Theoretical considerations on the retirement consumption puzzle and the optimal age of retirement

Nicolas Drouhin*

April 13, 2018

Abstract

Defining retirement as a discontinuity in the labor supply of the agent, this paper resolves the retirement consumption puzzle in a very general model of intertemporal choice of consumption and savings of a fully rational, forward looking, agent. Building on a specific version of Bellman (1957) principle of optimality, it provides a very general and parsimonious formula for determining the optimal age of retirement taking into account the possible discontinuity of the optimal consumption profile at the age of retirement.

Code JEL: C61 D91 J26

Key words : life cycle theory of consumption and saving; optimal retirement, retirement consumption puzzle, discontinuous optimal control.

*Ecole normale supérieure Paris-Saclay and CREST (UMR CNRS 9194)

1 Introduction

In this article, I build a model that address at the same time the retirement consumption puzzle and and the optimal age of retirement.

Since Hamermesh (1984a) many empirical studies document a drop in consumption at retirement, the *retirement consumption puzzle* (Banks et al., 1998; Bernheim et al., 2001; Battistin et al., 2009, among others). This phenomena is seen as puzzling and "paradoxical" because it seems in contradiction with the idea that, within the intertemporal choice model, which is the backbone of modern economics, when preferences are convex, consumption smoothing is the rule. Then, explanation of this paradox has been searched in relaxing some assumptions of the model of fully rational forward looking agent. For example the agent may systematically underestimate the drop in earnings associated with retirement(Hamermesh, 1984a). Or, the agent may not be fully time consistent as in the hyperbolic discounting model (Angeletos et al., 2001).

Without denying that those phenomena may be important traits of "real" agents behavior, building on an insight of Banks et al. (1998) in their conclusion, this paper will emphasize the point that a closer look at the intertemporal choice model of consumption and savings in continuous time allows to understand that what is smooth in the model is not necessarily consumption, but marginal utility of consumption. Of course, if consumption is the only variable of the utility function the two properties are equivalent. But if utility is multi-variate, any discontinuity in a dimension, may imply an optimal discontinuity response in the others. I will illustrate that insight into a very general model of inter-temporal choice that can be considered as a realistic generalisation of the basic one. Two ingredients will be required. First, I will assume a bi-variate, additively intertemporaly separable utility function that depends on consumption and leisure. Second weI will assume, realistically, that retirement is not a smooth process with a per period duration of labor that tend progressively to zero, but a discontinuous process.

I will show that, as long as the per period utility function is not additively separable in consumption and leisure, discontinuity of the consumption function is the rule in this general model. However, as insightful is the preceding statement, it is not so easy to prove it formally, with all generality because our assumptions imply a discontinuous payoff function, a case that is not standard with usual intertemporal optimization techniques in continuous time. I will provide a general and simple lemma that will make the problem tractable and it's resolution at the same time rigourous and insightful.

So, if we want to solve the paradox within a quite standard model of intertemporal choice, we have to drop additive separability of utility of consumption and leisure. And if we want to extend the problem to the choice of the optimal retirement age, we have to carry on with this non-separability. However, as pointed by d'Albis et al. (2012) most of the study addressing the question has been made precisely under the assumption of additive separability in consumption and leisure (see d'Albis and Augeraud-Véron, 2008; Bloom et al., 2014; Boucekkine et al., 2002; Hazan, 2009; Heijdra and Mierau, 2012; Heijdra and Romp, 2009; Kalemli-Ozcan and Weil, 2010; Prettner and Canning, 2014; Sheshinski, 1978, among others). And if there are some important papers that study a general life cycle model of consumption and savings, without additive separability of consumption and leisure (Heckman, 1974, 1976; Bütler, 2001) they are mostly focused on the the explanation of co-movement of earnings and consumption all over the life-cycle. Hamermesh (1984b) and Chang (1991) study the retirement decision with non separability of consumption and leisure, but they fully endogenize the work decision, without any granularity concerning the per-period duration of worktime, and thus without any discontinuity of per period labor supply, implying model that are unable to explain at the same time retirement consumption paradox and the retirement decision. The model I propose can easily be expanded to endogenize the retirement decision and provide very general condition that fulfills the optimal age of retirement. I will show that when optimal consumption is discontinuous at the age of retirement, this condition is qualitatively very different than in the traditional case.

2 A general Life Cycle Model solving the retirement consumption paradox

Let's assume that we are in a very standard continuous time life-cycle model of consumption and savings with preference for leisure and retirement.

$$\mathcal{P} \quad \begin{cases} \max\limits_{c} \int_t^T e^{-\theta(s-t)} u\left(c\left(s\right), l(s)\right) \mathrm{d}s \\ s.t. \forall s \in [t,T], \quad \dot{a}(s) = ra(s) + w(s)(1-l(s)) + b(s) - c(s) \\ a(t) \text{ given and } a(T) \geq 0 \end{cases}$$

t is the decision date and T is life duration, u is the per-period bi-variate utility function that depends on consumption and leisure. c, is the intertemporal consumption profile, the control variable of the program, l, is the intertemporal leisure profile, that I will assume, in a first stance, to be exogenous. a, is a life-cycle asset, the state variable of the program, that brings interest at the rate r. w, is labor income per period when the individual spend all this time working. b is social security income profile, interpreted as social security benefit when positive (typically after retirement) and social security contribution when negative (typically before retirement). l, c, w and b are assumed to be piecewise continuous and a is assumed to be piecewise smooth, assumptions that are fully compatible with the use of standard optimal control theory.

I assume that the utility function includes standard minimum requirements of the microeconomic theory of consumption/leisure trade-off: $u_1 > 0$, $u_2 > 0$, $u_{11} < 0$, $u_{22} < 0$ and quasi-concavity (i.e. the indifference curves are convex). It implies that $-u_{11}u_2^2 + 2u_{12}u_1u_2 - u_{22}u_1^2 > 0$. It is important to notice that, without further assumptions, the sign of the second order crossed derivative is undetermined.

I will assume that there exists a retirement age t_R such that:

$$\begin{cases} \forall s \in [t, t_R), & l(s) = \kappa < 1 \\ \forall s \in [t_R, T], & l(s) = 1 \end{cases}$$

Of course this assumption is a simplification, but it allows to characterize directly the central idea of the paper: retirement is fundamentally a discontinuity in the labor/leisure profile. This assumption seems much more realistic than usual idea that retirement is the smooth process with per period work duration tending to zero at the age of retirement.[1]

I denote c^* the optimal consumption profile, solution of the program $\mathcal{P}$ and a^* the associated value of the state variable. Of course those optimal functions are parameterized by all the given of the problem $(t, t_R, T, a(t), r, w, b, l)$.

I denote V^* the optimal value of the problem i. e.

$$V^*(t, t_R, T, a(t), r, w) = \int_t^T e^{-\theta(s-t)} u\left(c^*(t, t_R, T, a(t), r, w, b, l, s), l(s)\right) \mathrm{d}s$$

Because of the discontinuity of the instantaneous payoff function in t_R, the problem is non standard. Therefore it is useful to decompose the problem in two separate ones:

$\mathcal{P}^0$

$$\begin{cases} \max\limits_{c} \int_t^{t_R} e^{-\theta(s-t)} u\left(c\left(s\right), \kappa\right) \mathrm{d}s \\ s.t. \\ \dot{a}(s) = ra(s) + (1-\kappa)w(s) + b(s) - c(s) \\ a(t), a(t_R) \text{ given} \end{cases}$$

$\mathcal{P}^1$

$$\begin{cases} \max\limits_{c} \int_{t_R}^{T} e^{-\theta(s-t)} u\left(c\left(s\right), 1\right) \mathrm{d}s \\ s.t. \\ \dot{a}(s) = ra(s) + b(s) - c(s) \\ a(t_R) \text{ given and } a(T) \geq 0 \end{cases}$$

[1]This idea could be generalized by endogenizing per period work duration taking into account a granularity assumption. In general for organizational reason work duration can be zero or something significantly different from zero.

I denote c^0 and c^1 the optimal consumption profile, respective solution of programs $\mathcal{P}^0$ and $\mathcal{P}^1$. As c^* they are also implicit functions of the parameter of their respective program and I can define the value function of $\mathcal{P}^0$ and $\mathcal{P}^1$.

$$V^0(t, t_R, a(t), a(t_R), r, w, b, \kappa) = \int_t^{t_R} e^{-\theta(s-t)} u\left(c^0(t, t_R, a(t), a(t_R), r, w, b, \kappa), \kappa\right) \mathrm{d}s$$

$$V^1(t_R, T, a(t_R), r, b) = \int_{t_R}^{T} e^{-\theta(s-t)} u\left(c^1(t_R, T, a(t_R), r, b, s), 1\right) \mathrm{d}s$$

The two programs are linked by the asset level at the age of retirement. By application of the optimality principle, I can deduce:

Lemma 1 (A Principle of Optimality)**.**
If (c^*, a^*) is an admissible pair solution of program $\mathcal{P}$ then we have:

1. $V^*(t, t_R, T, a(t), r, w) = V^{0*}(t, t_R, a(t), a^*(t_R), r, w) + V^{1*}(t_R, T, a^*(t_R), r, w)$

2. $a^*(t_R) = \underset{a(t_R)}{\operatorname{argmax}}\{V^0(t, t_R, a(t), a(t_R), r, w) + V^1(t_R, T, a(t_R), r, w)\}$

Proof: It is a direct application of Bellman (1957) principle of optimality. □

I have now all the material to solve the program $\mathcal{P}$.

Proposition 1 (Discontinuity of the consumption profile)**.**
If I denote $c^0(t_R) \stackrel{def}{=} \lim_{s \to t_R} c^0(s)$, and restrict my analysis to per period utility with a second order cross derivative that is either, everywhere strictly positive, everywhere strictly negative or everywhere equal to zero:

1. *The optimal consumption profile solution of program $\mathcal{P}$ is unique.*

2. *The optimal consumption profile solution of program $\mathcal{P}$ is continuous for every age s in $[t, t_R) \bigcup (t_R, T]$.*

3. *In t_R, $u_1(c^0(t_R), \kappa) = u_1(c^1(t_R), 1)$ and the continuity of the optimal consumption profile is determined solely the cross derivative of the per period utility function.*

(a) $c^0(t_R) > c^1(t_R) \Leftrightarrow u_{12}(c,l) < 0$

(b) $c^0(t_R) = c^1(t_R) \Leftrightarrow u_{12}(c,l) = 0$

(c) $c^0(t_R) < c^1(t_R) \Leftrightarrow u_{12}(c,l) > 0$

Proof: Relying on Lemma 1, I start by solving the program $\mathcal{P}^0$ and $\mathcal{P}^1$ for a given $a(t_R)$. Denoting μ^0 the costate variable, the Hamiltonian of the Program $\mathcal{P}^0$ is:

$$\mathcal{H}^0(c(s), a(s), \mu^0(s), s) = e^{-\theta(s-t)} u\left(c\left(s\right), \kappa\right) + \mu^0(s)\left[r\, a(s) + (1-\kappa) w(s) + b(s) - c(s)\right] \tag{1}$$

According to Pontryagin maximum principle the necessary condition for optimality is:

$$\forall s \in [t, t_R), \frac{\partial \mathcal{H}^0(\bullet)}{\partial c(s)} = 0 \Rightarrow \mu^0(s) = e^{-\theta(s-t)} u_1\left(c\left(s\right), \kappa\right) \tag{2}$$

$$\forall s \in [t, t_R), \frac{\partial \mathcal{H}^0(\bullet)}{\partial a(s)} = -\dot{\mu}^0(s) \Rightarrow \dot{\mu}^0(s) = -r\, \mu^0(s) \tag{3}$$

$$\forall s \in [t, t_R),\ \dot{a}(s) = ra(s) + (1-\kappa)w(s) + b(s) - c(s) \tag{4}$$

Moreover by construction of the Hamiltonian and *Pontryagin maximum principle* it is well known that:

$$\frac{\partial V^0(t, t_R, a(t), a(t_R), r, w, b, \kappa)}{\partial a(t_R)} = -\mu^0(t_R) \tag{5}$$

Similarly for program $\mathcal{P}^1$, we have:

$$\mathcal{H}^1(c(s), a(s), \mu^1(s), s) = e^{-\theta(s-t)} u\left(c\left(s\right), 1\right) + \mu^1(s)\left[r\, a(s) + b(s) - c(s)\right] \tag{6}$$

$$\forall s \in (t_R, T], \frac{\partial \mathcal{H}^1(\bullet)}{\partial c(s)} = 0 \Rightarrow \mu^1(s) = e^{-\theta(s-t)} u_1\left(c\left(s\right), 1\right) \tag{7}$$

$$\forall s \in (t_R, T], \frac{\partial \mathcal{H}^1(\bullet)}{\partial a(s)} = -\dot{\mu}^1(s) \Rightarrow \dot{\mu}^1(s) = -r\, \mu^1(s) \tag{8}$$

$$\forall s \in (t_R, T],\ \dot{a}(s) = ra(s) + b(s) - c(s) \tag{9}$$

$$\frac{\partial V^1(t, t_R, a(t), a(t_R), r, b)}{\partial a(t_R)} = \mu^1(t_R) \tag{10}$$

Moreover, $\mathcal{P}^1$ being a constrained endpoint problem, we have to fulfill the transversality condition:

$$\mu^1(T)a(T) = 0 \Rightarrow a(T) = 0 \tag{11}$$

$\mathcal{P}^0$ and $\mathcal{P}^1$ verifying the standard strict concavity condition of their respective Hamiltonian, they both admit continuous and unique solution on their respective domain.

Let us now turn to the solution problem of the optimal value of the asset at retirement date $a^*(t_R)$. Relaying on the principle of optimality (Lemma 1), a necessary condition for $a^*(t_R)$ to be a maximum of $(V^0(\bullet) + V^1(\bullet))$ is:

$$\frac{\partial V^0(\bullet)}{\partial a(t_R)} + \frac{V^1(\bullet)}{\partial a(t_R)} = -\mu^0(t_R) + \mu^1(t_R) = 0 \Leftrightarrow u_1(c^0(t_R), \kappa) = u_1(c^1(t_R), 1) \tag{12}$$

It is easy to check that the left hand term of the last equality is increasing in $a(t_R)$ while the right hand one is decreasing, assuring the uniqueness of $a^*(t_R)$. If for all c, l in $\mathbb{R}^+ \times [0,1]$, $u_{12} < 0$, then $u_1(c^0(t_R), \kappa) < u_1(c^0(t_R), 1)$. Because $u_{11} < 0$, we can only have $u_1(c^0(t_R), \kappa) = u_1(c^1(t_R), 1)$, if and only if $c^0(t_R) > c^1(t_R)$. The reasoning is the same for the two other cases. □

In this setting, a negative cross derivative of the per period utility of consumption and leisure is necessary to obtain a discontinuous drop in consumption at the age of retirement, *i.e.* to resolve *the retirement consumption puzzle*. It means that, if we believe that the model is a proper simplification of the intertemporal choice of agent in the real world, the observation of that kind of drop, informs us on the negative sign of the cross derivative. It may seems strange because many workhorse utility function in labor economics such as the cobb-Douglas or the CES utility function are characterized by a positive cross derivative.

However, it is important to notice that relying on a different model of intertemporal choice with full endogeneity of labor, Heckman (1974) also conclude that a negative cross derivative of the per period utility of consumption and leisure was required to explain the hump shape of the intertemporal consumption profile.

In this part, I have given a complete theoretical treatment of an idea that was alluded in Banks et al. (1998) and in the "back-of-the-envelope calculation" in Battistin et al. (2009). This calculation was grounded on the following parametrical form:

$$u(c,l) = \frac{\left(c^{\alpha} l^{1-\alpha}\right)^{1-\gamma}}{1-\gamma}$$

with $\gamma > 0$ interpreted as the reciprocal of the intertemporal elasticity of substitution. They rightfully conclude that to solve the retirement consumption puzzle in this model, $\gamma > 1$ is required, but they miss the right insight for explaining that. Because, in this model, γ fully capture the intensity of the response of consumption to a variation of the rate of interest only when leisure is fully endogenous, but in this case there will be no discontinuity in the consumption function. As we have shown, explaining such a discontinuity, requires leisure to be exogenous at the age of retirement[2], then it is $-c\,u_{11}/u_1 = \alpha(\gamma - 1) + 1$ that will capture the intensity of response of consumption to a change of the rate of interest. Moreover, if the model is based on a Cobb-Douglass utility function, it is in fact a power transformation of a Cobb-Douglass, a transformation that can alter the sign of the second order cross derivative. We have $u_{12} = \alpha(1-\alpha)(1-\gamma)c^{\alpha(1-\gamma)-1} l^{(1-\alpha)(1-\gamma)-1}$. With this special parametrical form, the sign of the cross derivative of utility is fully given by the position of γ with respect to unity. When γ is higher than one, this cross derivative is negative explaining the downward discontinuity in consumption, as confirmed by the general statement of Proposition 1.3. The effect has nothing to do with the the intertemporal elasticity of substitution *per se*.

[2]Or at least a constraint for a minimum per-period work duration that is binding.

3 Optimal age of retirement

I have solved the program $\mathcal{P}$ with the age of retirement, t_R, being a parameter. I have all the material to characterize the optimal age of retirement, the one that maximizes the value of the program. In particular, the decomposition of the general Program in two sub-programs delimited by the age of retirement, allows to derive this optimal age of retirement in a parsimonious and elegant manner.

Proposition 2 (The optimal age of retirement)**.**
When an interior solution exists, and denoting $b^0(t_R) \stackrel{def}{=} \lim_{s \to t_R^-} b(s) < 0$*, the optimal age of retirement* $\hat{t}_R$ *is such that:*

$$u(c^1(\hat{t}_R), 1) - u(c^0(\hat{t}_R), \kappa)$$
$$= u_1(c^0(t_R), \kappa) \left[((1-\kappa)w(\hat{t}_R) + b^0(\hat{t}_R)) - b(\hat{t}_R)) + (c^1(\hat{t}_R) - c^0(\hat{t}_R))\right] \quad (13)$$

Proof: $\hat{t}_R$ is a solution of $\max_{t_R} V^*(t, t_R, T, a(t), r, w)$. Because V^* is continuous and differentiable in t_R, a necessary condition for having an interior solution is:

$$\frac{\partial V^*(t, t_R, T, a(t), r, w)}{\partial t_R} = 0 \quad (14)$$

Relying on Lemma 1 and noting that by construction of the Hamiltonian and Pontryagin maximum principle:

$$\frac{\partial V^0(t, t_R, a(t), r, w)}{\partial t_R} = \mathcal{H}^0(c^0(t_R), a^0(t_R), \mu^0(t_R), t_R)$$

and

$$\frac{\partial V^1(t_R, T, a(t), r, w)}{\partial t_R} = -\mathcal{H}^1(c^1(t_R), a^1(t_R), \mu^1(t_R), t_R)$$

we can easily conclude that $\hat{t}_R$ is such that:

$$\mathcal{H}^0(c^0(\hat{t}_R), a^*(\hat{t}_R), \mu^0(\hat{t}_R), \hat{t}_R) = \mathcal{H}^1(c^1(\hat{t}_R), a^*(\hat{t}_R), \mu^1(\hat{t}_R), \hat{t}_R) \quad (15)$$

Using the definitions of the Hamiltonian and first order conditions of program $\mathcal{P}^0$ and $\mathcal{P}^1$ and remembering that, in any case, a is continuous in t_R, we get the right hand side. □

This is a standard marginal condition for optimality. The left hand side of Equation (13) is the direct cost in utility of a marginal increase in the retirement age, while the right hand side is the indirect gain in utility due to supplementary resources generated by a longer work duration. The important and innovative point is that when taking into account the retirement consumption puzzle, the endogenous drop of consumption implies that less resources are required to maintain a same level of utility. Thus the earnings differential can be higher when the agents decide to retire.

Proposition 2 provides a very general characterisation of the optimal retirement age. Moreover, when expanding consumption before and after retirement as implicit function of the parameters of the problem, and when endogenizing the budgetary constraint of the social security system, it allows to derive comparative static results on the optimal age of retirement.

4 Conclusion

This short paper provides a general methodology to resolve the retirement consumption puzzle and the choice of the optimal age of retirement. The principle is illustrated in a simple model of intertemporal choice in which utility depend on consumption and leisure with certain horizon. To solve the puzzle we need only two assumptions: 1. retirement implies a discontinuity in the leisure intertemporal profile and, 2. the cross-derivative of the utility function is negative. But the method is general and can easily be extended in more realistic models with uncertain lifetime[3].

[3]In a companion paper, I am actually working on a calibrated version of the model taking into account a realistic modeling of uncertain lifetime in the spirit of Drouhin (2015) and the possibility of a non stationary intertemporal utility, allowing for per period utility to change with age in the spirit of Drouhin (2017)

References

Angeletos, G.-M., D. Laibson, A. Repetto, J. Tobacman, and S. Weinberg (2001). The hyperbolic consumption model: Calibration, simulation, and empirical evaluation. *The Journal of Economic Perspectives 15*(3), pp. 47–68.

Banks, J., R. Blundell, and S. Tanner (1998). Is there a retirement-savings puzzle? *American Economic Review*, 769–788.

Battistin, E., A. Brugiavini, E. Rettore, and G. Weber (2009). The retirement consumption puzzle: Evidence from a regression discontinuity approach. *The American Economic Review 99*(5), pp. 2209–2226.

Bellman, R. (1957). Dynamic programming. *Princeton University Press 89*, 92.

Bernheim, B. D., J. Skinner, and S. Weinberg (2001). What accounts for the variation in retirement wealth among us households? *American Economic Review*, 832–857.

Bloom, D. E., D. Canning, and M. Moore (2014). Optimal retirement with increasing longevity. *The Scandinavian journal of economics 116*(3), 838–858.

Boucekkine, R., D. de la Croix, and O. Licandro (2002). Vintage human capital, demographic trends, and endogenous growth. *Journal of Economic Theory 104*(2), 340 – 375.

Bütler, M. (2001). Neoclassical life-cycle consumption: a textbook example. *Economic Theory 17*(1), 209–221.

Chang, F.-R. (1991). Uncertain lifetimes, retirement and economic welfare. *Economica*, 215–232.

d'Albis, H. and E. Augeraud-Véron (2008). Endogenous retirement and monetary cycles. *Mathematical Population Studies 15*(4), 214–229.

d'Albis, H., S.-H. P. Lau, M. Sánchez-Romero, et al. (2012). Mortality transition and differential incentives for early retirement. *Journal of Economic Theory 147*(1), 261–283.

Drouhin, N. (2015). A rank-dependent utility model of uncertain lifetime. *Journal of Economic Dynamics and Control 53*(0), 208 – 224.

Drouhin, N. (2017). Non stationary additive utility and time consistency. working paper - halshs-01238584v3.

Hamermesh, D. S. (1984a). Consumption during retirement: The missing link in the life-cycle. *The Review of Economics and Statistics 66*(1), 1–7.

Hamermesh, D. S. (1984b). Life-cycle effects on consumption and retirement. *Journal of Labor Economics 2*(3), 353–370.

Hazan, M. (2009). Longevity and lifetime labor supply: Evidence and implications. *Econometrica 77*(6), 1829–1863.

Heckman, J. (1974). Life cycle consumption and labor supply: An explanation of the relationship between income and consumption over the life cycle. *The American Economic Review 64*(1), pp. 188–194.

Heckman, J. J. (1976). A life-cycle model of earnings, learning, and consumption. *Journal of Political Economy 84*(4), pp. S11–S44.

Heijdra, B. J. and J. O. Mierau (2012). The individual life-cycle, annuity market imperfections and economic growth. *Journal of Economic Dynamics and Control 36*(6), 876–890.

Heijdra, B. J. and W. E. Romp (2009). Retirement, pensions, and ageing. *Journal of Public Economics 93*(3-4), 586–604.

Kalemli-Ozcan, S. and D. Weil (2010). Mortality change, the uncertainty effect, and retirement. *Journal of Economic Growth 15*(1), 65–91.

Prettner, K. and D. Canning (2014). Increasing life expectancy and optimal retirement in general equilibrium. *Economic Theory 56*(1), 191–217.

Sheshinski, E. (1978). A model of social security and retirement decisions. *Journal of Public Economics 10*(3), 337–360.