IZA Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16372

Eggs When Young, Chicken When Old. Time Consistency and Addiction over the Life Cycle

Luca Piccoli
Silvia Tiezzi

AUGUST 2023

DISCUSSION PAPER SERIES

IZA DP No. 16372

Eggs When Young, Chicken When Old. Time Consistency and Addiction over the Life Cycle

Luca Piccoli
University of Trento and IZA

Silvia Tiezzi
University of Siena

AUGUST 2023

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Eggs When Young, Chicken When Old. Time Consistency and Addiction over the Life Cycle

This paper tests whether young and adult smokers have different time preferences, in particular with respect to time consistency. The recent introduction of Tobacco 21 law in the US were in part motivated by allegedly inconsistent time preferences of the young consumers. This research empirically tests this hypothesis using individual cigarettes consumption longitudinal data from RLMS, estimating a quasi-hyperbolic discounting rational addiction model for young and adult smokers separately. While our test rejects time inconsistency in the form of present-bias for both population groups, young smokers are found to discount future utilities much more than adults. From a life-cycle perspective, this is still a form of time inconsistency, which provides partial empirical support to the T21 law motivation, but also highlights how the quasi-hyperbolic discounting formulation might not be able to properly capture long-run time preferences.

JEL Classification: C23, D03, D12

Keywords: rational addiction, smoking behavior, time inconsistency, young vs adult smokers, GMM

Corresponding author:
Luca Piccoli
Department of Sociology and Social Research
University of Trento
Via Verdi, 26
38122 Trento
Italy

E-mail: luca.piccoli@unitn.it

1. Introduction

In the analysis of intertemporal consumption models, the assumption of exponential discounting has been criticized as early as in the middle of the past century (Strotz, 1956). To overcome such critics and maintain a parsimonious specification, the well known quasi-hyperbolic discounting has been proposed, which is characterized by present-biased intertemporal preferences where all future utilities are further discounted by a constant term β (see, for instance, Pollak, 1968; Laibson, 1997). This model allows consumers to be time inconsistent, meaning that their present optimal consumption plan will not be automatically respected in the future, because when the future becomes the present, present-biased preferences will kick in and the previous choices will not be optimal anymore. This happens to naïve consumers that do not account for the fact that in the future they will still be present-biased.[1] This is still a very relevant research topic today (see Grossman, 2022), especially because until very recently, empirical tests of the quasi-hyperbolic discounting have been limited to the experimental setting (see Blow et al., 2021, for an overview of such attempts). Although the experimental method has some limits, the results generally confirm that quasi-hyperbolic discounting tends to outperform the exponential discounting model.

An empirical test of quasi-hyperbolic discounting using commonly available survey data, such as household expenditure surveys, has been recently introduced by Blow et al. (2021). The authors propose a revealed preference characterization of the quasi-hyperbolic consumption model and easy conditions for testing its validity against the standard exponential discounting assumption. They also show that, with their data, the behavior of naïve versus sophisticated consumers is empirically indistinguishable.

The hypothesis of time consistent agents is even more critical when analyzing the consumption of addictive goods, for which the Rational Addiction (RA henceforth) model (Becker and Murphy, 1988; Becker et al., 1994) has been widely criticized. In fact, the model's assumption of an addicted consumer being able to stick to an optimal long run consumption plan strucks with real life observations, such as people trying hard to stop smoking but not being able to, for instance. Quasi-hyperbolic discounting has been first introduced in the RA framework by Gruber and Köszegi (2001), where dynamic inconsistency can deliver radically different implications for government policies. In particular, while time consistency implies that the optimal tax on addictive goods should

[1] Sophisticated consumers that are fully aware of their present-biased preferences will face a different optimization problem, in the form of a sequential game played against future selves.

depend only on the externalities imposed on the society, time inconsistency suggests a much higher tax depending also on the "internalities" that addictive goods impose on consumers' future selves (Gruber and Köszegi, 2001; O'Donoghue and Rabin, 2006). Despite the notable relevance of such a model, because its empirical specification was believed to be indistinguishable from that of the general specification of the RA model,[2] no empirical tests for time inconsistency in consumption of addictive goods have emerged until Piccoli and Tiezzi (2021). The authors show how the general specification of the quasi-hyperbolic RA model allows to test for the β parameter being different from 1, and how to identify its upper bound. Using the Russian Longitudinal Monitoring Survey for years 2006 to 2018, the study cannot reject the null of a $\beta = 1$ for a sample of adult smokers, and finds an upper bound for the present bias parameter of 0.99.

In the present paper, we address a related question that emerged with the recent approval of the State Tobacco 21 laws in the US, which raised the minimum legal purchasing age of tobacco products to 21 years (Hansen et al., 2022). In addition to the usual health concerns of smoking, especially at young ages, one of the motivations that emerged in the debate was that young people are more likely to display time inconsistent and present-biased preferences (Crettez and Deloche, 2021), and thus tend to be unable to correctly evaluate the future consequences of smoking and more likely to develop addiction. Empirical evidence based on representative data, however, is still lacking. This paper fills the gap by testing whether young smokers aged 21 or less display evidence of quasi-hyperbolic discounting, and compare their time preferences to those of the adult population (aged 30 to 65) using data from the Russian Longitudinal Monitoring Survey for years 1994 to 2020.

The results confirm that adult Russian smokers are not time inconsistent and that their discount rate conforms with what is generally found in the empirical literature. Young smokers do not display present-biased preferences either, with an upper-bound limit for β equal to 0.99, identical to that of the adults. However, their discount factor is much smaller, 0.54 vs 0.97, indicating that they discount future utilities much more than the adults. This may suggest a more complex form of time inconsistency that cannot be captured by the quasi-hyperbolic discounting, at least in a life-cycle perspective.

[2] The general specification of the RA model, rarely estimated in the literature, embeds current, lead and lagged prices of the good in the demand equation. This is in contrast with the simplified specification, more frequently estimated, which only includes current prices but needs to impose that addiction is not persistent over time.

The remainder of the paper is structured as follows. Section 2 presents a short theoretical description of the model used in the empirical analysis. Section 3 describes the data used, sample selections and the empirical specification. Section 4 presents the results and Section 5 concludes.

2. Method and data

The model specification is as follows. Individuals are assumed to maximize the sum of lifetime discounted utility

$$Max\ U_t + \beta \sum_{i=1}^{\infty} \delta^i U_{t+i} = U_t + \beta\delta U_{t+1} + \beta\delta^2 U_{t+2} + ... \tag{1}$$

where $\delta = \frac{1}{(1+r)}$ is the long-run discount factor, r is the discount rate (that coincides with the interest rate in the budget constraints), and the extra discount parameter $\beta \in (0, 1]$ is intended to capture the essence of hyperbolic discounting, namely, that the discount factor between consecutive future periods (δ) is larger than that between the current period and the next ($\beta\delta$). If $\beta \neq 1$, time preferences in equation (1) are dynamically inconsistent, in the sense that the optimal consumption plan established in t is inconsistent with the one that will be established in $t + 1$. As shown by O'Donoghue and Rabin (2002), the equilibrium of both naïve and time consistent individuals solves the same optimization problem. Therefore, the demand equation that solves (1) applies to both time consistent and naïve consumers. As in the general specification of the RA model, the interaction between past and future consumption is modeled by the investment function $A_t = (1-\gamma)A_{t-1}+C_{t-1}$, with A_t being interpreted as an "addiction stock" that depreciates over time at rate $\gamma < 1$ and increases with current consumption. The consumer solves the maximization problem such that $C_0 = C^0$ and $(Y_t + P_t C_t) + \sum_{i=1}^{\infty} \delta^i (Y_{t+i} + P_{t+i} C_{t+i}) = W_0$, where C^0 measures the previous level of consumption.

Taking a quadratic utility function in the three arguments and solving the consumer problem produces the following Euler equation:[3]

$$C_t = \theta_0 + \theta_L C_{t-1} + \theta_F \delta C_{t+1} - \theta_1 \left[1 + (1-\gamma)^2 \delta\right] P_t + \theta_1 (1-\gamma) P_{t-1} + \theta_1 \delta (1-\gamma) P_{t+1} \tag{2}$$

[3] For a detailed derivation of the Euler equation, see Piccoli and Tiezzi (2021).

where:[4]

$$\Omega = \alpha_{CA}(1+\beta)\delta(1-\gamma) - \alpha_{AA}\beta\delta - \alpha_{CC}\left[1+\delta(1-\gamma)^2\right] > 0$$

$$\theta_0 = \frac{\gamma}{\Omega}\left[\kappa + \alpha_C\big(\delta(1-\gamma)-1\big) - \alpha_A\beta\delta\right]$$

$$\theta_L = \frac{1}{\Omega}\left[\alpha_{CA} - \alpha_{CC}(1-\gamma)\right] > 0 \tag{3}$$

$$\theta_F = \frac{1}{\Omega}\left[\beta\alpha_{CA} - \alpha_{CC}(1-\gamma)\right] > 0 \tag{4}$$

$$\theta_1 = \frac{\lambda}{\Omega} > .0$$

The RA model with quasi-hyperbolic discounting encompasses the original formulation with exponential discounting. In particular setting $\beta = 1$ the model reduces to the standard Becker, Grossman, Murphy (1994) RA model, which implies time consistency. Equation (2) can thus be used to test whether consumers are time consistent or not by testing the equality $\theta_L = \theta_F$, as the only difference between the two coefficients is the presence of β in θ_F.

Rewriting equation (2) as a reduced form equation leads to

$$C_{it} = \phi_0 + \phi_L C_{it-1} + \phi_F C_{it+1} + \varphi_T P_{it} + \varphi_L P_{it-1} + \varphi_F P_{it+1}, \tag{5}$$

which, if no parameters restrictions are imposed, allows us to identify all needed structural parameters to perform the standard battery of tests for rational addiction and to test for time consistency. The time consistency test verifies the null[5]

$$\phi_L\varphi_F - \phi_F\varphi_L = 0\ . \tag{6}$$

If the test rejects the null hypothesis, then $\beta \neq 1$ and the data do not support time consistent preferences, as implied by the RA model, in favor of quasi-hyperbolic discounting for naïve consumers.

Given the parametric specification of equation (2) and the corresponding reduced-form equation (5), it is not possible to directly point-identify the value of the present bias parameter β. Nevertheless, it is possible to find an upper bound for β compatible with the estimated coefficients. The

[4]In what follows, α_x are parameters of the quadratic utility function.

[5]Quick proof: since $\delta = \frac{\varphi_F}{\varphi_L}$, $\theta^L = \phi_L$, and $\theta^F = \frac{\phi_F}{\delta} = \frac{\phi_F\varphi_L}{\varphi_F}$, from equations (3) and (4) it is possible to note that $\beta = 1$ if and only if $\theta_L = \theta_F$, i.e., $\phi_L\varphi_F = \phi_F\varphi_L$, or $\phi_L\varphi_F - \phi_F\varphi_L = 0$

upper bound for the present-bias parameter is $\beta_{max} = \frac{\phi_F \varphi_L}{\phi_L \varphi_F}$ (for a proof see Piccoli and Tiezzi, 2021).[6]

3. Empirical strategy

We estimate two empirical demand equations, for young and adult individuals, of the form:

$$C_{it} = \phi_0 + \phi_L C_{it-1} + \phi_F C_{it+1} + \varphi_T P_{it} + \varphi_L P_{it-1} + \varphi_F P_{it+1} + \eta_X X_{it} + v_i + d_t + u_{it} \tag{7}$$

where C_{it} is the number of smoked cigarettes by individual i in period t, P_{it} is cigarettes real price, X_{it} is a vector of exogenous economic and socio-demographic variables that affect cigarettes consumption, v_i are individual fixed effects capturing time invariant preferences that are correlated with lead and lagged consumption and probably with other determinants of consumption, d_t are time fixed effects, and $u_{it} = \xi_1 e_t + \xi_2 e_{t+1}$ is the idiosyncratic error term.

There are two problems that will bias OLS estimates of equation (7). First, OLS estimates can suffer from an omitted variable bias due to unaccounted demand shifters that may also be serially correlated (Becker et al., 1994). Second, there is measurement error when we use actual values of C_{it+1}. C_{it+1} in equation 7 should be interpreted as planned cigarette consumption at time $t+1$ using the information at time t. However, planned and actual consumption at time $t+1$ might differ. Since we use actual cigarette consumption at time $t+1$ to proxy planned consumption, C_{it+1} might be affected by measurement error which will enter the idiosyncratic error term (Picone, 2005). The disturbance term in equation 7 would be:

$$u_{it} = \xi_1 e_{it} + \xi_2 e_{it+1} + \beta\delta r_{it+1} \tag{8}$$

where $r_{it+1} = C^*_{it+1} - C_{it+1}$ is the difference between planned and actual cigarette consumption. There are two consequences of this measurement error. First, r_{it+1} is correlated with P_{it+1} and P_{it}, so both P_{it+1} and P_{it} should be treated as endogenous. Second, $E(P_{it-1}u_{it}) = 0$, which is the assumption that P_{it-1} is a predetermined random variable.

The standard route to correct for the endogeneity bias is to follow Arellano and Bond (1991) in using a GMM procedure to obtain the vector of parameters. The idea is to take first-differences to deal with the unobserved fixed effects and then use the suitably lagged levels of the endogenous

[6] Extensive simulations of the structural model for a wide range of plausible values of the parameters suggest that the true value of the β parameter is always very close to its upper bound β_{max}.

and predetermined variables as instruments for the first-differenced series, under the assumption that the error term in levels is spherical and taking into account the serial correlation induced by the first-difference transformation. This idea extends to the case of lags and leads of the dependent variable and to the case where serial correlation already exists in the error term of the original model, as in equation (7).

We need a set of instruments Z_{it} that are uncorrelated with the first-differenced error term Δu_{it} and correlated with the regressors. By definition

$$\Delta u_{it} = \xi_1 \Delta e_{it} + \xi_2 \Delta e_{it+1} + \beta\delta\Delta r_{it+1} \tag{9}$$

for $i = 1, ..., N$ and $t = 3, ..., T-1$. Given (9), the following moment conditions are available: $E(C_{it-s}\Delta u_{it}) = 0$ for $t = 4, ..., T-1$ and $s \geq 3$. This allows the use of lagged levels of the observed consumption series dated $t-3$ and earlier as instruments for the first-differenced equation (10):

$$\Delta C_{it} = \phi_1 \Delta C_{it-1} + \phi_2 \Delta C_{it+1} + \phi_3 \Delta P_{it} + \phi_4 \Delta P_{it-1} + \phi_5 \Delta P_{it+1} + \phi_6 \Delta X_{it} + \phi_7 \Delta d_t + \Delta u_{it} \tag{10}$$

The moment restrictions can be written in matrix form as $E(Z_i' \Delta u_i) = 0$ for $t = 4, ..., T-1$, where Δu_i is the $(T-4)$ vector $(\Delta u_{i4}, \Delta u_{i5}, ..., \Delta u_{iT-1})'$. $\Delta u_i = u_{it} - u_{it-1}$ and Z_i is a $(T-4) \times g$ block diagonal matrix, whose i^{th} block is:

$$Z_i = \begin{pmatrix} C_{i1}, P_{i1}, P_{i2} & 0 & 0 & \dots & 0 & \dots & 0 & \Delta W_{i4}' \\ 0 & C_{i1}, P_{i1}, P_{i2} & C_{i2}, P_{i2}, P_{i3} & \dots & 0 & \dots & 0 & \Delta W_{i5}' \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{i1}, P_{i1}, P_{i2} & \dots & C_{iT-4}, P_{iT-4}, P_{iT-3} & \Delta W_{iT-1}' \end{pmatrix}.$$

where the block diagonal structure at each time period exploits all of the instruments available, concatenated to one-column of first differenced exogenous regressors $\Delta W_{it}' = (\Delta X_{it})$ that act as instruments for themselves (Arellano and Bond, 1991).

The first-differenced GMM estimator is known to be poorly behaved in terms of finite sample properties (bias and imprecision) when instruments are weak. This can occur here given that the lagged levels of consumption are usually only weakly correlated with subsequent first-differences. Better finite sample properties can be obtained with system-GMM (Arellano and Bover, 1995; Blundell and Bond, 1998). This estimator exploits additional moment conditions, which are valid under the "constant correlated effects"assumption (Bun and Sarafidis, 2013). This yields $(T-4)$ further linear moment conditions, $E(\Delta C_{it-2} u_{it}) = 0$ for $t = 4, ..., T-1$, which allow the use of equations in levels with suitably lagged first-differences of the series as instruments. The complete system of moment conditions available can be expressed as $E(Z_i^{+'} u_i^+) = 0$, where $u_i^+ = (\Delta u_{i4}, ..., \Delta u_{iT-1}, u_{i4}, ..., u_{iT-1})'$.

The instrument matrix, Z_i^+, for this system is

$$Z_i^+ = \begin{pmatrix} Z_i & 0 & 0 & \dots & 0 \\ 0 & \Delta C_{i2}, \Delta P_{i3} & 0 & \dots & 0 \\ 0 & 0 & \Delta C_{i3}, \Delta P_{i4} & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & \Delta C_{iT-3}, \Delta P_{iT-2} \end{pmatrix}$$

where the first row is the set of valid instrument for the equation at time $t = 4$.

In finite samples, such a large instruments collection generates a bias/efficiency trade-off (Biørn and Klette, 1998; Roodman, 2009b; Ziliak, 1997). For this reason, after experimenting with the instruments, we use two parsimonious matrices of instruments, Z_i^Y and Z_i^A, for the young and adults equations, respectively, where we use only a subset of the available instruments, discarding the use of some valid ones either by collapsing or curtailing the instrument matrix (Kiviet, 2020).

As to the choice of the transformation used to remove individual effects in GMM estimators, while first differencing (FD) is one option, Arellano and Bover (1995) propose forward orthogonal deviations (FOD) as an alternative transformation for models with predetermined instruments, involving subtracting the mean of all future observations for each individual. The FOD transformation does not introduce a moving average process in the disturbance, i.e. orthogonality among errors is maintained, and preserves the sample size in panels with gaps, as in our case, where FD would reduce the number of observations (Roodman, 2009a). Valid moment conditions for the FOD model in presence of endogenous regressors (Kripfganz, 2019) are: $E(C_{it-s} \Delta u_{it}) = 0$ for $t = 4, ..., T-1$ and $s \geq 2$. This allows the use of lagged levels of the observed consumption series dated $t-2$ and earlier as instruments for the FOD equation.

3.1. Data

The estimation of the general specification of the RA model for cigarettes consumption requires sufficiently long individual longitudinal data on the number of cigarettes smoked along with prices at local level. This requirement is not easy to fulfill. For instance most US longitudinal data, including PSID and NLSY97, only include repeated information on whether the individual smokes or not, not the number of cigarettes, and this hampers the estimation of a RA model. In addition, the longer individuals are followed, the more information is available for the estimation of the dynamics of smoking habits, thus leaving out rotational panel from the pool of candidates. To the best of our knowledge, the longest longitudinal survey that collects individual information on the number of cigarettes smoked along with local level prices and that follows individuals for as long as possible is the Russia Longitudinal Monitoring Survey (RLMS-HSE), which started in 1994 and is till ongoing. The survey is conducted by the Higher School of Economics and ZAO Demoscop, together with the Carolina Population Center, and follows individuals and their families from childhood to adulthood.[7] Households participating in the survey were selected through a multistage probability sampling procedure to guarantee cross-sectional national representativeness. Within each of the 38 primary sample units (PSUs), the population was stratified into urban and rural substrata to guarantee the representativeness of the sample in both areas. The survey covers approximately 5,000 hh, 12,000 adults and 2,000 children (aged up to 15 years) per wave.

The empirical analysis of cigarette addiction is thus based on waves 5 to 29 (corresponding to years 1994 to 2020) of the RLMS-HSE. For each individual aged 13 years and above, the survey asks whether she/he smokes and, if so, the number of cigarettes smoked per day. This is the main consumption measure used in our study. The price variable is computed from the community questionnaire, where interviewers go to local stores in the community and check minimum and maximum prices of a large sample of commodities, including domestic- and foreign-branded cigarettes. Because several missing values are recorded at the community level (if, for instance, no store had a particular item or if the store was closed), the price was averaged across communities within the same primary sample units to reduce the impact of measurement errors. Because the prices are at the current level, and the survey does not provide consumer price indices to deflate prices, we compute a consumer price index at the PSU level following the Törnqvist procedure (Törnqvist, 1936). The reference price is that of the Moscow PSU in 1998, and the index is computed on a wide set of food commodities, excluding tobacco and alcohol items. Cigarette prices are then deflated using this consumer price index.

In order to be able to observe young smokers still living with their parents, but avoid complex multi-family households, we selected smokers living in households with up to 4 components. the Young's equation is estimated for individuals no older than 21, for which about two thousand ob-

[7] More information can be found in the RLMS-HSE site: http://www.cpc.unc.edu/projects/rlms-hse.

servations and about one thousand individuals are available. The Adults' equation is estimated for individuals aged 30 to 65, corresponding to about 31 thousand observations and 6.4 thousand individuals. We use identical specifications for both equations with a minimal set of strictly exogenous covariates, X_{it}, which includes: waves' indicators, gender of the respondent ($female = 1$), age (in years) of the respondent and its square. Time dummies have been specified in the level model only to avoid redundancy (Kripfganz, 2019).[8]

4. Results

Table 1 reports estimates of model (7) for young and adult individuals. The table also reports the instruments count and the p-value of the Hansen test for the joint validity of all instruments (Hansen p-value), together with the Arellano-Bond test for second and third-order serial correlations in the residuals.[9]

Estimates are consistent with the RA framework for both young and adult individuals. In both equations past consumption has a significant positive effect. Future consumption also has a significant positive effect, supporting the idea that smokers' behavior is forward looking. The coefficient of lagged price is greater than the coefficient of lead price, determining a positive discount rate. We obtain a negative coefficient on the current price and a positive coefficient on both past and future prices. So, the signs on the two consumption variables and on the three price variables conform to theoretical predictions. The p-values of the Hansen J statistic for over-identifying restrictions for the full model are consistent with the null hypothesis of no-overidentification. Finally, the Arellano-Bond test for third-order autocorrelation in the residuals does not detect third-order serial correlation in the residuals of either of the two equations.

The conditions necessary for stability of the second-order difference equation in current consumption include that the sum of the coefficients on past and future consumption is less than unity (Chaloupka, 1990). In both equations, the sum of coefficients on past and future consumption is less than unity (0.523 for young individuals; 0.869 for Adults).

As to the covariates, being female has a negative and statistically significant impact on the number of cigarettes smoked. Age and smoked cigarettes display a non linear relationship in both

[8]For estimations we used the xtabond2 command in STATA 17.

[9]Because in our model current consumption depends on both past and future consumption, this is an autoregressive process of order 2 (AR2) and we have second-order serial correlation by construction. So, for the validity of our instrument set, we need to detect no serial correlation of order 3 in the residuals.

Table 1: General Rational Addiction Model Estimates: Young & Adult smokers

Variables	**Young** ($age \leq 21$)	**Adult** ($30 \leq age \leq 65$)
C_{t-1}	0.339***	0.442***
	(0.079)	(0.076)
C_{t+1}	0.183***	0.427***
	(0.071)	(0.088)
P_t	-0.060**	-0.033***
	(0.029)	(0.013)
P_{t-1}	0.046*	0.017**
	(0.026)	(0.007)
P_{t+1}	0.025**	0.016*
	(0.010)	(0.009)
Gender	-1.220**	-0.780**
	(0.472)	(0.314)
Age	0.683**	0.059**
	(0.311)	(0.027)
Age^2	-0.012**	-0.001**
	(0.006)	(0.000)
Time dummies	Yes	Yes
Hansen p-value full	0.100	0.537
p-value Arellano-Bond test for AR(2)	0.001	0.000
p-value Arellano-Bond test for AR(3)	0.146	0.358
# Obs	2,037	31,189
Instruments count	172	136

Notes: Robust SE in parentheses using Windmeijer correction. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

equations: one additional year of age is associated with an increase in the number of cigarettes smoked per day but at a decreasing rate. This non-linear relationship is stronger for young smokers for whom we obtain larger coefficients (in absolute value) for both Age and Age squared signalling that, compared to adult individuals, one additional year of age is associated with a larger increase in the number of cigarettes smoked per day and with a larger decrease in marginal consumption.

Table 2 shows the results of the time consistency tests and time preferences parameters computed from the estimated parameters. The test consists of testing the null hypothesis $\phi_1\phi_5 - \phi_2\phi_4 = 0$. For the non-linear test, under the null, the test statistics has a χ^2 distribution with 1 degree of

Table 2: Time consistency tests

	Young (age $\leq$ 21)	**Adult** (30 $\leq$ age $\leq$ 65)
β_{max}	0.992	0.997
Time -consistency nl-test $\chi^2(1)$ (p-val)	0.00 (0.993)	0.00 (0.997)
Discount factor	0.543	0.969
Discount rate	0.842	0.032

freedom. We obtain a test statistics of $\chi^2(1) = 0.00$ with a $Prob > \chi^2 = 0.992$ for young individuals, and a test statistics of $\chi^2(1) = 0.00$ with a $Prob > \chi^2 = 0.996$ for adult individuals. This means that the null of time consistency cannot be rejected for either group. The estimated upper bound for the present bias parameter is $\beta_{max} = \frac{\widehat{\theta}^+}{\widehat{\theta}^-} = \frac{\phi_2\phi_4}{\phi_1\phi_5} = 0.992$ for young individuals and 0.997 for adult individuals. However, we cannot reject the null hypothesis $\beta_{max} = 1$ for either sample.

As to the estimated discount factor and discount rate, for young individuals we estimate a discount factor $\delta = \phi_5/\phi_4 = 0.543$ and a discount rate of 0.842, while for the adults we estimate a discount factor of $\delta = 0.969$ and a discount rate of 0.032.

One possible issue in estimating the general version of the RA model when assuming $\beta = 1$, as supported by our results, is overparametrization. The problem is that the same structural parameters would be derived by different empirical parameters. More specifically, the discount factor would be obtained by the ratios $\frac{\phi_F}{\phi_L}$ and $\frac{\varphi_F}{\varphi_L}$, with no guarantee that the result is identical, and generally will not. One possibility, would be to constrain one of the parameters, although this is not easily achievable using the available statistical programs and commands. Most of the empirical literature, instead, facing difficulties in properly estimating the general version of the RA model, resorted to the restricted version. The restricted version excludes from the empirical specification lead and forward prices. From a theoretical perspective, this corresponds to assume complete depreciation of the addiction stock in each time period, i.e. $\gamma = 1$ in the addiction motion equation and in equation (2). This may be reasonably considered implausible, knowing how persistent addiction can be in the real word. However, this specification still allows to recover the discount factor as $\frac{\phi_F}{\phi_L}$, avoids overparametrization, and, considering that our time period is one year, it may not be completely out of logic to assume that in one time period addiction fully

Table 3: Restricted model estimation

	Young (age $\leq$ 21)	**Adult** (30 $\leq$ age $\leq$ 65)
C_{t-1}	0.361***	0.497***
	(0.074)	(0.057)
C_{t+1}	0.194***	0.492***
	(0.076)	(0.063)
P_t	-0.044*	-0.010**
	(0.023)	(0.0048)
Discount factor	0.538	0.988
Discount rate	0.859	0.012

Notes: Robust SE in parentheses using Windmeijer correction. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

depreciates.[10]

Thus, to check the robustness of our results we also estimate the restricted version of the RA model on the samples of young and adult smokers with the same set of covariates. The results, presented in Table 3 are indeed very close to those obtained for the general specification. In particular, with the restricted specification, the discount factor for young smokers is just 1% smaller than with the general specification, and for the adults is 2% larger, confirming our main result.

Summarizing, the estimation results suggest that both young and adults Russian are rationally addicted smokers with non-present-biased time preferences, but young individuals discount future utilities much more than adults. Observing such a large difference in time preferences at different ages rises some doubts on the validity of the quasi-hyperbolic discounting model, which assumes that both the β and the δ are constant over the life cycle, as discussed later.

There are very few works to which we may compare our results, because estimates of quasi-hyperbolic preferences using observational data have been achieved only very recently in the literature. In a previous article, Piccoli and Tiezzi (2021) found very similar results in terms of present-bias and discount factor using the same data but with a different sample selection and reference period. For a sample of adults aged 22-74 and years 2006-2018, with a slightly different

[10]This would mean that a smoker that stops smoking today and never smokes for one year could be considered addiction free in the next period.

empirical specification and covariates set, they still reject present bias, with $\beta_{max} = 0.99$, and find a discount factor of 0.988, remarkably close to the 0.969 found in the present study. No estimates were proposed for young smokers, though.

More interesting is the comparison with Blow et al. (2021). Although their setting does not account for addiction in consumption, and the empirical test is very different, as they apply a revealed preference approach, for a sample of Spanish households they find a remarkably close discount factor (0.957). Where they differ substantially is about present bias. They find very poor support for time consistent exponential discounting, with only 2% of households in their sample passing the test, but much better support for quasi-hyperbolic discounting, with a 45% pass-rate. Their average present bias parameter β is 0.836, and a relevant proportion of the sample shows even larger present bias: about 10% of families show a present-bias parameter of 0.729 or less. While 45% might not seem a stellar number, for sure quasi-hyperbolic discounting performs much better than exponential discounting in their sample.

One possible explanation may reside on the reference period. O'Donoghue and Rabin (2015) pointed out how present bias (i.e. $\beta < 1$) is all about noticeable short-term discounting, such as daily discounting, that is to say it is about, e.g., comparing utility now and in one day from now, versus comparing utilities in two adjacent future days. Instead, here we estimate yearly discount factors, i.e. how much individuals care today about utility in one year. In this long-run perspective, finding no evidence of present bias seems plausible. As to Blow et al. (2021), they use quarterly data, i.e. their data frequency is 4 times faster than ours, and this might be sufficient to let some present-bias emerge.

5. Discussion

Time inconsistency of young individuals is one of the motivations behind the adoption of Tobacco 21 Law in recent years in the US, which bans tobacco sales to individuals younger than 21. In addition, as Gruber and Köszegi (2002, 2004) point out, when agents are time inconsistent, positive taxation is optimal even in the absence of externalities, as time inconsistency will imply self-control problems and the optimal future consumption path planned at time t will not be realized by the agent, because not optimal anymore in $t + 1$ and onward. Hence, in the case of time inconsistent agents, taxes on addictive goods are substantially larger than those for time consistent consumers.

In the present paper we test whether young and adult smokers have different time preference, in particular with respect to time consistency, using individual cigarettes consumption longitudinal data. We reject present-bias for both groups and find that they display a rational addictive behavior, in the sense that they are both addicted and forward looking, although young smokers discount future utilities much more than the adults.

While we reject present-bias as a form of time inconsistency, the much larger discounting of future utilities of the young suggests a more complex form of life-cycle time inconsistency –that cannot be properly accounted for using quasi-hyperbolic discounting– where young individuals are much more present-oriented than the adults. This has relevant implications both from a policy and research perspective.

These results are consistent with Crettez and Deloche (2021) suggesting that young individuals discount the future more heavily than the adult population. Adults display a much smaller discount rate and a much larger discount factor, i.e. they give future actions and future utility a much larger weight than young individuals. This gives empirical support to one of the motivations behind the Tobacco 21 law even using large scale consumption survey data (in addition to the evidence based on experimental methods). Clearly, this also supports larger taxes on addictive goods due to the unforeseen consequence of addiction on future selves. This is despite the fact that the empirical test of present-biased preferences is rejected, and more complex forms of time inconsistency could not be formally tested.

Further extension would be needed to model time inconsistent preferences in the presence of addiction in a life-cycle perspective. The economic literature has generally proposed to treat the discount factor as endogenous and dependent on some choice or state variables determined by the model. The first proposal dates back to Becker and Mulligan (1997) who consider an endogenous discount factor that depends on some endogenous investment/effort choice from the individual. The idea is that the individual seeks to become more forward looking but to do so she must make some effort and invest some time or resources in that direction. While this hypothesis may well be supported by the evidence produced in our analysis, it is empirically difficult to implement. In fact, it is based on unobservable choice variables that would be quite difficult to proxy. To the best of our knowledge no empirical results from these types of models using survey data have been produced by the literature yet.

An interesting alternative is based on recent work by Strulik (2018), in which smoking accelerates health deficit accumulation and reduces the survival probability, which in turns affects time preferences through the discount factor. While such proposal may work very well for explaining the long term impact of addiction on time preferences, it seems less useful to explain why young individuals display such a smaller discount factor compared to adults. Besides, it is also based on unobserved state variables that may limits its empirical applicability. In fact, if survival probability may well be proxied by health status and other socio-economic variables, the stock of addiction is an artificial construct that is difficult to proxy using observational data.

An alternative specification recognizes that the addictive stock itself may impact individuals' time preferences (Shi and Epstein, 1993). The idea is that people who are severely addicted could discount future utilities much more, as they "need" to consume now. Hence, a larger stock of addiction would reduce the discount factor. While high degree of addiction of young consumers and oscillatory behavior of consumption are compatible with the model predictions (Perali and Piccoli, 2022), the existence of a unique steady state suggests that these kind of models may be unable to properly account for the much lower discount factors of young smokers. Interestingly, however, empirical applications of such models start to emerge in the literature. In particular, in a recent working paper, Hai and Heckman (2022) estimate a model where the discount factor depends on the stock of addiction using a simulated GMM estimator. This way of estimating more complex models with inconsistent time preferences might well be one way forward.

Finally, we foresee that a more pragmatic modeling of inconsistent time preference might be possible. Given that age seems to be one of the most relevant factor in explaining the discount rate,[11] and it is readily available information in most surveys that also collect information on consumption of addictive goods, the discount factor might be modeled directly as a function of age. A naïve consumer knows that she has a certain discount factor but she ignores that in the future it will change with age. This would produce time inconsistent consumption plans similar to the naïve present-biased preferences, but would account for a possibly non-linear relationship of the discount factor with age. The properties of such a model, the existence of a steady state, its stability and its empirical tractability will be subject of future research.

[11] Other factors, closely related with age, might be important as well. For instance, getting married, having children, or retire from work may be important life cycle events that may affect time preference.

References

Arellano, M., Bond, S., 1991. Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. Review of Economic Studies 58, 277–297.

Arellano, M., Bover, O., 1995. Another look at the Instrumental Variable Estimation of error-component Models. Journal of Econometrics 68, 29–51.

Becker, G., Grossman, M., Murphy, K., 1994. An empirical analysis of cigarette addiction. The American Economic Review , 396–418.

Becker, G., Murphy, K., 1988. A theory of rational addiction. Journal of Political Economy 96, 675–701.

Becker, G.S., Mulligan, C.B., 1997. The endogenous determination of time preference 112, 729–758.

Biørn, E., Klette, T., 1998. Panel data with errors-in-variables: essential and redundant orthogonality conditions in gmm-estimation. Economics Letters 59(3), 275–282.

Blow, L., Browning, M., Crawford, I., 2021. Non-parametric analysis of time-inconsistent preferences. Review of Economic Studies 88, 2687–2734.

Blundell, R., Bond, S., 1998. Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics 87, 115–143.

Bun, M., Sarafidis, V., 2013. Dynamic panel data models. UvA-Econometrics Discussion Paper - University of Amsterdam 01.

Chaloupka, F., 1990. Rational Addictive Behaviour and Cigarette Smoking. NBER Working Paper 3268.

Crettez, B., Deloche, R., 2021. Time-inconsistent preferences and the minimum legal tobacco consuming age. Rationality and Society 33, 176–195.

Grossman, M., 2022. The demand for health turns 50: Reflections. Health Economics 31, 1807–1822.

Gruber, J., Köszegi, B., 2001. Is Addiction Rational? Theory and Evidence. Quarterly Journal of Economics CXVI, 1261–1303.

Gruber, J., Köszegi, B., 2002. A theory of government regulation of addictive bads: optimal tax levels and tax incidence for cigarette excise taxation. NBER Workink Paper 8777.

Gruber, J., Köszegi, B., 2004. Tax incidence when individuals are time-inconsistent: the case of cigarette excise taxes. Journal of Public Economics 88, 1959–1987.

Hai, R., Heckman, J.J., 2022. The causal effects of youth cigarette addiction and education. NBER Working Paper .

Hansen, B., McNichols, D., Sabia, J., Bryan, C., 2022. Do state tobacco 21 laws work? Available at SSRN: https://ssrn.com/abstract=4182293.

Kiviet, I.F., 2020. Microeconometric dynamic panel data methods: Model specification and selection issues. Econometrics and Statistics 13, 16–45.

Kripfganz, S., 2019. Generalized method of moments estimation of linear dynamic panel data models. Proceedings of the 2019 London Stata Conference.

Laibson, D., 1997. Golden eggs and hyperbolic discounting. Quarterly Journal of Economics 62, 443–478.

O'Donoghue, T., Rabin, M., 2006. Optimal sin taxes. Journal of Public Economics 90, 1825–1849.

O'Donoghue, T., Rabin, M., 2002. Addiction and Present Biased Preferences. Working Paper Department of Economics, University of California at Berkeley E02-312.

O'Donoghue, T., Rabin, M., 2015. Present bias: Lessons learned and to be learned. The American Economic Review 105, 273–279.

Perali, F., Piccoli, L., 2022. An extended theory of rational addiction. Mathematics 10, 2652.

Piccoli, L., Tiezzi, S., 2021. Rational addiction and time-consistency: An empirical test. Journal of Health Economics 80, 102546.

Picone, G., 2005. Gmm estimators, insturments, and the economics of addiction. Unpublished Manuscript, University of South Florida.

Pollak, R., 1968. Consistent planning. Review of Economic Studies 35(2), 201–208.

Roodman, D., 2009a. How to do xtabond2: an introduction to difference and system gmm in stata. The Stata Journal 9, 86–136.

Roodman, D., 2009b. A note on the theme of too many instruments. Oxford Bullettin of Economics and Statistics 71, 135–158.

Shi, S., Epstein, L.G., 1993. Habits and Time Preference. International Economic Review 34, 61–84.

Strotz, R., 1956. Myopia and inconsistency in dynamic utility maximization. The Review of Economic Studies 23, 165–180.

Strulik, H., 2018. Smoking kills: An economic theory of addiction, health deficit accumulation, and longevity. Journal of Health Economics 62, 1–12.

Törnqvist, L., 1936. The Bank of Finland's consumption price index .

Ziliak, J., 1997. Why do people drink? tests of the rational addiction model. Journal of Business and Economics Statistics 15, 419–431.