components of permanent income.

4.1 The Environment

In Section 3. I restrict actual consumption to be function of permanent income: $w_t = a_t + y_t + \sum_{k=1}^{T-t} R^{-k} y_{t+k}$. Here, I allow the actual consumption to respond different components of permanent income differently. In other words, mistakes in actual consumption rules may also include inefficiently differential responses to different components of permanent income.

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

$$c_t(a_t, s_t) = \phi_t^a a_t + \phi_t^y \left(y_t + \sum_{k=1}^{T-t} \omega_{t,k} R^{-k} y_{t+k} \right) + \bar{c}_t, \quad (19)$$

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 $\omega_{t,k} < 1$, it means the consumer excessively discounts future income k periods later. Finally, \bar{c}_t in (19) is an exogenous constant capturing the level of self t 's actual consumption, of which the exact value will not influence the deliberate MPCs calculated below.

The actual consumption rule in (19) allows differential mistakes in response to different components of permanent income. Similar to (9), I use $\lambda_t = (\lambda_t^a, \{\lambda_{t,k}^y\}_{k=0}^{T-t})$ to capture self t 's mistakes. That is, how the actual MPCs in (19) deviate from the deliberate MPCs $\phi_t^{\text{Deliberate}}$ and $\{\phi_t^{\text{Deliberate}} \omega_{t,k}^{\text{Deliberate}}\}_{k=0}^{T-t}$ introduced below in (21). Specifically, the mistakes λ_t^a and $\{\lambda_{t,k}^y\}_{k=0}^{T-t}$ are given by: for all $t \in \{0, \dots, T-1\}$ and $k \in \{0, \dots, T-t\}$

$$\phi_t^a = (1 - \lambda_t^a) \phi_t^{\text{Deliberate}}, \quad \phi_t^y = (1 - \lambda_{t,0}^y) \phi_t^{\text{Deliberate}}, \quad \text{and} \quad \phi_t^y \omega_{t,k} = (1 - \lambda_{t,k}^y) \phi_t^{\text{Deliberate}} \omega_{t,k}^{\text{Deliberate}}, \quad (20)$$

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 (9), a positive λ means under-reaction and a negative λ means over-reaction. As in Section 3, the mistakes λ_t^a and $\{\lambda_{t,k}^y\}_{k=0}^{T-t}$ are treated as exogenous now but will be connected 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 $k \in \{0, \dots, T-t\}$. That is, the fungibility case analyzed above is a special case in which mistakes in response to different components of permanent income are restricted to be the same.

Based on Definition 1, from 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 $t \in \{0, \dots, T-1\}$, each self t 's deliberate consumption rule is given by:*

$$c_t^{\text{Deliberate}}(a_t, s_t) = \phi_t^{\text{Deliberate}} \left(a_t + y_t + \sum_{k=1}^{T-t} \omega_{t,k}^{\text{Deliberate}} R^{-k} y_{t+k} \right) + \bar{c}_t^{\text{Deliberate}}. \quad (21)$$

In (21) $\phi_t^{\text{Deliberate}}$ is a function of $\{\lambda_{t+l}^a\}_{l=1}^{T-t-1}$, δ , R . And $\omega_{t,k}$ is a function of $\{\lambda_{t+l}^a\}_{l=1}^{T-t-1}$, $\{\lambda_{t+l,k-l}^y\}_{l=1}^{\min\{k, T-t-1\}}$, δ , R .

In (21), $\phi_t^{\text{Deliberate}}$ captures the MPC of the 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 $\{\lambda_{t+l,k-l}^y\}$ because $\omega_{t,k}$ is about self t 's response to future income y_{t+k} and the relevant future mistake is $\{\lambda_{t+l,k-l}^y\}$, i.e., how the future self $t+l$ responds to income y_{t+k} .

In this Section, I will establish two general results about how future consumption mistakes impact current MPCs. First, the above high MPCs result remains to be true: as long as future consumption responds inefficiently to changes in savings/borrowings ($\lambda_{t+l}^a \neq 0$), current deliberate MPCs, i.e., $\phi_t^{\text{Deliberate}}$ in (21), 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 the current sophisticated 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 also violate the fungibility principle and respond differentially to changes in different components of permanent income.

4.2 High Current MPCs

Here, I show that the main results in Section 3, i.e., how future consumption mistakes lead to excess concavity of the continuation value function and high current MPCs, remain true. I further emphasize that the key behind these result is the inefficient responses of future consumption to changes in savings/borrowings.

Similar to Lemma 1, I use $\Gamma_t > 0$ to denote the ‘‘concavity’’ of the consumer’s continuation value function in (5): $\frac{\partial^2 V_t(a_t, s_t)}{\partial a_t^2} \equiv u'' \cdot \Gamma_t$.

Proposition 5. *i. Future consumption mistakes lead to excess concavity of the continuation value function: for $t \in \{0, \dots, T-2\}$, Γ_{t+1} strictly increases with $\{|\lambda_{t+l}^a|\}_{l=1}^{T-t-1}$.*

ii. Future consumption mistakes lead to high current deliberate MPCs: for $t \in \{0, \dots, T-2\}$, $\phi_t^{\text{Deliberate}} \geq \phi_t^{\text{Frictionless}}$ and increases with each of $\{|\lambda_{t+l}^a|\}_{l=1}^{T-t-1}$.

The intuition behind part (i) of Proposition 5 is similar to Lemma 1. Larger $\{|\lambda_{t+l}^a|\}_{l=0}^{T-t-1}$ 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 $\{\lambda_{t+l}^a\}_{l=0}^{T-t-1}$ 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 the concavity Γ_t .

The intuition behind part (ii) of Proposition 5 is similar to Proposition 2. With future consumption mistakes (larger $\{|\lambda_{t+l}^a|\}_{l=1}^{T-t-1}$), 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 $\{\lambda_{t+l}^a\}_{l=1}^{T-t-1}$ for the high current MPCs result are future selves' inefficient responses to changes in savings/borrowings. This principle has an independent use: for a behavior bias causing inefficiently differential responses of future consumption to different components of permanent income, it helps predict whether it contributes to the high current MPCs. For example, in the context of inattention, Corollary 9 below shows how future imperfect perception of wealth (i.e., savings/borrowings) increases current MPCs. On the other hand, as explained in the Appendix, if future selves are *only* inattentive to income, the current MPCs will not be influenced. In the Appendix B, I also use this principle to study when future selves' distorted expectations (Mullainathan, 2002; Rozsypal and Schlafmann, 2017; Bordalo et al., 2018; Azeredo da Silveira and Woodford, 2019) will lead to high current MPCs. For example, if future selves over-extrapolate based on their wealth, current MPCs will be higher.

4.3 Future Non-fungibility Begets Current Non-fungibility

Now, I turn to a new prediction.

Proposition 6. *Generically, the deliberate consumption in (21) violates the fungibility principle. That is, for $t \in \{0, \dots, T-2\}$ and $k \in \{0, \dots, T-t\}$, generically, $\omega_{t,k}^{\text{Deliberate}} \neq 1$. Here, generically is in the sense of the Euclidean measure of the product space generated by future selves'*

$$\text{mistakes} \left(\left\{ \lambda_{t+l}^a \right\}_{l=1}^{T-t-1}, \left\{ \lambda_{t+l,k-l}^y \right\}_{l=1}^{\min\{k, T-t-1\}} \right).$$

This result means that the inefficiently differential responses of future consumption to different components of permanent income, by itself, suffices 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 3 is rather special. There, future actual consumption exhibits the same degree of mistakes in responses to changes in different components of permanent income

$$\lambda_{t+l} = \lambda_{t+l}^a = \lambda_{t+l,k-l}^y \quad \forall l, k \quad (22)$$

In this case, the current deliberate consumption remains to follow the fungibility principle. Away from (22), generically, the 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 the fungibility. That is, 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 $l \in \{1, \dots, T-t-1\}$ and $k \in \{l, \dots, T-t+l\}$ (recall a larger λ means a smaller MPC). This case is consistent with the empirical evidence on small MPC 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), which will be further discussed below.

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 $l \in \{1, \dots, T-t-1\}$ and $k \in \{l, \dots, T-t+l\}$.*

The current deliberation consumption in (21) has the following properties: for $k \in \{0, \dots, T-t\}$,

- (i) $\omega_{t,k}^{\text{Deliberate}} \leq 1$. *That is, the current self excessively discounts future income.*
- (ii) $\omega_{t,k}^{\text{Deliberate}}$ *decreases with each $\{\lambda_{t+l}^a\}_{l=1}^{T-t-1}$ (i.e., increases with future selves' actual MPCs out of wealth) and increases with each $\{\lambda_{t+l,k-l}^y\}_{l=1}^{\min\{k, T-t-1\}}$ (i.e., decreases with future selves' actual MPC out of income).*

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

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 the Proposition, note: when future selves mistakenly respond too little to changes in savings/wealth (a larger $\{\lambda_{t+l}^a\}_{l=1}^{T-t-1}$), the excess concavity in Propo-

sition 6 means that the current self will be less willing to change her saving. As a result, the current self is less willing to adjust her current consumption in response to changes in future income, since the response of current consumption to future income requires changes in savings. As a result, there is excess discounting ($\omega_{t,k}^{\text{Deliberate}} < 1$) and $\omega_{t,k}^{\text{Deliberate}}$ decreases with each $\{\lambda_{t+l}^a\}_{l=1}^{T-t-1}$.

On the other hand, $\omega_{t,k}^{\text{Deliberate}}$ increases with each $\{\lambda_{t+k,l-k}^y\}_{k=1}^{\min\{l, T-t-1\}}$. That is, if future selves' mistakenly respond too little to changes in future income y_{t+k} (a larger $\{\lambda_{t+k,l-k}^y\}_{k=1}^{\min\{l, T-t-1\}}$), 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 empirical 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." In response to changes in future income y_{t+k} , the further away from period $t+k$, the more discounting. This is because the mechanism behind the excess discounting accumulates over the distance between current consumption and future income.

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

In the Appendix, I also study the case that mistakes in future consumption take the form of inefficiently large MPCs out of wealth. Though this case is potentially empirically irrelevant, the main lesson in Proposition 6 remains true: the non-fungibility of future consumption leads to the non-fungibility of the 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 statistics in part (ii) of Proposition 7.

4.4 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 6, the consumer's MPC remains the same with gradual resolution of income uncertainty.

Specifically, with a gradual resolution of the income uncertainty, the actual consumption rule

of each self $t \in \{0, \dots, T-1\}$ can now be written as

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

where $E_t[y_{t+k}] = E_t[y_{t+k}|s_t]$ captures the expected future income based on the current income state s_t . Self t 's mistakes λ_t^a and $\{\lambda_{t,k}^y\}_{k=0}^{T-t}$ are still given by (20).

Based on future selves' actual consumption rules $\{c_{t+k}(a_{t+k}, s_{t+k})\}_{k=0}^{T-1-t}$, each self t 's deliberate consumption rule defined in (3) will take the following form.

Corollary 7. *For $t \in \{0, \dots, T-1\}$, each self t 's deliberate consumption rule is given by:*

$$c_t^{Deliberate}(a_t, s_t) = \phi_t^{Deliberate} \left(a_t + y_t + \sum_{k=1}^{T-t} \omega_{t,k}^{Deliberate} R^{-k} E_t[y_{t+k}] \right) + \bar{c}_t^{Deliberate}, \quad (23)$$

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

The $T \rightarrow \infty$ and hand-to-mouth limit. Similar to Corollary 5, I can establish a simple limit for the deliberate consumption rule in (21) when the consumer's horizon T goes to infinity.

Corollary 8. *Fix a self t . Let $\lambda_{t+l}^a = \lambda^a$ with $|\lambda^a| < (\delta^{-1/2} R^{-1})$ and $\lambda_{t+l,k-l}^y = \lambda^y$ for all k and h . We have, when $T \rightarrow +\infty$,*

$$\begin{aligned} \phi_t^{Deliberate} &\rightarrow \phi^{Deliberate} \equiv \frac{\delta R^2 - 1}{\delta R^2 (1 - (\lambda^a)^2)}. \\ \omega_{t,k}^{Deliberate} &\rightarrow (\omega^{Deliberate})^k \equiv \left(1 - \frac{(\delta R^2 - 1) \lambda^a (\lambda^a - \lambda^y)}{1 - (\lambda^a)^2} \right)^k \end{aligned} \quad (24)$$

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

$$\phi^{Deliberate} \rightarrow 1 \quad \text{and} \quad \omega^{Deliberate} \rightarrow 0. \quad (25)$$

The limit in (25) is effectively a “hand-to-mouth” limit. When the current self is so 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 current income level ($c_t \neq y_t$).

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

4.5 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, which 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 λ s in (20). That is, why future consumption may exhibit differential responses to different components of income and wealth. The result in Propositions 6 and 7 then follows 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 1, I accommodate inattention within the fungible framework in Section 3. There, actual consumption is decided based on the same degree of attention to all components the permanent income, as in (13). 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 . 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 under inattention to wealth/savings a_t but full attention to the income state s_t . Specifically, similar to Corollary 1 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 also let each self perfectly perceive her current income state s_t . That is, for $t \in \{0, \dots, T-1\}$,

$$a_t^p(a_t) = (1 - \lambda_t^a) a_t + \lambda_t^a a_t^d \quad \text{and} \quad s_t^p(s_t) = s_t, \quad (26)$$

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 which the exact value does not matter for the MPCs). Also similar to Corollary 1, 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, as explained in Appendix.

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 their financial wealth changes and fail to adjust accordingly. Moreover, the recent literature on Fintech shows that providing information about a consumer's total wealth by aggregating her financial account 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 imperfect perception of wealth without the tool. Likewise, Carlin, Olafsson and Pagel (2017) study the introduction of 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 as well as other bank fees.

Based on the perceived wealth in (13), the actual consumption rule of each self t is given by

$$c_t(a_t^p(a_t), s_t) = \arg \max_{c_t} u(c_t) + \delta V_{t+1}(R(a_t^p(a_t) + y_t - c_t), s_{t+1}), \quad (27)$$

where the continuation value function V_{t+1} is defined similarly to above, based on future selves' actual consumption rules.

The deliberate consumption is decided as in (6), based on the correct current wealth and taking future selves' imperfect perception of wealth as given:

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

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

Corollary 9. *For each $t \in \{0, \dots, T-1\}$, self t 's deliberate consumption rule is given by*

$$c_t^{Deliberate}(a_t, s_t) = \phi_t^{Deliberate} \left(a_t + y_t + \sum_{k=1}^{T-t} \omega_{t,k}^{Deliberate} R^{-k} y_{t+k} \right) + \bar{c}_t^{Deliberate} \quad (28)$$

where $\phi_t^{Deliberate}$ and $\{\omega_{t,k}^{Deliberate}\}_{l=0}^{T-t}$ are given by the formula in Lemma 2 with λ_{t+l}^a given by (26) and all λ^y is zero. Moreover,

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

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 then naturally translate to properties of the actual MPCs.

Corollary 10. *For each $t \in \{0, \dots, T-1\}$, self t 's actual consumption rule in (19) has the following properties:*

- (i) The MPC out of current income $\phi_t^y \geq \phi_t^{Frictionless}$ and increases with future selves' degrees of imperfect perception of wealth $\{\lambda_{t+l}^a\}_{l=1}^{T-t-1}$.
- (ii) The MPC out of wealth is given by $\phi_t^a = (1 - \lambda_t^a) \phi_t^y$
- (iii) For $k \in \{0, \dots, T-t\}$, there is extra discounting of future income $\omega_{t,k} \leq 1$ and it decreases with future selves' degrees of imperfect perception of wealth $\{\lambda_{t+l}^a\}_{l=1}^{T-t-1}$.

In other words, Corollary 10 shows how current and anticipated future imperfect perception of wealth provides 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 liquid wealth than out of current income, and excess discounting of future income.

Finally, one can use the limit result in Corollary 8 to gauge the magnitudes of 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 choose to back out λ^a from relevant moments of MPCs in the data.

Specifically, in the $T \rightarrow \infty$ studied in Corollary 10, we have $\phi^a/\phi^y = 1 - \lambda^a$.⁸ This ratio ϕ^a/ϕ^y between the MPC out of wealth and the MPC out of current income is also 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, they imply $\lambda^a = 1 - 1/7 = 6/7$.⁹ 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.¹⁰

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.

5 Other Applications

The main application in the paper is to show future consumption mistakes can explain high-liquidity consumers' high MPCs and non-fungibility, as studied above. The key mechanism behind those 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 3 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 $-\frac{\frac{\partial^2 V_t}{\partial w_t^2}}{\frac{\partial V_t}{\partial w_t}}$,

⁸This assumes perfect attention to the income state s_t ($\lambda^y = 0$) as in (26). For a given ϕ^a/ϕ^y , if I allow inattention to income state, the implied degree of imperfect perception of wealth ($\lambda^a = 0$) will be larger.

⁹In their published version (Maggio, Kermani and Majlesi, 2019), their estimates of MPC out of current income are slightly larger than their NBER version above, which implies an even higher larger λ^a and hence even large frictions. To be conservative, I use their estimates in the NBER working paper version above.

¹⁰Using 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.

both proportional to $\frac{\partial^2 V_t}{\partial w_t^2}$. From Lemma 1 about excess concavity, we know that consumption mistakes lead to 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 $T \rightarrow +\infty$ limit in Corollary 5. In this limit, we have $\Gamma_t \equiv \frac{\partial^2 V_t}{\partial w_t^2} / u'' \rightarrow \Gamma = \frac{\delta R^2 - 1}{\delta R^2 (1 - \delta R^2 \lambda^2)}$. With the calibration of λ used in Section 3 ($\lambda = 0.56$ 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). The 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, the continuation value function is excessively concave; the current self is less willing to respond to change her savings and 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 .¹¹

Proposition 8. *The response of deliberate consumption to interest rate changes, $\left| \frac{\partial c_t^{\text{Deliberate}}}{\partial R_t} \right|$, decreases with each future self's mistake $\left\{ \left| \lambda_{t+k}^a \right| \right\}_{k=1}^{T-t-1}$.*

6 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 from liquidity constraints and helps resolve the empirical puzzles on high liquidity consumers' high MPCs. The main approach, using “wedges” to capture behavioral mistakes and deriving robust predictions independent from the exact psychological cause of these mistakes, can also be useful in other contexts.

¹¹The zero net saving condition guarantees that responses to interest rate changes are 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 results in response to income changes.

For example, in ongoing works, I am exploring how mistakes in future consumption levels briefly discussed after Proposition 4, can robustly impact current decisions.

Appendix A: Proofs

Proof of Proposition 1. The definition of deliberate consumption in (3) at t together with the definition of the value function in (5) at $t + 1$ lead to (6). The recursive formulation for the value function in (7) follows directly from the definition of the value function in (5).

Now, consider consumption rules and value functions $\{c_t^{\text{Deliberate}}(a_t, s_t), c_t(a_t, s_t)\}_{t=0}^{T-1}$ and $\{V_t(a_t, s_t)\}_{t=0}^T$ satisfy (4), (6), (7), 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 (3) – (5).

Proof of Lemma 1 and Proposition 2. We work with backward induction. At T , we have:

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

For each $t \leq T - 1$, from (6), the deliberate MPC is given by

$$\phi_t^{\text{Deliberate}} = \frac{\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}}. \quad (29)$$

From (9), the actual MPC is given by

$$\phi_t = \frac{(1 - \lambda_t) \delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}}. \quad (30)$$

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

$$\frac{\partial V_t(w_t)}{\partial w_t} = \phi_t u'(c_t(w_t)) + (1 - \phi_t) \delta R \frac{\partial V_{t+1}(w_{t+1})}{\partial w_{t+1}}. \quad (31)$$

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

$$\begin{aligned} \Gamma_t &= (\phi_t)^2 + (1 - \phi_t)^2 \Gamma_{t+1} \delta R^2 \\ &= (1 + \Gamma_{t+1} \delta R^2) \left(\phi_t - \frac{\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}} \right)^2 + \frac{\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}} \\ &= \frac{(\delta R^2 \Gamma_{t+1})^2}{1 + \delta R^2 \Gamma_{t+1}} \lambda_t^2 + \frac{\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}}. \end{aligned} \quad (32)$$

Lemma 1 and Proposition 2 then follow directly.

Proof of Proposition 3. Here, I follow O'Donoghue and Rabin (2001) and introduce “partial sophistication.” Specifically, for $k \in \{1, \dots, T - t - 1\}$, the deliberate consumption $c_t^{\text{Deliberate}}$ is determined with a partial understanding of her future selves' mistakes:

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

where $s_t \in [0, 1]$ captures deliberate self t 's degree of sophistication and $\tilde{\lambda}_{t,t+k}$ captures deliberate self t 's perceived mistake in her actual consumption at $t + k$.

With similar notations, I use $\tilde{c}_{t,t+k}(w_{t+k})$ to denote deliberate self t 's perceived actual consumption rule at $t + k$. At t , the deliberate consumption $c_t^{\text{Deliberate}}(w_t)$ is then given by

$$c_t^{\text{Deliberate}}(w_t) \equiv \arg \max_{c_t} u(c_t) + \sum_{k=1}^{T-t-1} \delta^{k-1} u(\tilde{c}_{t,t+k}(w_{t+k})) + \delta^{T-t} v(w_T), \quad (33)$$

with $w_{t+k} = R(w_{t+k-1} - \tilde{c}_{t,t+k-1}(w_{t+k-1}))$. Furthermore, $\tilde{c}_{t,t+k}(w_{t+k})$ is given by

$$\tilde{c}_{t,t+k}(w_{t+k}) = \mathcal{S}(\tilde{c}_{t,t+k}^{\text{Deliberate}}(w_{t+k}), \tilde{\lambda}_{t,t+k}), \quad (34)$$

where $\tilde{c}_{t,t+k}^{\text{Deliberate}}(w_{t+k})$ captures the consumption self $t + k$ would have chosen based on deliberate self t 's perceived actual consumption rule $\{\tilde{c}_{t,t+k+l}(w_{t+k+l})\}_{l=1}^{T-t-k-1}$:

$$\tilde{c}_{t,t+k}^{\text{Deliberate}}(w_{t+k}) \equiv \arg \max_{c_{t+k}} u(c_{t+k}) + \sum_{l=1}^{T-t-k-1} \delta^{k-1} u(\tilde{c}_{t,t+k+l}(w_{t+k+l})) + \delta^{T-t} v(w_T), \quad (35)$$

with $w_{t+k+l} = R(w_{t+k+l-1} - \tilde{c}_{t,t+k+l-1}(w_{t+k+l-1}))$. Finally, let $\tilde{V}_{t,t+1}(\cdot)$ be self t 's perceived continuation value function based on her perceived future actual consumption rules:

$$\tilde{V}_{t,t+1}(w_{t+1}) = \sum_{k=1}^{T-t-1} \delta^{k-1} u(\tilde{c}_{t,t+k}(w_{t+k})) + \delta^{T-t} v(w_T), \quad (36)$$

with $w_{t+k} = R(w_{t+k-1} - \tilde{c}_{t,t+k-1}(w_{t+k-1}))$.

From (33) - (35), we know that $\tilde{V}_{t,t+1}(\cdot)$ coincides with the actual continuation value function in (5) if future selves' actual mistakes are given by $\{\tilde{\lambda}_{t,t+k}\}_{k=1}^{T-t-1}$. Proposition 3 then follows directly.

Proof of Corollary 1. From (14), we have

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

while

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

Because both u and V_{t+1} are quadratic, u' and V'_{t+1} are linear. Together with (13), we know, in this case, $\phi_t = (1 - \lambda_t) \phi_t^{\text{Deliberate}}$, where λ_t is the degree of inattention in (13). Corollary 1 then follows directly.

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

The value function in (7) is now given by $V_t (a_t, s_t) = E_t [u (c_t (a_t, s_t)) + \delta V_{t+1} (R (a_t + y_t - c_t (a_t, s_t)), s_{t+1})]$ where $E_t [\cdot]$ averages over the potential realizations of actual consumption rule. The deliberate consumption in (6) is unchanged.

In the proof of Proposition 2, the deliberate MPC is still given by (6), but (7) becomes

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

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 2 then follows directly from (6).

Proof of Corollary 3. From (15) and (16), we have

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

and

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

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

$$\phi_t = \frac{\delta \beta_t R^2 \Gamma_{t+1}}{1 + \delta \beta_t R^2 \Gamma_{t+1}} \quad \text{and} \quad \phi_t^{\text{Deliberate}} = \frac{\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}}, \quad (37)$$

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,

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

where

$$\lambda_t = \frac{1 - \beta_t}{1 + \delta \beta_t R^2 \Gamma_{t+1}}. \quad (38)$$

Substitute into (32), we have

$$\Gamma_t = \frac{(\delta R^2 \Gamma_{t+1})^2}{1 + \delta R^2 \Gamma_{t+1}} \left(\frac{1 - \beta_t}{1 + \delta \beta_t R^2 \Gamma_{t+1}} \right)^2 + \frac{\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}}. \quad (39)$$

Define $f(x, \beta) = \frac{x^2}{1+x} \left(\frac{1-\beta}{1+\beta x} \right)^2 + \frac{x}{1+x}$. We have $f_\beta(x, \beta) = -\frac{2(1-\beta)x^2}{(1+x\beta)^3}$ and $f_x(x, \beta) = \frac{1+x\beta(-1+2\beta)}{(1+x\beta)^3}$. As a result, for $x \geq 0$ and $\beta \in [\frac{1}{2}, 1]$, we have $f_\beta(x, \beta) \leq 0$ and $f_x(x, \beta) \geq 0$. Using these properties in (39), we know Γ_t decreases with $\{\beta_{t+k}\}_{k=0}^{T-t-1}$. Corollary 3 then follows from (2).

Proof of Corollary 4. This case is not directly nested in Proposition 2, as the actual consumption rule is stochastic. But the proof is essentially unchanged.

The value function in (7) is now given by

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

where $E_t[\cdot]$ averages over the potential realizations of actual consumption rule. The deliberate consumption in (6) is unchanged.

In the proof of Proposition 2, the deliberate MPC is still given by (6), but (7) becomes

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

As a result, Γ_t increases with $\{\text{Var}(\varphi_{t+k})\}_{k=0}^{T-t-1}$. Corollary 2 then follows directly from (6).

Proof of Corollary 5. From (32), we know that $\Gamma_t = \frac{(\delta R^2 \Gamma_{t+1})^2}{1 + \delta R^2 \Gamma_{t+1}} \lambda^2 + \frac{\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}} \equiv f(\Gamma_{t+1})$, with $f(x) \equiv \frac{\delta R^2 x}{1 + \delta R^2 x} + \frac{(\delta R^2 x)^2}{1 + \delta R^2 x} \lambda^2 = \frac{\delta R^2 x}{1 + \delta R^2 x} (1 + \lambda^2 \delta R^2 x)$. 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 $\Gamma > \Gamma_t = f^{(T-t)}(\Gamma_T) > f^{(T-t-1)}(\Gamma_T) > \dots > \frac{v''}{u''} = \Gamma_T$. As a result, $\Gamma_t = f^{(T-t)}(\Gamma_T)$ converges to the fix point Γ with $T \rightarrow +\infty$.

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

Together, one way or another, as long as $0 \leq \lambda < \delta^{-1/2}R^{-1}$, $\Gamma_t \rightarrow \Gamma$ with $T \rightarrow +\infty$. From (29), we then have, with $T \rightarrow +\infty$.

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

Proof of Corollary 6. With graduate resolution of uncertainty, the optimal deliberate consumption in (6) becomes

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

while the recursive formulation for the value function in (7) becomes

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

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 of Proposition 1 remains unchanged, except (31) 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 (29), (31), and (32) remain unchanged. So Corollary 6 follows directly.

Proof of Proposition 4. The recursive formulation in Proposition 1 remains to hold. The optimal deliberate consumption now is given by¹²,

$$u'(c_t^{\text{Deliberate}}(w_t)) = R \delta V'_{t+1}(R(w_t - c_t^{\text{Deliberate}}(w_t))). \quad (40)$$

We henceforth have:

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

¹²This 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.

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

$$\frac{\partial c_t^{\text{Deliberate}}(\tilde{w}_t)}{\partial w_t} = \frac{R^2 \delta \frac{\partial^2 V_{t+1}(\tilde{w}_{t+1})}{\partial w_{t+1}^2}}{u''(c_t^{\text{Deliberate}}(\tilde{w}_t)) + R^2 \delta \frac{\partial^2 V_{t+1}(\tilde{w}_{t+1})}{\partial w_{t+1}^2}}. \quad (41)$$

From (6):

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

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

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

At \tilde{w}_t , because $c_t(\tilde{w}_t) = c_t^{\text{Deliberate}}(\tilde{w}_t) = \tilde{c}_t$, from (40), 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,

$$\frac{\partial^2 V_t(\tilde{w}_t)}{\partial w_t^2} = \left(\frac{\partial c_t(\tilde{w}_t)}{\partial w_t}\right)^2 u''(c_t(\tilde{w}_t)) + \left(1 - \frac{\partial c_t(\tilde{w}_t)}{\partial w_t}\right)^2 \delta R^2 \frac{\partial^2 V_{t+1}(\tilde{w}_{t+1})}{\partial w_{t+1}^2}. \quad (42)$$

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 (41) and (42), 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

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

Proposition 4 then follows.

Proof of Lemma 2. Similar to (10), we define $\{\Gamma_t, \Gamma_{t,k}^y\}_{t \in \{0, \dots, T\}, k \in \{0, \dots, T-t\}}$ based on

$$\frac{\partial V_t}{\partial a_t} \equiv u'' \cdot \left(\Gamma_t a_t + \sum_{k=0}^{T-t} \Gamma_{t,k}^y R^{-k} y_{t+k} + \bar{\Gamma}_t \right). \quad (43)$$

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 (6), the deliberate consumption is given by

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

Together (43) at $t+1$, we have

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

with

$$\phi_t^{\text{Deliberate}} = \frac{\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}} \quad (44)$$

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

$$\omega_{t,k}^{\text{Deliberate}} = \frac{\delta R R^{-(k-1)} \Gamma_{t+1,k-1}^y}{1 + \Gamma_{t+1} \delta R^2} / (\phi_t^{\text{Deliberate}} R^{-k}) = \frac{\Gamma_{t+1,k-1}^y}{\Gamma_{t+1}}. \quad (45)$$

Now, from the recursive formulation of the value function in (6), we have:

$$\frac{\partial V_t(a_t, s_t)}{\partial a_t} = \phi_t^a u'(c_t(a_t, s_t)) + (1 - \phi_t^a) \delta R \frac{\partial V_{t+1}(a_{t+1}, s_{t+1})}{\partial a_{t+1}}. \quad (46)$$

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

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

Together with (20), we have, for all $t \in \{0, \dots, T-1\}$:

$$\begin{aligned} \Gamma_t &= \phi_t^a (\phi_t^a - (1 - \phi_t^a) \delta R^2 \Gamma_{t+1}) + (1 - \phi_t^a) \delta R^2 \Gamma_{t+1} \\ &= (1 + \delta R^2 \Gamma_{t+1}) \left(\phi_t^a - \frac{\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}} \right)^2 + \frac{\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}} \\ &= \frac{(\delta R^2 \Gamma_{t+1})^2}{1 + \delta R^2 \Gamma_{t+1}} (\lambda_t^a)^2 + \frac{\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}}, \end{aligned} \tag{47}$$

and

$$\begin{aligned} \Gamma_{t,0}^y &= \phi_t^y (\phi_t^a - (1 - \phi_t^a) \delta R^2 \Gamma_{t+1}) + (1 - \phi_t^a) \delta R^2 \Gamma_{t+1} \\ &= (1 + \delta R^2 \Gamma_{t+1}) \left(\phi_t^a - \frac{\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}} \right) \left(\phi_t^y - \frac{\beta R^2 \Gamma_{t+1}}{1 + \beta R^2 \Gamma_{t+1}} \right) + \frac{\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}} \\ &= \frac{(\delta R^2 \Gamma_{t+1})^2}{1 + \delta R^2 \Gamma_{t+1}} \lambda_t^a \lambda_{t,0}^y + \frac{\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}}, \end{aligned} \tag{48}$$

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

$$\begin{aligned} \Gamma_{t,k}^y &= \phi_t^y \omega_{t,k} (\phi_t^a - (1 - \phi_t^a) \delta R^2 \Gamma_{t+1}) + (1 - \phi_t^a) \delta R^2 \Gamma_{t+1,k-1}^y \\ &= (1 + \delta R^2 \Gamma_{t+1}) \left(\phi_t^a - \frac{\delta R^2 \Gamma_{t+1}}{1 + \delta R^2 \Gamma_{t+1}} \right) \left(\phi_t^y \omega_{t,k} - \frac{\delta R^2 \Gamma_{t+1,k-1}^y}{1 + \delta R^2 \Gamma_{t+1}} \right) + \frac{\delta R^2 \Gamma_{t+1,k-1}^y}{1 + \delta R^2 \Gamma_{t+1}} \\ &= \frac{(\delta R^2)^2 \Gamma_{t+1} \Gamma_{t+1,k-1}^y}{1 + \delta R^2 \Gamma_{t+1}} \lambda_t^a \lambda_{t,k}^y + \frac{\delta R^2 \Gamma_{t+1,k-1}^y}{1 + \delta R^2 \Gamma_{t+1}}. \end{aligned} \tag{49}$$

Lemma 2 follows from (44) – (49).

Proof of Proposition 5. From Lemma 2, we know the expressions for ϕ_t^a , $\phi_t^{\text{Deliberate}}$, and Γ_t here are identical to those in Lemma 1 and Proposition 2, with $\{\phi_t^a\}_{t=0}^{T-1}$ replacing the role of $\{\phi_t\}_{t=0}^{T-1}$ and $\{\lambda_t^a\}_{t=0}^{T-1}$ replacing the role of $\{\lambda_t\}_{t=0}^{T-1}$. Proposition 5 then follows directly from Lemma 1 and Proposition 2.

Proof of Proposition 6. From (45), (47), and (48), for $t \in \{0, \dots, T-2\}$, we have

$$\omega_{t,1}^{\text{Deliberate}} = \frac{\delta R^2 \Gamma_{t+2} \lambda_{t+1}^a \lambda_{t+1,0}^y + 1}{\delta R^2 \Gamma_{t+2} (\lambda_{t+1}^a)^2 + 1} = 1 - \frac{\delta R^2 \Gamma_{t+2} \lambda_{t+1}^a (\lambda_{t+1}^a - \lambda_{t+1,0}^y)}{\delta R^2 \Gamma_{t+2} (\lambda_{t+1}^a)^2 + 1}, \quad (50)$$

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

From (45), (47), and (49), for $t \in \{0, \dots, T-2\}$ and $k \in \{2, \dots, T-t\}$, we have

$$\begin{aligned} \omega_{t,k}^{\text{Deliberate}} &= \frac{\Gamma_{t+1,k-1}^y}{\Gamma_{t+1}} = \frac{\delta R^2 \Gamma_{t+2} \lambda_{t+1}^a \lambda_{t+1,k-1}^y + 1}{\delta R^2 \Gamma_{t+2} (\lambda_{t+1}^a)^2 + 1} \frac{\Gamma_{t+2,k-2}^y}{\Gamma_{t+2}} \\ &= \left[1 - \frac{\delta R^2 \Gamma_{t+2} \lambda_{t+1}^a (\lambda_{t+1}^a - \lambda_{t+1,k-1}^y)}{\delta R^2 \Gamma_{t+2} (\lambda_{t+1}^a)^2 + 1} \right] \omega_{t+1,k-1}^{\text{Deliberate}}. \end{aligned} \quad (51)$$

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

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 $l \in \{1, \dots, T-t-1\}$ and $k \in \{l, \dots, T-t+l\}$.

(i) This comes directly from (50) and (51).

(ii) The comparative statics with respect to $\{\lambda_{t+l,k-l}^y\}_{l=1}^{\min\{k, T-t-1\}}$ come directly from (50) and (51). To prove comparative statics with respect to $\{\lambda_{t+l}^a\}_{l=1}^{T-t-1}$, define:

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

We have

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

As a result, $\frac{\partial f}{\partial \lambda^a}(\Gamma, \lambda^y, \lambda^a) \leq 0$ if $\lambda^a \geq \lambda^y$ and $\lambda^a \geq 0$. Applying this result in (50) and (51), we know $\omega_{t,k}^{\text{Deliberate}}$ decreases with each $\{\lambda_{t+l}^a\}_{l=1}^{T-t-1}$.

(iii) This comes directly (51).

Proof of Corollary 7. With graduate resolution of uncertainty, the optimal deliberate consumption in (6) 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 (7) 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 (44) – (49) remain to be true. So Corollary 7 follows directly.

Proof of Corollary 8. Similar to Corollary 5, we have, when $T \rightarrow +\infty$,

$$\begin{aligned} \phi_t^{\text{Deliberate}} &\rightarrow \phi^{\text{Deliberate}} \equiv \frac{\delta R^2 - 1}{\delta R^2 (1 - (\lambda^a)^2)} \\ \Gamma_t &\rightarrow \Gamma \equiv \frac{\delta R^2 - 1}{\delta R^2 (1 - \delta R^2 (\lambda^a)^2)} \end{aligned}$$

From (50) and (51), we know

$$\omega_{t,k}^{\text{Deliberate}} \rightarrow (\omega^{\text{Deliberate}})^k,$$

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

Proof of Corollary 9 and Corollary 10. From (26) and (27), 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 (26) and $\lambda_{t,k}^y = 0$ for all t, k . Corollary 9 and Corollary 10 then follow from (27), Lemma 1, and Propositions 5 – 7.

Proof of Proposition 8. Consider the environment in Section 4. 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., $\tilde{a}_{t+1} = 0$.¹³

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 (5). Self t 's deliberate consumption is given by

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

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

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

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 path¹⁴ 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

Noisy signal approach to inattention.

In the inattention cases studied in Corollaries 1 and 1, 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 1 as an example to illustrate. The non-fungibility case in 1 follows similarly. I assume Normally distributed exogenous fundamentals, i.e. $w_0 \sim \mathcal{N}(0, \sigma_{w_0}^2)$.¹⁵

¹³On this path, actual consumption coincides with the deliberate consumption $\tilde{c}_t = c_t(\tilde{a}_t, \tilde{s}_t) = c_t^{\text{Deliberate}}(\tilde{a}_t, \tilde{s}_t)$.

¹⁴This comes from (46) 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 $\tilde{c}_{t+1} = c_{t+1}^{\text{Deliberate}}(\tilde{a}_{t+1}, \tilde{s}_{t+1})$.

¹⁵This together with the linear actual consumption rule in (55) guarantees that each w_t is Normally distributed

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 $\epsilon_t \sim \mathcal{N}(0, \sigma_{\epsilon_t}^2)$ and is independent from 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.

$$E[w_t | x_t] = (1 - \lambda_t)x_t, \quad (52)$$

where $\lambda_t = \frac{\text{Var}(\epsilon_t)}{\text{Var}(w_t) + \text{Var}(\epsilon_t)} \in [0, 1]$ 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

$$c_t(x_t) = \arg \max_{c_t} u(c_t) + \delta E[V_{t+1}(R(w_t - c_t)) | x_t], \quad (53)$$

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 11. *Each self t 's deliberate MPC $\phi_t^{\text{Deliberate}}$ is the same as that in Corollary 1, based on $\{\lambda_{t+k}\}$ defined above.*

Proof of Corollary 11. The value in (7) is now given by

$$V_t(w_t) = \int [u(c_t(w_t + \epsilon_t)) + \delta V_{t+1}(R(w_t - c_t(w_t + \epsilon_t)))] f_t(\epsilon_t) d\epsilon_t, \quad (54)$$

where $f_t(\cdot)$ is the p.d.f. given $\epsilon_t \sim \mathcal{N}(0, \sigma_{\epsilon_t}^2)$. Similar to (10), 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 (11) and (6). For the actual consumption in (53), we have

$$\begin{aligned} c_t(x_t) &= (1 - \lambda_t)\phi_t^{\text{Deliberate}}(w_t + \epsilon_t) + \bar{c}_t^{\text{Deliberate}}, \\ &= \phi_t w_t + \phi_t \epsilon_t + \bar{c}_t^{\text{Deliberate}} \end{aligned} \quad (55)$$

where I use (52) and $\phi_t \equiv (1 - \lambda_t)\phi_t^{\text{Deliberate}}$ as in the main text.

too.

From (54), 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 (32) 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 11 then follows.

Hyperbolic discounting.

Here I establish some additional results regarding the hyperbolic discounting in Corollary 3.

First, I let $\beta_t = \beta$ for all t and consider the $T \rightarrow \infty$ limit. From (39), we know $\Gamma_t \rightarrow \Gamma$ 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 (38), 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 3.3.

Second, let me derive the hyperbolic Euler equation in Harris and Laibson (2001) based our framework here. From (15) 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 (31), 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

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

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).¹⁶ 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 4, I capture inattention to income similar to (26). That is, for $t \in \{0, \dots, T-1\}$, 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 $t \in \{0, \dots, T-1\}$,

$$s_t^p(s_t) = (1 - \lambda_t^y) s_t + \lambda_t^y s_t^d \quad \text{and} \quad a_t^p(a_t) = a_t, \quad (56)$$

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 of which the exact value does not matter for the MPCs). Recall that, in the environment here, since all income uncertainty is resolved in period 0, $s_t = (y_t, \dots, y_T)$.

Based on the perceived income state in (56), the actual consumption rule of each self t is given by

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

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,

¹⁶Sims (2003) also studies the inattention to exogenous initial wealth, which effectively plays the same role as exogenous income.

the deliberate consumption is decided as in (6), 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) 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 12. *For each $t \in \{0, \dots, T-1\}$, self t 's actual consumption is given by*

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

where $\{y_{t+k}^p\}_{k=0}^{T-t}$ 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 $\{\lambda_{t+k}^a\}_{k=0}^{T-t-1}$. 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 11. The case here is nested by the general case studied in Lemma 1 with $\lambda_t^a = 0$ given by (26) and $\lambda_{t,k}^y = \lambda_t^y$ for all $t \in \{0, \dots, T-t\}$ and $k \in \{0, \dots, T-t\}$

In the proof of Propositions 5 and 6, if all $\lambda_t^a = 0$, we have $\phi_t^{\text{Deliberate}} = \phi_t^{\text{Frictionless}}$ and $\omega_{t,k}^{\text{Deliberate}} = 0$ for $k \in \{0, \dots, T-t\}$. Corollary 11 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). 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 et al. (2018).

The fungibility case. Let me start from the simple fungibility case in Section 3, I summarize such a friction by letting each self t 's perceived permanent income be given by

$$w_t^p(w_t) = w_t + \theta_t (w_t - w_t^d) \quad \forall t \in \{0, \dots, T-1\}, \quad (57)$$

where θ_t captures self t 's degree of distorted expectations and w_t^d captures the default (an exogenous constant of which the exact 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.¹⁷

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

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

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

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

taking future selves' consumption mistakes as given, driven by future selves' distorted expectations. As a corollary of Proposition 2, these future consumption mistakes lead to a high MPC of the current deliberate consumption.

Corollary 13. *For $t \in \{0, \dots, T-2\}$, $\phi_t^{\text{Deliberate}} \equiv \frac{\partial c_t^{\text{Deliberate}}}{\partial w_t} > \phi_t^{\text{Frictionless}}$ 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 4. Similar to the discussion after 5 and the inattention case studied in Corollaries 9 and 12, 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 (57), 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 13.

¹⁷In fact, when $\theta_t < 0$, the case here is the same as the inattention case studied above.

On the other hand, if future selves' distorted income expectations are fully driven by incomes and independent from savings/wealth (e.g. Mullainathan, 2002; Rozsypal and Schlafmann, 2017; Azeredo da Silveira and Woodford, 2019), these future mistakes will not directly impact current MPCs. This is the same as result in Corollary 12: future selves' inattention to their income will not impact current MPCs.

For example, consider the following variant of (57) regarding distorted expectations about future income:

$$y_{t+k}^p(y_{t+k}) = y_{t+k} + \theta_{t,k}(y_{t+k} - y_{t+k}^d) \quad \forall t \in \{0, \dots, T-1\},$$

where $\theta_{t,k}$ captures self t 's degree of distorted expectations with regard to y_{t+k} and y_{t+k}^d 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 14. *For $t \in \{0, \dots, T-1\}$, each self's deliberate consumption in (21) coincides with the frictionless one. That is, $\phi_t^{\text{Deliberate}} = \phi_t^{\text{Frictionless}}$ and $w_{t,k}^{\text{Deliberate}} = 1$ for all $k \in \{0, \dots, T-t\}$.*

Proof of Corollary 11. This case is nested in Proposition 2 with $\lambda_t = -\theta_t$.

Proof of Corollary 14. This case is nested in Lemma 2 with $\lambda_t^a = 0$ and $\lambda_{t,k}^y = -\theta_{t,k}$ for all $t \in \{0, \dots, T-1\}$ and $k \in \{0, \dots, T-t\}$. In the proof of Propositions 5 and 6, if all $\lambda_t^a = 0$, we have $\phi_t^{\text{Deliberate}} = \phi_t^{\text{Frictionless}}$ and $\omega_{t,k}^{\text{Deliberate}} = 0$ for $k \in \{0, \dots, T-t\}$. Corollary 14 follows.

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.
- Azeredo da Silveira, Rava, and Michael Woodford.** (2019) “Noisy Memory and Over-Reaction to News.” Vol. 109, 557–61.
- 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.
- Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer.** (2018) “Over-reaction in Macroeconomic Expectations.” *NBER Working Paper No. 24932*.
- 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.
- Kahneman, Daniel.** (2011) *Thinking, Fast and Slow*. Macmillan.
- 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.
- 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.
- 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 Schläpman.** (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.
- Thaler, Richard.** (1990) “Anomalies: Saving, Fungibility, and Mental Accounts.” *Journal of Economic Perspectives*, 4(1): 193–205.