The Negative Natural Rate of Interest in the Modern Theories of Liquidity Trap and Secular Stagnation: Back to Böhm-Bawerk via Samuelson

Stefano Di Bucchianico
Department of Economics and Statistics
University of Siena
stefano.dibucchianico@unisi.it

Abstract

The negative natural rate of interest is since two decades eliciting theoretical and policy debates. It re-emerged, after a relatively long time, in Krugman’s Liquidity Trap model. Later, it was placed at the hearth of the Secular Stagnation theory by Summers. It is argued that Krugman’s negative natural rate of interest ensues from theoretical premises analogous to those present in Samuelson’s overlapping-generations model. In turn, Samuelson obtained a negative equilibrium interest rate by opportunely recasting Böhm-Bawerk’s three causes for a positive rate of interest. The present paper illustrates and analyses this neglected line of thought, until its recent developments.

Keywords: Secular Stagnation; Liquidity Trap; negative natural rate of interest; Krugman; Samuelson.

JEL: B13; B20; B30.
"As Knut Wicksell, Böhm-Bawerk’s admirer, pointed out: in the end we merely accept what technology is and what human decisions are." (P. A. Samuelson)

1. **Introduction**

The negative natural rate of interest (NNRI) is, since more than two decades, at the centre of the attention of economic theory and policy. In fact, the modern theories of Liquidity Trap and Secular Stagnation crucially ground on this concept. As for the former, Krugman’s (1998), one of the most important contributions in macroeconomic theory in the last decades, intended to modernise Hicks’ (1937) Liquidity Trap (and to understand the cause of Japan long-lasting stagnation). In this new vestige convention-al monetary policy loses effectiveness due to the presence of an NNRI and of the zero lower bound (ZLB) on the nominal rate of interest. As for the latter, Summers’s (2014, 2015, 2018) resurrection of Hansen’s (1939) Secular Stagnation hypothesis opened a new strand of literature that is based on the supposition that the NNRI permanently features advanced capitalist economies. This second line of enquiry was preceded by Krugman’s contribution (Backhouse and Boianovsky 2016a, p. 963) and it closely tracks its logical structure (Di Bucchianico 2020a). Krugman himself has explicitly made the connection clear (Krugman 2013).

On the policy side, the fact that the NNRI cannot be reached because of the ZLB en-couraged the study of novel strategies of intervention. On the monetary policy side, exceptional policy crackdowns such as unconventional monetary actions were explicit-ly encouraged (Krugman 2000; Eggertsson and Woodford 2003; Fiebiger and Lavoie 2020) and, since some years, the possibility of negative interest rate policy is subject to mounting debate (Rogoff 2017). The rationale beneath the case for implementing this kind of policy is crucially based on the supposition that the natural rate of interest is negative (Di Bucchianico 2020c). On the fiscal policy side, other authors vocally buttress the case for active fiscal policy given monetary policy ineffectiveness due to the NNRI (Summers 2016, Eggertsson et al. 2019).

The objective of this paper is to uncover and analyse a line of connection in the his-tory of economic thought that links the debate about the reasons why we can expect the rate of interest to be positive to those recent contributions in which we find the suppo-sition that the natural rate of interest has turned negative. This contribution aims to en-rich the study of the origins of the NNRI, which enjoy a key role in the Liquidity Trap and the Secular Stagnation theories (Backhouse and Boianovsky 2016a, 2016b; Boianovsky 2004, 2016).
According to our reconstruction, the NNRI in Krugman’s (1998) model emerges in response to assumptions which closely resemble the analytical background employed by Samuelson (1958) when setting forward his overlapping-generations model. Samuelson, in turn, obtained in this model a negative equilibrium rate of interest by ignoring or reversing the famous ‘three causes’ for a positive rate of interest listed by Böhm-Bawerk (1889 [1930]). Read in this perspective, the model of Krugman can be indicated as the linchpin through which the modern theories of Liquidity Trap and Secular Stagnation are linked to the history of the NNRI. We, therefore, contend for a more comprehensive understanding of the modern reappraisal of the negative value of the natural rate of interest to be achieved when this (to the best of our knowledge, unnoticed) line of connection is taken into account. Nevertheless, we maintain that Krugman’s remarkable analytical enterprise did not allow for a robust advancement in the analysis of the NNRI. Summers’ contribution, by introducing capital in the analysis, brings the theoretical treatment one step forward but cannot avoid the problems entailed by the use of an NNRI.

The article proceeds as follows: Section 2 presents Krugman’s (1998) model and maps its ramifications until Summers’ (2014) proposal; Section 3 discusses the line of connection linking Krugman to Böhm-Bawerk via Samuelson; Section 4 investigates whether Krugman brought the analysis of the NNRI forward and offers a ‘map’ of the connections we discuss; and Section 5 concludes.

2. The Negative Natural Rate of Interest in Krugman’s Liquidity Trap and Summers’ Secular Stagnation Theory

Krugman (1998) intended to revive, with suitable modifications, the Keynesian-Hicksian Liquidity Trap. This served to explain the long-lasting stagnation of Japan (Krugman 2000). According to the 2008 Nobel Prize recipient, a Liquidity Trap emerges when monetary policy can no longer steer the rate of interest by controlling the money supply, “because nominal interest rates are at or near zero: injecting monetary base into the economy has no effect, because base and bonds are viewed by the private sector as perfect substitutes” (Krugman 1998, p. 141). The reason why the central bank should try to steer the nominal interest rate as low as possible lies in the emergence of an NNRI. The natural interest rate is supposed to turn negative when potential output is assumed to decrease over time (Krugman 1998, p. 147):
The condition under which the required real interest rate is negative is straightforward in this simple endowment economy. Market clearing will require a negative real interest rate if the marginal utility of consumption in period two is greater than that in period one, which will be the case if the economy’s future output is expected to be sufficiently less than its current output.

Therefore, in a simple Fisher-equation description of the real interest rate, we see that:

\[ r = i - \pi_e, \quad r^* < 0 \]  

(1)

If the real interest rate required for equilibrium \( r^* \) is negative, the market real interest rate \( r \) needs to be negative as well. Given inflation expectations, one strategy might be for the monetary authorities to set a negative nominal rate of interest (Rogoff 2017; Di Bucchianico 2020c). Unfortunately, given the ZLB, Krugman’s contention was that the NNRI is out of reach. The alternative he envisaged contemplates management of inflation expectations. Despite the ZLB, raising inflation expectations can generate an appropriately negative real interest rate, provided that the central bank’s policy is considered credible by agents. However, the public is convinced that monetary authorities will, sooner rather than later, revert to conservative policy. The central bank is hence compelled to ‘credibly promise to be irresponsible’, remarkably raising the inflation target and engaging in a ruthless crackdown to steer agents’ expectations in the desired direction.

2.1. A Sketch of the Model

In the model, the representative agent has a utility function of the following type:

\[ U = \frac{1}{1 - \rho} \sum c_t^{1-\rho} \beta^t \]  

(2)

where \( c_t \) is consumption, \( \rho \) is the relative risk aversion, \( \beta \) is the discount factor. The latter, as commonplace, is supposed to be positive but lower than unity. There are no savings, and consumption is the only component of output. The representative agent, subject to a ‘cash in advance’ constraint, trades cash for one-period bonds, for the sake of obtaining the only available consumption good. Money supply \( M \) is exogenous and the price level today \( P \) is given by the Quantity Theory equation:

\[ P = \frac{M}{y} \]  

(3)

where output today (equal to endowment today) \( y \) is given. The velocity of circulation of money is (implicitly) supposed to be unity. In strict analogy, the price level tomorrow \( P^* \) is fixed at \( P^* = \frac{M^*}{y^*} \) once both \( M^* \) and \( y^* \) are given.
From (2), by exploiting the Euler equation for optimal intertemporal consumption allocation, Krugman links the price level today and the nominal interest rate:

\[ 1 + i = \frac{P^*}{\beta P} \left( \frac{y^*}{y} \right)^{\rho} \]  

(4)

In equation (4) there are five givens: the discount factor and the relative risk aversion (preference parameters); the two endowments for today and tomorrow (output); and the future price level. When the current price level rises, the nominal interest rate falls.

To derive the natural interest rate, we rearrange equation (4). The natural rate of interest \( r \) is exogenously set given preferences and endowments:

\[ 1 + r = \frac{1}{\beta} \left( \frac{y^*}{y} \right)^{\rho} \]  

(5)

It turns negative when future output is expected to shrink sufficiently when compared to current output. Assuming the relative risk aversion equal to unity for simplicity, the condition boils down to:

\[ \frac{y^*}{y} < \beta \]  

(6)

Condition (6) entails that the marginal utility of consuming tomorrow is higher than today since future output is lower than current output. The other element to be taken into account is the ZLB:

\[ i \geq 0 \]  

(7)

The Central Bank can fix the price level in (3) by injecting an appropriate amount of money and, indirectly, the nominal interest rate in (4) as well. Nevertheless, this mechanism operates, in the graph below, down to point 2 at most (the ZLB). If an NNRI emerges, point 3 cannot be achieved (it would entail a negative nominal interest rate).
In a flexible price environment, Krugman argues, the price level today would drop, thereby raising inflation expectations: the NNRI can be achieved despite the ZLB. However, if a rigidity on the current price level is introduced, price stickiness prevents the economy from hitting the NNRI through deflation. We thus have all the pieces needed in (i) to describe Krugman’s intuition synthetically: an NNRI, the ZLB, and, with given prices today and tomorrow, a fixed $\pi^e$. Since price rigidity prevents deflation, the system is stuck in an underemployment equilibrium. The ensuing policy prescription is thus in favour of an effort to raise $P^*$: if it can be increased aptly, the required inflation expectations allow to hit the NNRI.

Krugman (1998, pp. 150-151) then introduces investment in a stylised overlapping generations example. In it, the young cohort cultivates land and uses the proceeds to buy it from the old. The elder cohort consumes what it earns from selling the land. The expected rate of return $r_t$ on buying land is:

$$1 + r_t = \frac{R_{t+1} + q_{t+1}}{q_t}$$  \hspace{1cm} (8)

where $R_{t+1}$ is the marginal product of land and $q_t$, $q_{t+1}$ are the prices of land today and tomorrow. Krugman states that a sufficient population decline, causing the price of land to fall, can make the rate of return negative:

$$r_t < 0: q_t - q_{t+1} > R_{t+1}$$  \hspace{1cm} (9)

thereby showing the possibility for an NNRI to emerge even when investment is introduced.

2.2. From Liquidity Trap to Secular Stagnation
Krugman (1998) exerted a strong influence, which is still present in the background of central banks’ operative strategies (Fiebiger and Lavoie 2020), on subsequent literature dealing with monetary policy at the ZLB, liquidity trap and stagnation. A synthetic timeline can be depicted (Di Bucchianico 2020a). In a first phase, the publication of Krugman (1998) kick-started a series of contributions which focused on the unconventional monetary policy way out of a liquidity trap. Among the closest successors we find the works of Eggertsson and Woodford (2003, 2004), which heavily ground on Krugman (1998) and expand its main features by means of fully-fledged intertemporal models.

In a second phase, Japan’s stagnation and the unravelling of the Great Recession in the USA contributed to shift attention to fiscal policy. Krugman (2005) himself contributed to this shift, claiming for the necessity to rethink fiscal policy when monetary policy is constrained by the ZLB. This tendency can be appreciated in several contributions neatly expressing renewed interest in deficit spending fiscal policy at the ZLB (see among others Woodford 2011; Eggertsson and Krugman 2012).²

During the third phase, lasting stagnation led to progressively move the attention from the theory of Liquidity Trap to that of Secular Stagnation (Summers 2014, 2015, 2016, 2018). In Summers’ view “this experience was consistent with a negative Wicksellian natural rate of interest, implying that full-employment saving exceeded investment at any nonnegative interest rate” (Backhouse and Boianovsky 2016a, p. 946). On the one hand, the tendency of major firms to progressively use less physical capital, the falling relative price of investment goods in relation to consumption goods, and shrinking labour force growth curb the demand for investment. On the other hand, rising income shares accruing to the top 1% and ageing population intensify the supply of savings. Their joint effect on the natural rate of interest is to cause its continuous fall until it reaches negative values (from point A to point B in Fig. 2).
As in Krugman’s (1998) contribution, if the policy-controlled nominal interest rate is constrained by the ZLB, the NNRI can remain out of reach (Summers 2016, p. 3):

Following the Swedish economist Knut Wicksell, it is common to refer to the real interest rate that balances saving and investment at full employment as the “natural,” or “neutral,” real interest rate. Secular stagnation occurs when neutral real interest rates are sufficiently low that they cannot be achieved through conventional central-bank policies. At that point, desired levels of saving exceed desired levels of investment, leading to shortfalls in demand and stunted growth.

Summers’ proposal grounds on the logical structure of Krugman’s Liquidity Trap, but it makes the NNRI a permanent feature of an economy endowed with capital, and decidedly turns the attention to deficit-spending fiscal policy.

3. The Negative Natural Rate of Interest: Back to Samuelson and Böhm-Bawerk

In this section we aim to highlight the resurgence in Krugman’s (1998) seminal paper, after a relatively long time, of the issue of what causes the interest rate to be either positive or negative starting from premises analogous to those that can be found in previous contributions. In particular, while the origin of the enquiry on the reasons why we normally expect the rate of interest to be positive can be traced back to Böhm-Bawerk (1889 [1930]), Samuelson (1958) showed the possibility to get the opposite result grounding on the same reasons Böhm-Bawerk discussed. Therefore, while the general structure of the model proposed by Krugman (1998) draws heavily on Hicks’s (1937) reasoning in terms of intertemporal price elasticity and money/bonds substitutability (let alone the choice of renovating the liquidity trap concept) (Boianovsky 2004), the case involving an NNRI links
that seminal paper also to those other two authors. Let us, therefore, try to reconstruct the rationale for our claim.³

At first, we list Krugman’s (1998) main assumptions, those that play the most significant role in order to get an NNRI. As seen, Krugman reaches this fundamental result in an intertemporal, two-period model. In it:

i. the economy is supposed to be poorer tomorrow than today, given that consumption endowments decrease in the second period;

ii. agents strictly prefer consumption today to consumption tomorrow, as they are characterised by a positive (and lower than unity) time discount factor $\beta$;

iii. in the model there is only one consumption good, no capital is present.⁴

Keeping these hypotheses in mind, we now expand on our proposed connection. We focus initially on the link between Böhm-Bawerk’s causes for a positive rate of interest and Krugman’s (1998) modern Liquidity Trap.

Böhm-Bawerk, in his *The Positive Theory of Capital* (1889 [1930]) offered his “vision of how the interest rate might be determined by the interplay of systematic time preference (‘impatience’) and time-phased technology’s productivity” (Samuelson 1994, p. 202). He deemed the origin of a positive rate of interest to be owed to three causes (Böhm-Bawerk 1889 [1930]; Blaug 1997, pp. 480-488):

i. better provision for wants expected in the future than in the present;

ii. undervaluation of future wants;

iii. the superiority of more roundabout methods of production.

His investigation on the nature and causes of interest obviously entailed many other facets besides the discussion of the three causes.⁵ For example, Böhm-Bawerk also needed to devise a measure of an economy’s endowment of capital independent of the rate of interest. This, among other things, led him to construct a measure of the ‘average period of production’, which however failed its aim (Gehrke and Kurz 2009).

Samuelson (2001, pp. 302-304; emphasis in the original) described the three causes very clearly as follows:⁶

Böhm’s *first cause* of positive interest [...] recognized that if in the future I will be richer than I am today, I can afford to pay a positive interest premium to borrow and thereby make my consumption stream more smooth. [...] Böhm’s *second cause* involved asymmetric time preference for present consumption versus future consumption: whether rationally or irrationally, most people would (other things equal) prefer the pair ($2000 now, $1000 later) to ($1000 now, $2000 later) [...] Böhm offered his important *third cause* for interest, namely the brute technological fact that more “time-
intensive,” more “round-about,” and more “capital-intensive” processes (somehow measured) allegedly do create extra consumable harvests from the same totals of labor and natural resources.

The formulation of a theory for the positivity of the rate of interest in these terms attracted much attention, and several authors engaged with a discussion of the ‘three causes’. One of the preeminent examples is Wicksell (1893 [1954, pp. 106-119]) who, besides discussing the issue, also reinstated the problem, transforming the ‘three causes’ framework “into an explicit theory of interest as the marginal productivity of waiting, coordinated with the marginal productivity theories of wages and rent” (Uhr 1951, p. 844). In this context, if capital relative scarcity determines the rate of interest, and the latter settles to a level higher than the rate of time preference, net capital formation ensues and with it rising per capita income (Gehrke and Kurz 2009, p. 75). However, in such a scenario the rate of interest would be determined by the third cause, while the first is operative only in consequence of a divergence between the interest rate and the rate of time preference that accounts for the second cause.

Going beyond the exclusive focus on Wicksell, Boianovsky (1998) illustrated the various objections that the Swedish economist, but also Ramsey, addressed to Böhm-Bawerk’s formulation, while Samuelson (1994, 2001) also brought Fisher and Hayek into the picture.

Let us now move towards the comparison between the tasks of Böhm-Bawerk and Krugman. Whereas the former wanted to understand why the rate of interest is generally positive, the latter’s target was to place an NNRI at the core of his theory. Therefore, to get acquainted with how the three causes can be changed and mixed to obtain results different from that of Böhm-Bawerk, we consider two simple examples where there is a zero rate of interest. We do so by referring at first to Blaug (1997), who argued that:

the rate of interest can only be zero when (1) the flow of income is constant through time; (2) time preference is neutral; and (3) the net product cannot be increased by postponing consumption for the sake of future production. (ibid., p. 487)

the first and the second reasons for a positive rate of interest are absent by definition under stationary conditions; if the third reason is also inoperative, the rate of interest will be zero. [...] So long as time preference is positive, however, zero net productivity of capital or synchronisation of production and consumption will not reduce the rate of interest to zero. (ibid., p. 502)

In addition to this example, it is possible to find another instance featuring a zero interest rate - in which the result is attained by referring to Böhm-Bawerk’s framework - in the well-known textbook Economics (Samuelson and Scott 1966). The two authors affirm that the rate of interest is determined by both ‘impatience’ and ‘productivity’. By means of a graphical example analogous to that we will show below (cf. Fig. 3), they demonstrated how these two
factors contribute to the formation of a positive rate of interest. Nevertheless, supposing their absence (as the authors did) would make the positive rate of interest unwarranted: “Having thus ruled out net productivity and time preference, we should find that the equilibrium interest rate must then be zero” (ibid., p. 670). These two examples allow to focus on the fact that, while the three causes (and their interactions) had been employed by Böhm-Bawerk to study the positivity of the rate of interest, at the logical level some peculiar assumptions may make the final result different. The case in which, according to a particular formulation of the three causes, the rate of interest settles at a zero level is however not only a curiosum. In the first chapter of Schumpeter’s Theory of Economic Development (1934), among several other things, a discussion can be found about the properties of an economic system settled to “a circular flow”, a stationary and synchronised economic process in which there is no uncertainty about the future” (Blaug 1997, p. 502). In such a condition, Schumpeter stated, the interest rate would turn out to be zero, and he maintained that by recalling (his former mentor) Böhm-Bawerk’s ‘three causes’ (Elliott 1985, p. 24):

In the strictly stationary conditions of the circular flow, Schumpeter argued, interest rates would be zero, because of the assumed absence of systematic time preference for the present over the future (based, in turn, on synchronisation of production and consumption) and because opportunities for additional profits through investment in more roundabout methods of production become exhausted when those methods are a matter of routine.

Therefore, the zero-interest rate result Schumpeter arrives at depends on the suppression of all the three causes: “Denying all of Bohm-Bawerk’s three causes, Schumpeter cogently denies positivity to the sans-development interest rate” (Samuelson 1994, p. 205). This kind of argument led also to a debate between Schumpeter and Böhm-Bawerk themselves precisely on whether in a stationary condition the rate of interest would be positive or nil (Elliott 1985, p. 25).

After the reflection on the zero-interest rate case, we can now move to that in which the rate of interest becomes negative. Samuelson (1958) attempted to determine the equilibrium interest rate in an economy in which, over the years, different generations overlap. He obtained the equilibrium interest rate in a market among generations in which loans are demanded only for consumption.9 The main assumptions of the model were:

i. each cohort of agents experiences a three-period life: in the first two periods, individuals work and produce, while in the third they retire and have to find a way
to keep on consuming without receiving any labour income. Thus, later years are not wealthier but poorer;

ii. the arguments of the ordinal utility function of the representative consumer are the three-period dated consumption quantities, but there is no formalisation for the subjective discount factor;

iii. no good can be used as a store of value because none keeps through time (in his words, an "extreme assumption").

During the introduction to the formal setup of the model, Samuelson recalled Böhm-Bawerk’s three causes, and then he stated the problem in these terms (ibid., p. 469):

(Thus Bohm’s second cause of interest may or may not be operative; it could even be reversed, men being supposed to overvalue the future!) In addition to ignoring Bohm’s second cause of systematic time preference, I am in a sense also denying or reversing his first cause of interest[...] Finally, recall our assumption that no goods keep, no trade with Nature being possible, and hence Bohm’s third technological cause of interest is being denied. Under these assumptions, what will be the equilibrium time path of interest rates?

After having demonstrated that in such a stylised economy the interest rate can be seen as a by-product of human fertility (in his words, a “biological theory of interest”, because the rate of interest is equal in magnitude to the rate of growth of population), Samuelson also analysed the multiplicity of possible equilibrium interest rates in a situation of stationary population. In his numerical example, the relevant root from which to extract the equilibrium interest rate delivered a negative valued solution (ibid., pp. 477-478). How did the author comment such a surprising outcome?

Is this negative interest rate a hard-to-believe result? Not, I think, when one recalls our extreme and purposely unrealistic assumptions. With Böhm’s third technological reason for interest ruled out by assumption, with his second reason involving a systematic preference for the present soft-pedaled, and with his first reason reversed (that is, with people expecting to be poorer in the future), we should perhaps have been surprised if the market rate had not turned out negative. [...] It incidentally confirms what modern theorists showed long ago but what is still occasionally denied in the literature, that a zero or negative interest rate is in no sense a logically contradictory thing, however bizarre may be the empirical hypotheses that entail a zero or negative rate. (ibid., p. 479, emphasis in the original).

In a more recent article on Böhm-Bawerk’s capital theory, Samuelson reiterated the argument, although making use only of the inversion of the first cause: “In a modern society where one can expect to live a long time in a retirement without earning power, Böhm’s logic might rationalize a negative real interest rate!” (Samuelson 2001, p. 303; emphasis in the original).
The theoretical picture drawn by Samuelson appears clear. Given the nature of his assumptions, which either reverse or set aside the three causes, the NNRI follows straight. It is now time to relate Samuelson’s (1958) article to Krugman’s (1998) contribution in order to close the line of connection. The Krugmanian model, it can be shown, can indeed host an NNRI precisely because of two of the three reasons listed by Samuelson. In it, in the basic version, the natural interest rate is arrived at in an economy with only consumption, through the intertemporal maximisation of utility by the representative agent. As in the case explored by Samuelson, the first cause envisaged by Böhm-Bawerk for a positive rate of interest was reversed, as the future level of endowments was assumed to be lower than today’s level. On the contrary, the second reason was present, because agents systematically prefer consuming today rather than tomorrow. Again, like for Samuelson, technological concerns were set aside because there was neither a production function nor investment; Böhm-Bawerk’s third cause for the appearance of a positive interest rate was neglected.

In order to better grasp the analogy we want to draw among the authors we have been treating, we summarise their arguments in Table 1.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First cause</strong></td>
<td>Future income higher than today</td>
<td>Future income lower than today</td>
<td>Future income lower than today</td>
</tr>
<tr>
<td><strong>Second cause</strong></td>
<td>Systematic preference for consumption today</td>
<td>Neutrality between consumption today and tomorrow</td>
<td>Systematic preference for consumption today</td>
</tr>
<tr>
<td><strong>Third cause</strong></td>
<td>Superiority of more roundabout processes</td>
<td>Technological superior possibilities ruled out</td>
<td>Technological superior possibilities ruled out</td>
</tr>
<tr>
<td><strong>Rate of interest</strong></td>
<td>Positive</td>
<td>Negative</td>
<td>Negative</td>
</tr>
</tbody>
</table>

Table 1 – A synoptic table of Böhm-Bawerk, Samuelson and Krugman approaches to the three causes. Source: author’s elaboration.

In it, we can more schematically see the analogy that we detect in their reasoning about the sign of the rate of interest. While Böhm-Bawerk obtained a positive rate of interest out of his three grounds, Samuelson was able to reverse that result by tackling all three facets. Krugman, later on, kept such a result unchanged despite the reintroduction of the second cause in line with the original formulation of Böhm-Bawerk.

4. The Re-emergence of the Negative Natural Rate of Interest in Hindsight
4.1. Where does Krugman’s Work leave the Discussion on the Three Causes?

After having introduced the connection between Krugman, Samuelson and Böhm-Bawerk, we will now further elaborate on their relationship. Before starting, it is important to note that Krugman’s reasoning is always carried on dealing with a dynamic economy, not a stationary condition. Blaug (1997, pp. 480-489) carefully differentiates the discussion of these two different cases. We note that, were the economy stationary, the interest rate would be determined given the discount factor of the representative agent, as Krugman (1998, p. 144) pointed out right away when introducing the reader to his model. In fact, if endowments between the two periods are equal, equation (5) becomes:

\[ r^* = \frac{1 - \beta}{\beta} \]  

(10)

Keeping the non-stationarity assumption in mind, let us now re-read the discussion carried out in the previous section through the graphical representations in Fig. 3 and Fig. 4. We elaborate on Samuelson (1994, p. 204), Samuelson and Scott (1966, p. 670), who utilised a ‘Fisherian’ graph to better grasp Böhm-Bawerk’s insights:

![Figure 3](image-url)

**Figure 3** – A graphical sketch of the first two causes: (a) the case of Böhm-Bawerk. Source: author’s elaboration on Samuelson (1994), Samuelson and Scott (1966).

Case (a) in Fig. 4 represents the conventional Böhm-Bawerk rationalisation of the first two causes. Line OT has slope 1 (45 degrees): this means that along this line, the amount of available present \( (C_1) \) and future \( (C_2) \) consumption is the same. A situation like that would be
represented by drawing a square whose diagonal is the OG line. The slope of the indifference curve 1 in point F is 1 in absolute value: the second cause is not operative. A Böhm-Bawerkian intertemporal provision of consumption such that OA is longer than OC causes the slope of the line DE to be higher than 1 in absolute value. Hence, the rate of interest turns out to be positive at the tangency point between the two curves at point B and this is due to the effect of the first cause.

The case of indifference curve 2 is different. In fact, this indifference curve is steeper than indifference curve 1. In this instance, the slope of the indifference curve at point G is higher than 1 in absolute value: the second cause is operative. The line D’E’ is parallel to the DE line. Despite the two curves having the same slope, meaning that the first cause is operative also in this case, a positive rate of interest emerges already at point G. Given that it is placed on the OT line, this signals the fact that the rate of interest is positive even if the two-period endowments were equal, i.e. even in the absence of the first cause. This is due to the second cause (the systematic preference for consumption today), the element conferring a vertical bias to indifference curve 2.

**Figure 4** – A graphical sketch of the first two causes: (b) the case of Krugman. *Source*: author’s elaboration.

Case (b) in Fig. 4 depicts the Krugmanian version. If the provision of goods will be lower tomorrow, line OC is longer than OA, and this means that the first cause has been reversed. Indifference curve 1 has a slope equal to 1 in absolute value, and hence the second cause is not operative. In the absence of a systematic preference for consumption today, even a slightly smaller endowment today than tomorrow (as in the graph) suffices to let a negative interest rate emerge. Indeed, in point B the slope of the DE line is lower than 1 in absolute value. Hence, a negative rate of interest materialises.
When, on the contrary, we have a steeper indifference curve such as curve 2, this means that the second cause is operative. Therefore, in order to obtain a slope lower than 1 in absolute value for the D’E’ line a higher-then-before rate of decrease of endowments is needed. In the graph, line OA’ is considerably shorter than the line OC’, and in point G the slope of the D’E’ is lower than 1 in absolute value. This replicates Krugman’s (1998) result: despite the presence of a systematic preference for consumption today, the negative rate of interest turns negative when endowments tomorrow decrease by a sufficient extent.

At this point what must be asked is what the meaning of the interest rate in a model with only consumption is, notwithstanding the possibility to coherently represent it in accordance with Böhm-Bawerk’s first two causes. Let us in fact suppose that the whole reasoning made by Krugman holds. Nonetheless, the kind of interest rate he discusses is problematic when interpreting its theoretical relevance. The interest rate, in that case, is the by-product of the reasoning involving dated quantities of delivered amounts of a single consumption good. Indeed, it can be maintained that the interest rate calculated in that scenario is not a proper rate of interest, but rather an ‘own’ rate of interest \( r_{t,t+1} \), expressed by the relative price \( p_t / p_{t+1} \) of a single good delivered between the current period and the subsequent point in time (similarly to the dated commodities to be found in Arrow-Debreu models) (Fratini in Bellino et al. 2017, Fratini 2020a, pp. 102-103):

\[
\frac{p_t}{p_{t+1}} \equiv 1 + r_{t,t+1} \quad (11)
\]

This means that, as Fratini (2020b) states, it is not the price of a factor ‘capital’, but rather a simple relative price of a single commodity delivered in two different dates, on the same footing as that which can be used to express the intertemporal trade of any other commodity. But if it is so, there would obviously be no particular reason why such an interest rate cannot turn negative: the quantity of the good delivered tomorrow can be lower than today, and so the own-rate of interest can be negative. This consideration, however, carries no explicative power with respect to what that rate of interest should mean in concrete terms. As Di Bucchianico (2020a, pp. 109-110) posits, the rate of interest we find in Krugman’s model can be basically read in these terms, and therefore is no more than a factor that expresses the rate of interest in terms of an exchange of the same commodity (the only single consumption good available) in different dates (today and tomorrow). This issue leads us straight to two crucial considerations we want to stress.

First, it can be doubted whether Krugman’s contribution truly brought the analysis of the ‘three causes’ to a more advanced level. Let us in fact refer to Samuelson (1958). Once he acknowledged the possibility to obtain an NNRI in his overlapping-generations model, he called for successive possible improvements over the basic framework. Specifically, he suggested to introduce (i) technological investment possibilities, (ii) technical innovations.
raising incomes on a secular basis, (iii) a subjective discount factor, (iv) imperfect competition, (v) uncertainty, and to take into account an intertemporal framework in models with capital (ibid., p. 479, 482). Point (ii), linked to Böhm-Bawerk’s first cause, has been left where Samuelson did sixty years ago: in Krugman (1998) there is no presence of technical innovations raising incomes in the longer run. Since the beginning of his enquiry, Krugman has introduced point (iii), namely “strong biases toward present goods and against future goods”, which is nothing else than the beta discount factor in equation (2). Such an improvement makes the condition about the realisation of an NNRI stricter than what would have been otherwise, as we have seen in Fig. 4.¹¹ Thus, we can say that Krugman’s insertion of a systematic preference for present consumption deals with Samuelson’s point (iii): such a factor renders the condition for the appearance of an NNRI more stringent. However, this aspect does not make a great deal of difference. The fundamental hypothesis remains the supposition of a poorer economy tomorrow. What changes is how strong the endowments’ rate of decrease has to be. As seen in sec. 2.1, Krugman also briefly tackles the introduction of technological investment possibilities, Samuelson’s point (i).¹² We noted that the exercise is done by supposing a model with labour and land, the latter being a factor of production utterly different with respect to capital. It is in fact given in fixed supply and yields rent instead of interest. Moreover, the fact that it is given in fixed supply rules out the possibility to save present goods for the sake of making the production process more ‘roundabout’, contrary to what would happen with the production of capital goods. One might thus say that point (i), referring to the third cause, has been basically disregarded. In addition, even setting aside this criticism, a negative rate of return as in equation (7), known to the agents with certainty, would drive them towards acquiring some kind of stores of value. Indeed, Samuelson (1958) explicitly referred to the absence of any possible good that might have served as store of value because if such a storage possibility is there, then the rate of interest cannot turn negative. This is particularly problematic in the case of Krugman because in there the ZLB is caused precisely by the presence of money, a store of value. As Boianovsky (2016, p. 44) notes, this issue was already clear also to Böhm-Bawerk, who “denied the possibility of a negative rate of interest even if the income stream is declining, unless there are no durable goods able to function as stores of value”. Moreover, reversing the first cause and ignoring technical innovations has heavy consequences in terms of plausibility of the explanation. Krugman in fact contended that the NNRI emerged due to a labour force shrinkage leading to potential
output contraction (the now familiar decrease of the endowments tomorrow). However, even in a basic Solow (1956) growth model shrinking population growth rates can be reconciled with a growing potential output if considering technical progress, something Krugman did not (Di Bucchianico 2020a).

Second, in more general terms, there is a remarkable difference between Samuelson and Krugman when it comes to elicit policy prescriptions from their analytical frameworks. Indeed, as seen, Samuelson looked at the solution involving an NNRI as a sort of basic result obtained under overtly restrictive assumptions, an initial step to be integrated with a long list of improvements. Samuelson was not keen on providing any concrete policy recommendation. Rather, he reflected on the role of a possible social contract according to which youngsters should commit to provide also for retired elders, under the agreement that they would in turn be provided by newly born generations in their future. In this view, Samuelson offered an interpretation of the role of money as a ‘social compact’ that, by allowing to keep purchasing power over the years, allows to overcome a market failure. Thus, even without a social contract interfering with the market mechanism, “society by using money will go from the non-optimal negative-interest-rate configuration to the optimal biological-interest-rate configuration” (Samuelson 1958, p. 482). Later he worked on refining the rationale of the ‘third cause’ by means of his famous ‘surrogate’ production function (Samuelson 1962) but, as Boianovsky (2020, p. 627) points out, while Diamond (1965) extended Samuelson’s (1958) framework, the 1970 Nobel Prize recipient did not pursue further the research on overlapping-generations modelling. Krugman, while on the one hand basically retained Samuelson’s framework, on the other hand wished to take a firm stance on the explanation for the Japanese lasting stagnation. Accordingly, after having established his basic theoretical results in the first half of the paper, in the second he moved forward to discuss the empirical side of the issue and to more clearly state his case in favour of unconventional monetary policy. Following his reasoning, “[i]f the central bank can credibly promise to be irresponsible - that is, convince the market that it will in fact allow prices to rise sufficiently - it can bootstrap the economy out of the trap” (Krugman 1998, p. 161).

Let us now, at last, bring the third cause back into the discussion. What happens when the possibility to increase net output in the future by resorting to more ‘roundabout’, ‘mechanised’, ‘capital-intensive’ methods of production is explicitly introduced in our discussion? This point, as seen, did not enjoy much consideration by Krugman but is visibly
present in the recent reappraisal of the Secular Stagnation theory (Summers 2014, 2015, 2018; Eggertsson et al. 2019), where an NNRI is considered despite the presence of capital and its marginal product. At this point the subject must be split in two different, albeit contiguous, logical stages.

The first involves the discussion of whether the introduction of the third cause can in actual fact contribute to justify the existence of an interest rate by means of its linkage to the degree of roundaboutness in production. As well-known, this relationship has been seriously questioned during the “Cambridge capital controversies” (Lazzarini 2011). To cut a long story short, so as to focus more specifically on the point of our interest, the main outcomes of that controversy showed, among other things, that ‘reswitching’ and ‘reverse capital deepening’ make

it impossible to say unambiguously either that (1) a fall in the rate of interest will always alter the rankings of the most profitable of all available techniques in a unidirectional manner, or that (2) it will always increase the capital-intensity of the economy by promoting a greater degree of ‘roundaboutness’. (Blaug 1997, p. 505)

Hence,

the simple tale told by Jevons, Böhm-Bawerk, Wicksell, and other neoclassical writers - alleging that, as the interest rate falls in consequence of abstention from present consumption in favor of future, technology must become in some sense more “roundabout,” more “mechanized,” and “more productive” - cannot be universally valid. (Samuelson 1966, p. 568)

Accordingly, Blaug (1997, p. 504), with a crude metaphor, labelled this kind of criticism “the final nail in the coffin of the Austrian theory of capital”. This of course represents a very strong criticism launched against some basic neoclassical principles and in particular to the ‘third cause’. In addition to this, Samuelson (1994, 2001) showed how the problem in linking the rate of interest in an inverse relation with the ‘degree of roundaboutness’ in production arises even when reswitching is ruled out. Samuelson (2001, p. 307) through his numerical examples claimed to

side with Sraffians to show how and why there can be no universal measure of “depth or duration of time-phased produced inputs” that can serve as simple apologetics for mainstream theories of interest. Unequivocal “capital deepening” just cannot be defined.

He illustrated the fact that, even when reswitching does not occur, yet in a simple Austrian framework it can be shown that an inverse trade-off between the real wage and the interest rate is always present, while the inverse relationship between the interest rate and the capital-labour ratio can be reversed any number of times. Similarly, Fratini (2010) shows that, in an
Austrian framework, a monotonically decreasing schedule of the demand for capital can co-exist with stretches in which net product per worker and the interest rate simultaneously increase. Furthermore, Fratini (2014, 2019a) also shows that recent attempts to rehabilitate the average period of production as a useful measure of the degree of roundaboutness of production, and as an index that can be supposed to have a positive relationship with the economy’s net product, continue to encounter difficulties similar to those faced by their predecessors.

The second encompasses the discussion of whether, even when the previous critiques are set aside, it can be argued that the rate of interest can turn negative when the ‘third cause’ is present. It can be shown that this is not the case. In fact, if the production process is described, as is commonplace, by means of CES production functions, even when the capital-labour ratio rises to enormous values the rate of interest can in the limit fall down to zero given that the marginal product of capital does not become negative (von Weizsäcker 2020; Serrano et al. 2020). Nevertheless, even supposing the marginal product of capital to turn negative, the demand curve for capital would not go in the negative territory insofar as this would mean that entrepreneurs rationally decide to employ a quantity of capital in production such that the last unit employed yields a negative net product (Di Bucchianico 2020b). These critiques to the NNRI could have been shared by Samuelson. In fact, Bernanke (2015) recalled in a discussion with Summers on Secular Stagnation and the role of the NNRI that

As Larry’s uncle Paul Samuelson taught me in graduate school at MIT, if the real interest rate were expected to be negative indefinitely, almost any investment is profitable. For example, at a negative (or even zero) interest rate, it would pay to level the Rocky Mountains to save even the small amount of fuel expended by trains and cars that currently must climb steep grades. It’s therefore questionable that the economy’s equilibrium real rate can really be negative for an extended period.

These discussions suggest that if the ‘third cause’ is absent (as in Krugman), this leaves the analysis of the NNRI incomplete in an important aspect. Nonetheless, if present (as in Summers), it causes problems of logical consistency and plausibility for the framework adopted by these two economists.

4.2. The Main Connections in a Map

The last part of the enquiry concerns a clarification of how we believe our suggested line of connection could be inserted in a more general picture. It must obviously be stressed that the ‘map’ in Fig. 5 cannot but serve as a mere indication, and by no means it can exhaust all the
ramifications that should be involved. In it, we will use Krugman (1998) as the key intersection of the web.

To begin with, Krugman (1998) cannot but be linked with the outstandingly famous ‘original’ Liquidity Trap of Hicks (1937), a connection he himself states at the very beginning of his article, where Hicks, the IS-LM model, and the liquidity trap concept are all mentioned. This leads in turn straight back to Keynes’ (1936) General Theory, where he discussed the role of the propensity to hoard and the preference for liquidity. This line of connection has been thoroughly analysed by Boianovsky (2004), who reconstructs the history of the Liquidity Trap from the Keynesian/Hicksian origins onwards. Other authors, such as Kregel (2003) and Taylor (2014), criticised Krugman’s Liquidity Trap reappraisal for having replaced in it the concept of liquidity preference with that of rational expectations, and for relying on a natural rate of interest to bring the economy in equilibrium.\textsuperscript{17} In this line of enquiry we obviously find also the works of, among others, Modigliani, Klein, and Tobin.

As already mentioned, Krugman (2013) explicitly links his contribution to that of Summers (2014) on Secular Stagnation, and the linkage is apparent when looking at their analytical structures (Di Bucchianico 2020a). Summers, in turn, borrowed many intuitions from Hansen (1939) ‘original’ hypothesis. Also in this instance, the long road from the original concept to the modern reappraisal has been reconstructed (Backhouse and Boianovsky 2016a, 2016b), and could encompass additional authors such, among others, Hobson, Steindl, and Sweezy.

To this already extensive picture we want to add the line connecting Krugman (1998) to Samuelson (1958) and going back to Böhm-Bawerk (1889 [1930]). Also this line of thought cannot be reduced to a straight avenue and, as seen, authors such as Wicksell and Schumpeter have an important role in it. And, to conclude, the suggested map can also be integrated by cross-linkages among the authors: for example, the idea of an NNRI can be traced already in the work of Wicksell (Boianovsky 2016). Again, Summers’s use of an NNRI can be also associated with that of Klein (1947), who was elaborating on Pigou (1943) (Backhouse and Boianovsky 2016a, pp. 948-949). In turn, Pigou (1943) deemed Hansen’s theory to rest on an NNRI and while developing the argument he singled out the fact that in a stationary state the interest rate is determined by the discount factor (ibid., pp. 952-953).\textsuperscript{18} This makes room for reintroducing also in this instance the ideas of Böhm-Bawerk, as the case supposed by Pigou was also the starting point of Krugman (1998) model (equation 10).
5. Conclusions

The negative natural rate of interest is a theoretical novelty that has continuously recurred in macroeconomic discussions since more than two decades. Indeed, it is present and plays a major role in the modern Liquidity Trap and Secular Stagnation theories. Moreover, it heavily informs the theoretical background of central banks that have progressively resorted to unconventional monetary policies and are now also implementing negative interest rate policies.

In this paper we tried to offer a contribution that can in our opinion shade a new light on the historical origins of the concept. After a relatively long time, the negative natural rate of interest re-emerged in Krugman’s (1998) modern Liquidity Trap model and it is now at the hearth of the Secular Stagnation theory. This contribution can be read in a line of thought in which Samuelson (1958) and Böhm-Bawerk (1889 [1930]) are two fundamental authors. Indeed, when dealing with the establishment of what is now labelled as the overlapping generations model, Samuelson derived a negative natural rate of interest by grounding on a set of assumptions which we have also found in Krugman’s seminal work. The analytical exercise of Samuelson is in turn directly rooted in either ignoring or reverting Böhm-Bawerk’s three causes for a positive rate of interest.

We have then pointed out the fact that, first, Krugman’s (1998) contribution, despite his brilliant intuitions, is not a solid improvement over the theoretical scenario depicted by Samuelson (1958). In fact, Krugman did not engage in adding to that basic scenario the
refinements that Samuelson suggested in his seminal work, with the exception of introducing a systematic preference for consumption today.

Second, there is also a deeper problem in the modern treatment of a negative natural rate of interest in recent models of Liquidity Trap and Secular Stagnation. In fact, it appears that the fundamental result of an NNRI can be obtained when the reasoning ignores Böhm-Bawerk’s ‘third cause’. Therefore, more recent attempts, such as that of Summers (2014), to reintroduce capital in the discourse should in principle be welcomed as a noticeable improvement over its Liquidity Trap predecessor. However, the moment they do so, they seem to jeopardise the possibility for their analytical framework to consistently host a negative natural rate of interest.

In conclusion, given the great deal of attention that the topics of the negative natural rate of interest, Krugman’s Liquidity Trap and Summers’ Secular Stagnation theories (rightfully) garnered over the years, we claim a more comprehensive understanding of the origins of that concept to be of utmost importance. We also claim that the hitherto unnoticed line of connection we have proposed can help to shine a light on some hidden subjects which have been the target of long-lasting debates and are still at play, albeit seldom acknowledged, in modern theories.

References


I wish to thank the participants in the 17th Annual STOREP Online Conference, 3-4 October 2020 and the participants in the ESHET Online Young Scholars Seminar, 24 September 2020 for stimulating comments. Federica Cappelli helped me throughout. Any errors are solely my own.

In the latter contribution the similarity with Krugman (1998) in the analytical reasoning leading to the liquidity trap is neat.

Borrowing Blaug’s (1999) dichotomy, our enquiry is akin to a ‘rational’ reconstruction.

The following case-study with land introduces an overlapping, two-generation structure. The price of the only product of the economy is supposed to decrease tomorrow. Again, agents systematically prefer consuming today rather than tomorrow. In this case, there is an attempt to deal with a productive asset such as land, delivering a positive marginal product.

For a comprehensive analysis, see Fillieule (2015).

We report the quote by altering the order in which Samuelson recalled the causes to keep it consistent throughout.

Wicksell’s attempt has been also interpreted as a case in which the focus on the production side led to ignore the consumption side, thereby leading to the underdetermination of his system due to an alleged ‘missing equation’ (Kurz 2000). Such a thread of literature involved contributions by, among others, Hirshleifer, Sandelin, Negishi, Larry Samuelson, and Malinvaud (Fratini 2013).

Uhr (1960, pp. 115-119) also reconstructs the evolution of Wicksell’s thinking concerning those Böhm-Bawerk’s insights. So, for instance, while in Value, Capital and Rent Wicksell denied the relevance of the first cause while accepting the second and the third, in the Lectures he also started questioning the emphasis put by Böhm-Bawerk on the third cause.

Without further additions, we single out the similarity between this structure and a recent overlapping generations model used to study Secular Stagnation (Eggertsson et al. 2019).

Similar graphs meant to represent Böhm-Bawerk’s framework can also be found in Bernholz (1993, p. 24, 27), Negishi (1982, pp. 165-167). In our representation, the resource constraint line represents the first cause, while the indifference curves contain the second cause. However, the graph can be easily used to represent all three, when the resource constraint describes the third cause and its outward shift depicts the first (Samuelson 1994, pp. 204-205). This last method is more in line with Fisher’s use of the graph. Further, we here ignore the fact that the curvature of the indifference curve, not only its steepness, contributes to determine the value of the rate of interest.

In fact, when in equation (6) beta is lower than one, the decrease of future endowments has to be sufficiently strong as to make the ratio between future and present endowments lower than the discount factor.

There seems to be no clear-cut distinction between the discussion of the difference between a rise in productivity ensuing from the choice of a more capital-intensive technique from a given list of technological possibilities and the introduction of a novel process. However, the point appears of minor relevance in the present context.

As Lee (2019) recalls, Samuelson also engaged in a deeper treatment of the role of population dynamics. For a broader perspective on Samuelson’s theoretical enterprise, see Kurz (2010), Backhouse (2017), Cord et al. (2019).

The controversy did not end there, but it also experienced a second stage. On this point, see Fratini (2019b).

Note that we are also setting aside the above-mentioned role for the stores of value for the sake of exclusively focus on the decision to implement more roundabout production processes.

For other critiques of the use of a natural rate of interest in the Secular Stagnation theory, see Bertocco and Kalajzić (2018), Palley (2019), Hein (2020).

On the connection between Keynes and Krugman, see also Harcourt (2008) and Steele (2012). Fantacci and Sanfilippo (2020) illustrate also the role of Dennis Robertson in the birth and development of the Liquidity Trap concept in several exchanges entertained with Keynes.

For a comparison between Pigou’s views on the topic of stagnation and the more recent theories, see Di Matteo (2020).