

Volume 1

Number 1

NEW PERSPECTIVES ON POLITICAL ECONOMY

ARTICLES

- Cwik: The Inverted Yield Curve and the Economic Downturn
- Ylönen: Institutions and Instability in Africa: Nigeria, Sudan,
and Reflections from Mises's Nation, State and Economy
- Michel: End Of The Warriors
- Koderová: Monetární teorie v českém ekonomickém myšlení 90. let
20. století

2005

New Perspectives on Political Economy
Volume 1, Number 1, 2005, pp. 1 – 37

The Inverted Yield Curve and the Economic Downturn

Paul F. Cwik, Ph. D.*

JEL Classification: B53 – Austrian Economics, E22 – Capital and Investment, E32 – Business Cycles, G19 – Yield Curves

Abstract: This paper presents an answer to why the yield curve tends to invert one year before a recession. The capital-based macroeconomic model used in this paper traces out the effects of an injection of short-term working capital into the model. There are two consequences of this injection: the Wicksell effect and the Fisher effect. The Wicksell effect entails the downward pressure on interest rates, while the Fisher effect entails the upward pressure on interest rates. The short-term credit can create both short- and long-term malinvestments in the social structure of production. These malinvestments are unsustainable and must be liquidated. The process of liquidation phase may take the form of a credit crunch, a real resource crunch, or a combination of the two. Each scenario culminates in an inverted yield curve approximately one year before the upper-turning point of a recession.

* Adjunct Scholar of the Foundation for Economic Education and of the Ludwig von Mises Institute

1 Overview

This paper addresses the question of why the yield curve tends to invert before a recession. It does not create a model to demonstrate that such a phenomenon exists, as this relationship has already been well established. This paper uses the capital-based macroeconomic approach set forth by Garrison (2001) to explain that a correlation exists between the yield curve's spread and real output. Accordingly, the topic is examined by disaggregating investment and capital-formation decisions. The capital-based approach of macroeconomic theory is well-suited for the examination of this question, since it is a theory of the upper-turning point of a business cycle.¹

Macroeconomic theories attribute economic downturns to either monetary or real factors. The capital-based approach allows for both. A disaggregated approach allows for analysis and insights that other theories cannot provide. Unlike the capital-based approach, most macroeconomic theories that examine the upper-turning point focus on the immediate causes of the downturn. They do not include the underlying capital structure as a part of the theory, because this structure is viewed as an unnecessary complication to the theory. By ruling out capital (and the malinvestments that could be built up during the expansionary phase), the leading macroeconomic theories focus on more aggregated causes—such as monetary or real shocks to the economy.

These models are too aggregated to properly answer the question of why the yield curve tends to invert before a recession. Prior to the 1990-91 recession, several economists called attention to the past performance of the spread as a predictor of a business cycle's upper-turning point. However many dismissed the signal, declaring it may be a false positive.² Another had argued that there would not be a recession in 1989 or 1990 because there was “an absence of the kind of gross imbalances in the economy that have typically preceded past recessions.”³ It is possible in retrospect to see that the imbalances were in the economy and were liquidated in the 1990-91 recession.⁴ The current approach, how-

¹ See Hayek (1969) where he states, “This theory [the Austrian Business Cycle Theory or ABCT] never claimed to do more than account for the upper turning point of the typical nineteenth-century business cycle.” p. 282.

² See Brown and Goodman (1991) and Estrella and Mishkin (1998).

³ See Keen (1989) p. 40.

⁴ See, for example, Hughes (1997) for an empirical analysis supporting this claim.

ever, lacks the ability to see the imbalances (malinvestments) created during the “boom” phase. It is here that Austrian theory can provide insight. This approach allows one to analyze and draw conclusions about the problem in a manner that is superior to more aggregated methods.

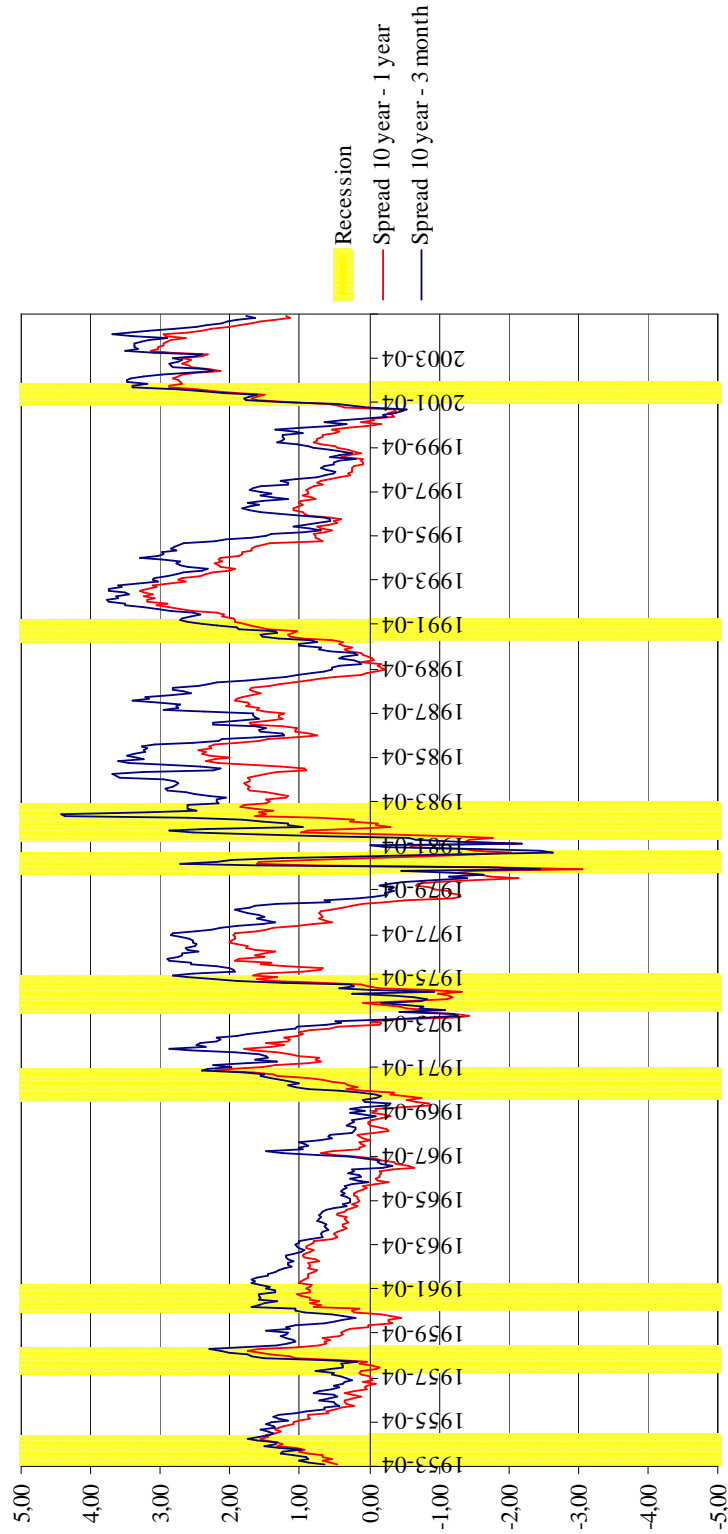
The remainder of this paper is divided into six sections. The next section reviews the empirical relationship and surveys the current models that attempt to explain the relationship. Section 3 presents the implication of monetary expansion. Section 4 continues the analysis by showing that monetary injections lead to a malinvestment boom. Section 5 establishes how such malinvestments are not sustainable and inevitably lead to the “crunch phase” of the business cycle and then demonstrates why the yield curve inverts during the crunch phase (prior to the upper-turning point). Section 6 summarizes and concludes.

2 Presentation of the Relationship and Current Research

Economists, government officials, and businessmen have long searched for accurate business cycle indicators. One strong predictor of the upper-turning point of a business cycle is the inverted yield curve. Chart 1 illustrates the 10-year Treasury Bond and the 1-year T-Bill spread and the 10-year T-Bond and the 3-month T-Bill spread between April 1953 and October 2003.⁵ (Please see Chart 1 below.)

⁵ The NBER dating of the recessions is used. All data for this paper were obtained from the Federal Reserve Bank of St. Louis' FRED II.

Chart 1: Yield Curve Spreads Between 1953-2005



Recessions are dated according to the NBER.
The data for interest rates were obtained from FRED II.

An inverted or humped yield curve has occurred no more than 5 quarters before every recession since the mid-1950s. Except for the Q3:1990-Q1:1991 recession, the yield curve has inverted in every recession since the mid-1960s. Prior to the Q3:1957-Q2:1958, Q2:1960-Q1:1961 and Q3:1990-Q1:1991 recessions, the 10-year/3-month spread did not become negative. The lowest points for this spread were 0.24% in February 1957, 0.20% in December 1959 and 0.13% in August 1989. Preceding these recessions, the yield curve was technically humped and not inverted. The 10-year/1-year spread was negative in December 1956 and from February through April 1957 and, according to McCulloch (1990), the 15-year/6-month spread (not shown in the chart) was negative from November 1956 through March 1957. The 10-year/1-year spread was also negative in the period of September 1959 through February 1960 and February through September 1989.

There is one instance where an inverted yield curve was not followed by a recession. From September 1966 through January 1967 the yield curve inverted, but no recession took place. Some refer to this occurrence as a false positive, but the second quarter of 1967 did experience a negative growth rate of -0.06% (real GDP). While this decline in real output did not constitute an official recession, it does confirm the relationship under study.⁶

The historical record does not show this connection to be only a post-WWII phenomenon. The yield curve inverted between June 1920 and March 1921 and again between January 1928 and November 1929.⁷ Data from the 19th century are incomplete and do not easily lend themselves to analysis.⁸ Nevertheless, support for the thesis of the yield curve as a predictor of business cycles can be traced as far back as the mid-1800s.⁹

The current research can be separated into two basic models: the consumption-based capital asset pricing model (CCAPM) and the Estrella models. The essential idea of the CCAPM is that investor's smooth income across business cycles. Criticism of the CCAPM has led to the development of an alternative theoretical model.

⁶ The true exception to this relationship is the Q2:1953-Q2:1954 recession, where the yield curve flattened but did not invert.

⁷ Cecchetti (1987) demonstrates that the observed bond market data from the 1930s and 1940s is distorted due to heavy government intervention. As an illustration of the degree of distortion, on December 31, 1932, institutional forces caused the 3½% US Liberty Bond's yield to fall to a nominal rate of -1.74%.

⁸ Davis (1971) examined U.S. capital markets from 1820 through 1930. He shows these markets were not integrated until after World War I.

⁹ See Keen (1989).

The second type of model builds on variations of the following economic tools: the Expectations Hypothesis, the Phillips curve, the IS curve, a monetary reaction function, and the Fisher Equation. This category of models is motivated by a monetary shock. The origin of these models is Estrella (1998). The Estrella model is also inadequate for understanding *why* the yield curve tends to invert before a recession.¹⁰ The Estrella models derive the relation between interest rates and real output from the Phillips curve and the IS curve. The Phillips curve (an empirical and not a theoretical relationship) fails to explain the connection between interest rates and real output. It only shows that a connection exists between nominal rates and output. However, the theoretical underpinnings needed to understand the relationship are not explained. Furthermore, the use of the Keynesian IS curve is insufficient to create a credible model.¹¹

Over the past 15 years, the debate on theory remains unresolved. No article during this period examines the effects of non-neutral monetary injections through a heterogeneous capital structure. The use of Austrian insights can provide an alternate (and fruitful) perspective to this debate.

3 Monetary Expansion

The capital-based approach posits that the initial disequilibrium of the business cycle is caused by monetary injections.¹² Credit is injected at the short-end of the yield curve and spreads through the economy causing non-neutral effects. The effect of the new credit is separated into the Wicksell effect and Fisher effect. These opposing effects distort the ability of price signals to transmit relative scarcities to entrepreneurs. As a result, monetary expansion lowers and alters interest rates that falsely signal entrepreneurs to embark upon malinvestments.

The effects of monetary expansion are traced through the yield curve, which was developed in Cwik (2004), Chapter 3. The modified Preferred-Habitat Theory of the yield curve is a combination of time-preference (in the Böhm-Bawerkian sense), expectations,

¹⁰ See Cwik (2004) Chapter 2.

¹¹ Ibid.

¹² Many factors can cause an economic downturn—war, sweeping changes in institutions, radical changes in monetary policy, etc., but these are outside of the scope of this topic. This paper specifically focuses on downturns caused by monetary injections.

liquidity-preference, and risk aversion (the preference for matching debt and equity).

Böhm-Bawerk's analysis is the basis for the formation of interest rates, since it satisfies both the essentialist and functionalist questions regarding interest. After an initial interest rate is established, a yield curve can be derived by adding expectations, liquidity-preference, and risk aversion to the analysis. As shown in Cwik (2004), the modified Preferred-Habitat Theory is consistent with the empirical observations of the yield curve.

When the monetary authority engages in a policy of monetary expansion, the new money is injected into the monetary system at specific points.¹³ The effect of additional liquidity is sometimes called the Wicksell effect.¹⁴ The Fisher effect is the change in interest rates caused by changes in the expectations of future inflation.¹⁵ The Wicksell effect and the Fisher effect are opposing forces. The Wicksell effect tends to lower interest rates while the Fisher effect tends to raise them.

With a policy of monetary expansion, the Wicksell effect first dominates interest rate movements. As money is injected into the short end of the yield curve (through the monetary base and thus the Fed funds rate) an initial lowering of short rates and a steepening of the slope of the yield curve results. Keeler (2002) states,

The liquidity effect of a monetary shock will lower interest rates in general and lower short-term rates relative to long-term rates. The yield curve will shift down and become steeper in slope. . . .¹⁶

Although Keeler is correct about the steepening of the yield curve, empirical observation does not support the tendency of the yield curve to shift, as long rates tend to remain stable relative to short rates. Bernanke and Blinder (1992) argue that the short rates move

¹³ The Federal Open Market Committee typically adjusts monetary policy through the use of open market operations and the use of the discount rate which change the aggregate level of depository institutions' reserves. Changes in these reserves induce changes to the Fed funds rate. The Fed funds rate is the interbank interest rate for short-term loans, usually overnight. See Miller and VanHoose (2001).

¹⁴ The phrase "Wicksell Effect" was first used in the Cambridge Capital Controversy of the 1960s. The phrase was divided into a "Real Wicksell Effect" and a "Price Wicksell Effect," describing the change in the relationship between the rate of profit and capital intensity in real or value terms. The phrase "Wicksell Effect" used in this paper refers to an "Interest Wicksell Effect" (or a liquidity effect) where money added to an economic system, by expanding the supply of investable funds, initially reduces the market rate of interest.

¹⁵ See Mishkin (2001) pp. 107-108.

¹⁶ Keeler (2002) p. 5. See also Keeler (2001) pp. 333 and 335.

while the long rates remain stable. The Fisher effect increases the forward short rates, thus applying upward pressure to long rates. However, the new money is arbitrated across the term structure. The Wicksell effect prevents the long rates from rising. Thus the yield curve rotates instead of shifting, as shown in Figure 1. The new yield curve is presented as the dashed curve.

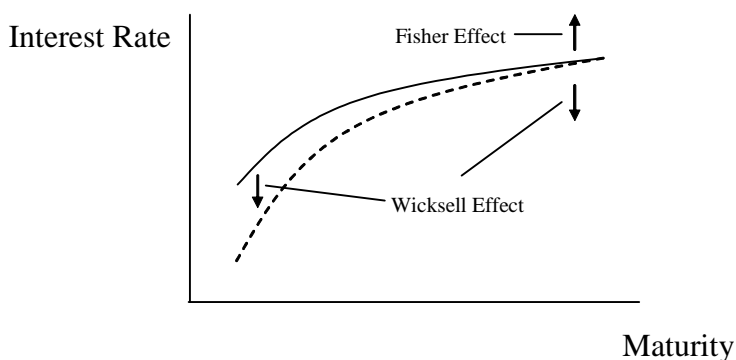


Figure 1: The Wicksell and Fisher Effects Combined

According to Keeler (2002), the steepening of the yield curve begins during the expansion phase of the business cycle. However, the yield curve is steepest at the lower turning point of the business cycle.¹⁷ These observations are not inconsistent with Keeler's observations if the recovery phase of the business cycle is also included as part of the expansionary phase of the next cycle.

In sum, the analysis begins with the Böhm-Bawerkian framework to establish an initial interest rate. Expectations, liquidity-preference and risk aversion are added to create a modified Preferred Habitat theory of the yield curve. With monetary expansion, the Wicksell effect shows the lowering of short rates due to an increase in the supply of investable funds. The long rates tend to remain stable relative to the short rates due to the interaction of the Wicksell and Fisher effects.

Since new money does not enter an economic system uniformly or at a steady rate, the already difficult job of entrepreneurs, to correctly read market signals, becomes even

¹⁷ This empirical observation was made as early as Kessel (1965) and has since been seen as a consistent pattern of the yield curve.

more difficult.¹⁸ Entrepreneurs need to correctly read these market signals to make profits, and as a consequence, coordinate the economy.¹⁹ Since the price changes resulting from an increase in the money supply are not uniform, entrepreneurs have difficulty determining whether the price change is a result of a change in relative scarcity or whether the price change is the result of inflationary pressures. In other words, they are not able to distinguish between relative price changes and inflationary price changes. As a result, the economy becomes more wasteful and less efficient.

This paper posits that the monetary authority injects money into the economy through short-term credit markets. The addition of credit lowers short rates, while the Fisher effect should increase long rates. However, the yield curve steepens and does not shift. The Wicksell effect is transmitted across the term structure of interest through the process of arbitrage and reduces the Fisher effect on long rates. Thus, monetary injections artificially lower interest rates across the entire yield curve. These false rates signal to entrepreneurs that consumers have shortened their time-preferences, leading to a malinvestment boom.

4 Malinvestment “Boom”

In the previous section, the idea of malinvestments was introduced. This section examines the nature of these malinvestments in the context of a capital-based macroeconomic approach. The crisis stems from the need to liquidate the malinvestments that are built during the boom. During this crisis, which is the upper-turning point of a business cycle, the yield curve inverts as a consequence of the liquidation process.

To clarify terminology, a distinction must be made between an individual project's period of production and the social period of production. Schmitz (2003) distinguishes between the individual and social periods of production. The individual period of production corresponds to the length of time that an entrepreneur's project will take until it yields output for the next stage of production. These projects are distributed throughout the structure of production. The social period of production corresponds to the degree of

¹⁸ Prices are packets of information that signal to entrepreneurs the relative scarcities of the various goods and services throughout the economy. See Hayek (1945).

¹⁹ See Mises (1980).

complexity of an economy. In other words, the overall degree of economy-wide roundaboutness is the social period of production. Individual projects' are divided into long and short terms and correspond to the long and short rates of the yield curve.²⁰ Both long and short-term projects are found at every stage in the structure of production.

Furthermore, capital is divided into two forms: working capital and fixed capital. Working capital, also known as circulating capital, refers to the funds that flow through the structure of production. Fixed capital is the capital equipment, buildings, machines, etc. that do not flow through the structure of production. Instead, fixed capital is embedded at the different stages within the structure of production. Through the investment process, working capital is used to purchase inputs such as labor and goods-in-process; additionally working capital is also transformed into fixed capital.

Expansion of Short-Term Working Capital

This paper seeks to demonstrate that monetary injections (in the form of working capital) into an economic system necessarily lead to an inverted yield curve prior to an economic downturn. As a result, this paper assumes the extreme case where monetary expansion initially takes the form of working capital in the short-term credit markets.²¹ As this assumption is relaxed, the argument is strengthened.

As previously demonstrated, the monetary injections during an economic boom (and also during a recovery) cause the yield curve to steepen.²² Short rates fall, while long rates tend to remain relatively stable. With a monetary injection at the short end of the yield curve, the modified Preferred-Habitat theory suggests that the yield curve should shift down instead of rotate. However, the empirical evidence shows that yield curve rotates and steepens, but shifts very little. This seeming inconsistency with the theory is due to the Wicksell effect explored in the previous section. The impact of the Wicksell effect on long-term and early stage malinvestments is discussed below.

²⁰ A long-term project may be financed through a series of short-term loans. A simplifying assumption of a hard link between the length of the project and the duration of the loan will not change the analysis.

²¹ This paper is additionally assuming that the traditional role of thrift institutions of "transforming maturities," by borrowing short and lending long, does not take place.

²² Cwik (2004), Chapter 3, section 3.6 provides empirical evidence that expansionary monetary policy is the typical case during the boom and the recovery phases of the business cycle.

The monetary injections, which rotate the yield curve, send opposite signals to entrepreneurs and consumers. As new short-term working capital is injected into the economy, the economy embarks upon a malinvestment “boom.” As the cost of borrowing decreases, the marginal borrowers (those previously excluded from the market) will now be able to obtain the wherewithal to fund their individual projects. Real resources are transferred to these borrowers and the working capital is converted into fixed capital as distinct production processes are added to the economy. Machlup (1932) illustrates this process,

The fresh borrowers employ the fresh capital—either for a new enterprise or for the expansion of an already existing one—by demanding means of production, partly original factors of production, partly intermediate goods. This increased demand will raise the price of production goods. Therefore the borrowers who are in the best position to compete are those who are less affected by the increased cost of intermediate goods than by the lowering of the rate of interest. This is *not* the case with investment in raw materials and goods in process, but it *is* the case with investments in *fixed* capital since in calculating the prospects of such investments the interest rate is of much greater importance than the price of the goods used.²³ (italics in the original)

After debating with Haberler, Machlup demonstrates “that the investment of fresh capital for an increase of production and output which might be technically possible without expanding fixed capital, is economically impossible.”²⁴ Machlup’s point is that in order to achieve an expansion of output, working capital must be transformed into fixed capital. In a later work, Machlup (1935) reinforces his conclusion that a decrease in interest rates leads to the formation of new investment in fixed capital.²⁵ While Machlup argues that the short-term funds will eventually be transformed into longer-term fixed capital, at this part of the analysis, the point to be emphasized is that working capital is transformed

²³ Machlup (1932) pp. 276-277. With expansionary monetary policy and an increase in output, Machlup concludes that, even with additional short-term funds, “the short-term use of capital is theoretically impossible.” p. 277.

²⁴ Machlup (1932) pp. 277.

²⁵ Machlup (1935) states, “As a cost factor, the interest rate has real significance only as it applies to *new investment in fixed equipment*.” p. 462. (italics in the original) “A decrease in the interest rate changes the comparative cost-calculation in favor of those methods of production which make the heavier demands on capital.” p. 462.

into fixed capital. These short-term projects are malinvestments and must be liquidated at a future date unless additional real savings are supplied.

In the meantime, the short-term projects in the late-stages of production (those stages closest to the consumers) are justified through increased profitability due to the increase in the demand for consumer goods. With a decline in short rates, the cost of financing short-term consumer purchases (through the use of credit instruments such as credit cards) falls. Thus, an immediate result of the monetary injection is an increase in the demand for consumer goods. The effect is seen in Figure 2.

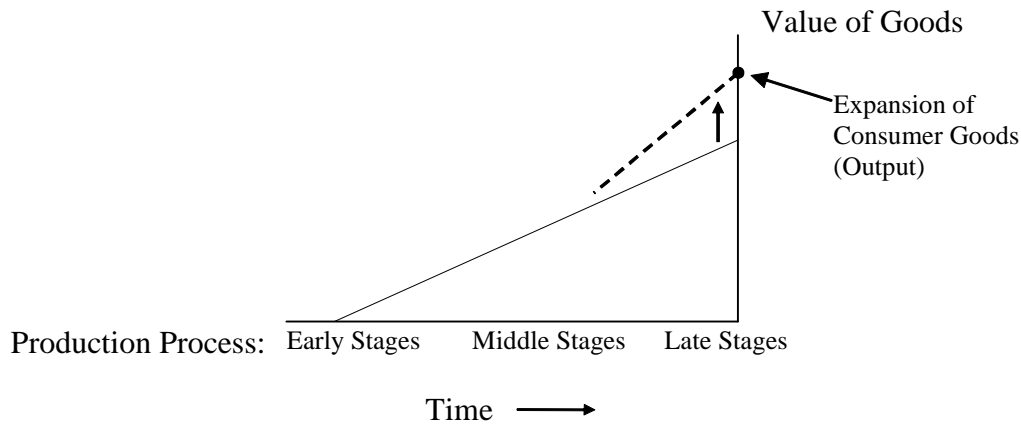


Figure 2: The First Effect from Additional Working Capital

The dashed line in Figure 2 represents the additional projects accumulating toward the late stages of production.

By restricting monetary injections to the addition of working capital in the short-term credit markets, the analysis leads to the conclusion that short-term projects at the late stages of production are built up. Machlup demonstrates that, over time, the working capital will be transformed into fixed capital. The fixed capital is combined with additional inputs to create consumer goods. These consumer goods are purchased with the new credit extended to consumers.

Long-Term Malinvestment

In this analysis, monetary injections have been restricted to the form of additions of working capital to the short end of the yield curve. As this working capital is transformed into distinct production processes to supply consumer goods, a portion of the working capital is applied to the purchase of inputs while the remainder is transformed into fixed capital. While the analysis has focused on an expansion of the late-stages of production, empirical observation of a boom is that early-stage markets experience larger swings relative to intermediate and consumer goods markets.²⁶ Such empirical evidence begs two particular questions: “If short-term rates fall relative to long rates thus increasing the amount of short-term projects, where is the long-term malinvestment?” and “Where is the early-stage malinvestment?”²⁷ To find answers the analysis will use Bastiat’s “unseen.”

When the monetary authority engages in monetary expansion, entrepreneurs increase their expectation of future inflation. Within the model of the modified Preferred-Habitat theory, the middle and long rates should rise in accordance with the Fisher effect, but empirically they do not. As previously observed, short rates fall relative to the long rates and the long rates tend to remain stable.

Nevertheless, long-term malinvestments emerge from the injection of short-term working capital. The arbitrage process from the shorter rates prevents the long rates from rising. In other words, credit is flooding into the long bond markets, keeping their yield from rising. The Wicksell effect counters the Fisher effect. The “unseen fact” is that many long-term projects would have been curtailed with an increase in long-term rates, but the Wicksell effect discourages their liquidation. These non-liquidated projects are now in a state of disequilibrium. They, too, are malinvestments. The degree to which the Wicksell effect inhibits long-term rates from rising corresponds with the extent to which of long-term malinvestment.

Machlup (1935) presents the rate of interest as a cost and capitalization factor in the production process. As interest rates (the short-term rates in particular) decline, the capi-

²⁶ See Skousen (1990) pp. 303-305.

²⁷ Long-term malinvestment and early-stage malinvestment are not the same. Long-term malinvestment refers to individual projects with a long-term planning horizon. Such projects may be found at any stage in the structure of production. Early-stage malinvestment refers to the projects at the higher-order stages of production. They may employ both long and short-term individual projects.

tal values of all fixed capital increases. The capitalization effect yields a greater return for the longer-lived capital equipment.²⁸ Such a change in the relative value of long-lived fixed capital encourages the expansion of long-term investments despite the increase in only short-term working capital. Thus, not only are some projects that should be liquidated not discontinued, but new long-term projects are started. As the newly expanded short-term malinvestments are added to the long-term malinvestments, the economy moves beyond the production possibility frontier curve—it has an unsustainable capital structure.

The argument so far is that the monetary authority has injected working capital into short-term credit markets. The addition of credit lowers the short rates, the yield curve steepens, and the arbitrage process prevents increases in the long-term rates; the Wicksell effect counters the Fisher effect. Short-term projects (malinvestments) are commenced due to the lowering of the short rates. These short-term projects embed various degrees of fixed capital into the structure of production. Up to this point of the analysis, the short-term projects have been located toward the late-stages of production, but the addition of short-term projects is not necessarily an addition of projects to a particular stage. Short-term projects may be added throughout the structure of production.

Simultaneously, the rotation and steepening of the yield curve is evidence that long rates are artificially held down. The prevention of increasing long rates delays the liquidation of some long-term projects, and new long-term projects are started because of the relative change in the return of long-lived fixed capital. These long-term projects are not supported by a foundation of real savings and will need to be liquidated at some future date. These long-term projects are malinvestments. Thus malinvestments (in both long and short-terms) emerge throughout the structure of production.

Despite demonstrating that long- and short-term malinvestments arise from monetary injections, one may be led to ask the second question posed above, “Where is the early-stage malinvestment?” While a hard link between long-term investments and the early-stages in the social structure of production cannot be technically proven, it is reasonable to assume that such an association exists. The case can be argued that individual long-term investment projects necessarily affect the social period of production to the

²⁸ See Machlup (1935) p. 465.

extent these long-term projects prompt capital lengthening. Nevertheless, to avoid using a conjecture that cannot be proven, this paper adopts the extreme case that long- and short-term projects may occur at any stage in the production process. To the extent that long-term investments lengthen the social period of production, the overall argument of this paper is bolstered.

Early-Stage Malinvestment

Despite all of the stringent assumptions about the monetary authority injecting working capital into the short end of the yield curve and that long and short-term projects may occur at any stage in the structure of production, a malinvestment boom in the early stages can be demonstrated.

In his seminal paper, Hayek (1945) demonstrates that entrepreneurs have only price signals to guide them to meet consumers' demands and make profits. Prices are information packets that not only signal to entrepreneurs the quantity and quality of the goods they are to produce, but also allow entrepreneurs to calculate which types of inputs and production processes are the most efficient. It is in this manner that the economy is coordinated. A network of prices ties the structure of production together. For a single interest rate model in equilibrium, the rate of interest equals the rate of profit. When the model is expanded to include a term structure of interest rates, the same principle applies, but the rate of profit for each individual project corresponds to the matching instrument in the yield curve. For example, a two-year project's rate of return should correspond to the yield of a two-year bond.^{29, 30} Thus when the rates across the yield curve fall (or are held down by the Wicksell effect), the cumulative effect is a change in the social period of production. The decrease in the short-term interest rates and the artificially low long rates signal to the entrepreneurs that the normal rate of economic profit has been

²⁹ The modified Preferred-Habitat Theory is able to accommodate the segmentation of the yield curve.

³⁰ While long-term projects may be financed through a series of short-term loans, the entrepreneur will use the maturity that matches the individual project as the relevant yield. With a positively sloped yield curve, the yield of rolling over short-term instruments is below that of the longer maturity instrument. However, the entrepreneur will not view the short-term instruments as a relevant substitute for the project. Instead, if he is looking to engage in a long-term project, he will look to the yield of the longer-term instrument as the opportunity cost of such an investment. For example, if the yield of a one-year bond is 4% and the yield of a two-year bond is 5%, the entrepreneur will regard the 5% yield as the opportunity cost of embarking upon a two-year project.

lowered.

To illustrate this process, the analysis begins with a single interest rate model. Suppose that all individual projects have a length of 3 months and there is a corresponding single interest rate for 3-month instruments. The monetary injections falsify the price signals to the entrepreneurs.³¹ The effect of the additional credit lowers the interest rate and also lowers the normal rate of return for these projects. In other words, the opportunity cost of each project is lowered.

As the rate of interest changes, so does the rate of return necessary to obtain normal economic profits.³² As the monetary injections lower the interest rate, two effects emerge. The first effect is that consumers, experiencing a decrease in their return on their savings, shift their wealth into consumer goods. Garrison (2001) refers to this situation as over-consumption.³³ As consumers dedicate less resources to their savings, retailers face an increase in the consumers' demand curves in their markets. As a result, retailers increase their demand of wholesaler goods in order to take advantage of the profit opportunities. The cumulative effect of the entrepreneurs' actions at the late-stages of production is to reduce the degree of roundaboutness in the economy. The effect is illustrated in Figure 2 above.

The second effect is that the injected money is lent simultaneously to entrepreneurs, thereby increasing the amount of investment throughout the structure of production. In this phase, the amount of investment is no longer equal to the amount of savings. A "tug-of-war," to use Garrison's phrase, arises between saving and investment. Garrison argues that "the conflict is resolved initially in favor of investment spending—because the investment community has more to pull with, namely, the new money that was lent into existence at an attractive rate of interest."³⁴ As a result of this conflict, the economy is pulled in the direction of more roundabout production processes by the investors and consumers pull the economy in the direction of less roundabout production processes.

³¹ The importance of Hayek's observation is that entrepreneurs only have price signals to guide them in their conduct. Evans (1987) argues that a "mis-assessment of risk" by investors can occur when the "true" risk structure of the economy is uncertain. Since the true risk structure is never precisely known, entrepreneurial error can occur even under the best conditions. When the monetary authorities manipulate price and interest rate signal, these errors are intensified.

³² The following analysis follows the analysis of Cwik (1998).

³³ See Garrison (2001) pp. 69-70.

³⁴ *Ibid.* p. 71.

A “dueling production structure”³⁵ emerges. Figure 3 illustrates this result.

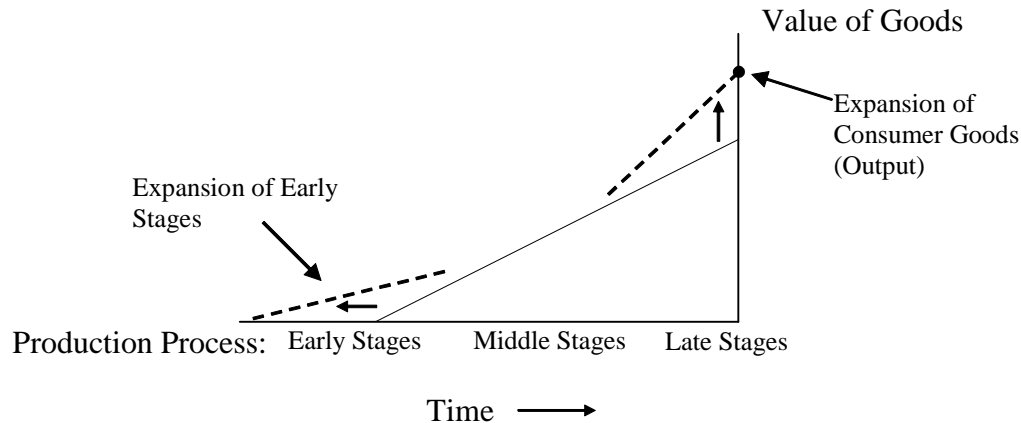


Figure 3: The Dueling Production Structure

Garrison is able to achieve this result by arguing that lower interest rates make longer-term investments more profitable. To the extent that longer-term investments are capital deepening, this result is correct. However, there is another, more fundamental reason why the economy becomes more roundabout: as the normal rate of profit falls, the effect of the decrease in the interest rate is compounded through the structure of production (via relative price changes) and yields its largest impact at the earliest stages of production.

With an increase in investable funds, the normal rate of profit for all businesses decreases. As firms react to compete for the new profits in the late-stages of production, they bid up input prices until they establish this new rate of profit. The cumulative effect of bidding up input prices creates windfall returns for the firms in the early stages. These economic profits attract new investment into the early stages creating the dueling production structure.

To illustrate this idea, suppose that there is a simple production process through which there are four stages. Each stage is 3-months long and the initial rate of interest is 10%. As a starting point, assume that the initial price of inputs is \$100. Under equilibrium, each stage meets the normal rate of profit of 10%. Using the discounted

³⁵ This phrase was coined by John Cochran (2001) p. 19.

present value formula, the price of the final output (one year later) is \$146.10. Figure 4 demonstrates this relationship.

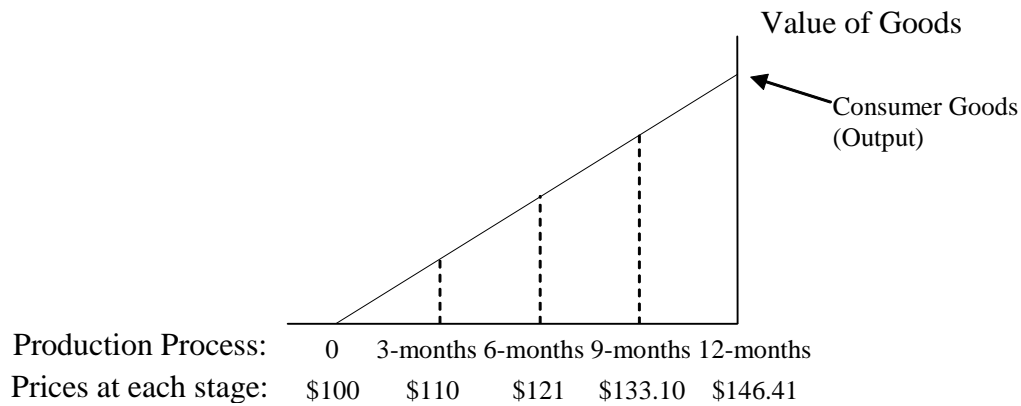


Figure 4: Prices for Each Stage Production

Holding, for the moment, the price of the final output constant, when the rate of interest falls to 8%, the price of 9-month goods will be bid up to \$135.56, thus yielding an economic profit of 1.85% to those firms operating at the wholesale stage of the production process. Table 1 shows the economic profit is highest for the owners of natural resources.

Table 1: The Effects of Relative Price Changes on Economic Profit

| Stage: | 0-months (Natural Resource Owner) | 3-months | 6-months | 9-months | 12-months (Final Output) |
|-----------------------------------|--|----------|----------|----------|--------------------------------|
| Prices with 10% interest rate. | \$100 | \$110 | \$121 | \$133.10 | \$146.41 |
| Prices with 8% interest rate. | \$107.61 | \$116.22 | \$125.52 | \$135.56 | \$146.41 |
| Rate of Economic Profit | 7.61% | 5.66% | 3.74% | 1.85% | 0% |

The rate of economic profit is expressed in Equation 1.

$$\text{Rate of Economic Profit} = \left(\frac{(1 + r_1)^{t_n - t_x}}{(1 + r_2)^{t_n - t_x}} - 1 \right) * 100. \quad \text{Equation 1}$$

where r_1 is the initial interest rate,
 r_2 is the interest rate after the monetary injection,
 t_n is the maximum number of stages in the production process, and
 t_x is the stage under examination.

Equation 1 shows that rate of economic profit is larger for the earlier-stages of production. This process demonstrates how the early-stage markets are able to expand while consumer market's demand curves are rising. While it may look as though the expansion of the early-stages depends upon longer-term investment (through the use of the discounted present value formula), the assumption early-stages having long-term investments is not necessary. The large swings in early-stage production processes result from changes in relative input/output prices.

The example provided above makes the unrealistic assumption of a specific number of stages that follow a precise order. In the real world, there is no method by which to determine where a firm is located in the structure of production. Additionally, there are many recursive loops in the structure of production, where a portion of a firm's output may be sold in both the consumer and early-stage markets. An example of such a product is the desktop computer. They are sold to research and development institutions and to consumers. Nevertheless, the principle illustrated above holds true when the economy is examined from the perspective of the social period of production.

When the assumption of holding the output price constant is relaxed and the output price is allowed to rise in accordance with the increased demand for consumer goods, the effect upon the level of economic profit is magnified. Furthermore, when the assumption of a single rate of interest is relaxed, the same formula and analysis can be applied and extended across the entire term structure of interest rates. The early-stage firms are able to derive economic profits from engaging in both long and short-term projects. The major difference for the long-term interest rates is that r_1 becomes the rate of interest

that would have materialized on the market if the Wicksell effect did not affect it.³⁶

Responding to the compounding effects of relative price and interest rates changes, entrepreneurs act as if the social period of production has changed and build more round-about processes. Keeler (2002) empirically demonstrates that relative prices are the key component of the propagation mechanism during the malinvestment boom phase of the business cycle. Furthermore, he establishes that investment reallocation and capacity utilization are expanded toward the early stages of production.³⁷

Additionally, Machlup (1932) points out that even short-term investment in working capital requires an array of higher-order capital (a superstructure) to support its production. Thus even if a short-term project is transformed into fully integrated fixed capital, it requires an additional array of higher-order capital to maintain its output. Machlup further argues that the effect on capital is compounded the “more distant” an individual production project is from consumers. Thus, a malinvestment boom in the early production stages occurs even when only short-term working capital is expanded.

A conclusion from this analysis is that the extent of the Wicksell effect corresponds to the degree of malinvestments. In other words, to the degree that the new credit is able to prevent an upward shift (or even cause a downward shift) in yields across maturities, one will see maintenance (and expansion) of disequilibrated capital projects. It is important to note that only a disaggregated approach can examine the capital structure in this manner. The more aggregated theories are unable to draw these conclusions. Thus the disaggregated, capital-based approach explains the malinvestment boom that Keen (1989), Brown and Goodman (1991) and Estrella and Mishkin (1998) missed prior to the 1990-1 recession.

A key aspect of the malinvestment boom is not the boom, but the malinvestment. The malinvestments maintained and created during the boom phase are *mal*investments because real savings does not support them.³⁸ As a result, they must be liquidated at a future date. These malinvestments are revealed during the crunch phase of the business

³⁶ The rate, r_2 , is still the observed rate of interest after the monetary injection has had its effect.

³⁷ See Keeler (2002) p. 15.

³⁸ Additional savings (entering the economy from abroad or through a tax cut on savings) will transform some (or maybe even all) malinvestments back into stable investments. However, additions to the money supply changes the international price of the currency and reduces the incentive of foreigners to invest in economy under examination.

cycle.

5 The Crunch Phase and an Inverted Yield Curve

As noted above, the theories of the upper-turning point of the business cycle center on either monetary or real factors as the primary cause of the downturn.³⁹ While the capital-based macroeconomic theory of the upper-turning point is not unique in describing the upper-turning point, the significance of this approach is that it is able to account for multiple factors. Robertson (1959) presents a classic observation on the phenomenon:

How is this cumulative upward process [of the economy] stopped and reversed? It seems to me unlikely that there is a single answer applicable to all occasions; there is a great variety of reasons why, in Haberler's language, the system may become more and more sensitive to "deflationary shocks" as expansion proceeds. Some interpreters have laid stress on purely monetary factors—the fact that the banks, finding their reserves slipping away through withdrawals of legal tender money to pay the enhanced wage-bills, etc., ultimately draw in their horns with a jerk. Others lay stress on the emergence of what they call a "shortage of saving," which no liberality on the part of the banks could remedy. According to this picture, windfall profits are eaten into by rising wages and interest rates, which at this stage no longer lag appreciably behind the rise in prices, and with the disappearance of windfall profits the main source of demand for instrumental goods is dried up. There turns out to be an overproduction of such goods in the sense, as Cassel puts it, of "an overestimate of the... amount of savings available for taking over the real capital produced."⁴⁰

Within the above passage, Robertson identifies two causes of the onset of a recession as either a "deflationary shock" or a shortage of savings. The capital-based approach identifies each of these factors as a potential and immediate cause of a recession, but the underlying factor in each case is the malinvestment built up during the boom phase. Monetary injections create disequilibria that cannot be maintained forever. The crunch phase of the business cycle, where the scramble to prevent the liquidation of malinvestments takes place, can come about in two ways—the credit crunch or the real resource

³⁹ See fn. 12 above.

⁴⁰ Robertson (1959) pp. 96-97. Robertson does not provide the specific cite for the Cassel quotation.

crunch. While each scenario may cause the economy to turn from boom to bust, they often occur together. However, the capital-based approach, by not over-aggregating the effects of monetary injections, shows that each of these “causes” have the same root-investments.

Since the mid-1960s, there have been six official recessions.⁴¹ Except for the 1990-1 recession, monetary policy was tightened in each instance. However, when tightening occurred after the recession started, it cannot be concluded that the recession was caused by a credit crunch.⁴² In five instances, 1966, 1969-70, 1973-5, 1981-2, and 2001, a credit crunch preceded an economic downturn. The recessions that are not preceded by a policy of credit tightening are: 1980, and 1990-1. These recessions were caused by a real resource crunch where economic pressure increased input prices which led to an economic downturn.

Credit Crunch

The credit crunch occurs when the monetary authority determines inflation (or expected inflation) is too high and “slams on the monetary brake.” The monetary authority’s actions force short-term rates to rise. The yield curve rotates instead of shifts because the rate of future inflation is expected to fall. The Wicksell effect dominates the Fisher effect at the short-end of the yield curve and they negate each other at the long-end. Thus the yield curve tends to invert prior to a recession, as seen in Figure 5.

⁴¹ In addition to the six recessions, the second quarter of 1967 experienced a negative growth rate. It was preceded by an inverted yield curve and a credit crunch in 1966.

⁴² Owens and Schreft (1995) argue that there was tight credit in 1966, 1969-70, 1973-74, first half of 1980, 1981-2, and early 1990-2. However, they state that the 1990 crunch was market induced, while the others stem from actual policies of monetary tightening or threats of increased regulatory oversight.

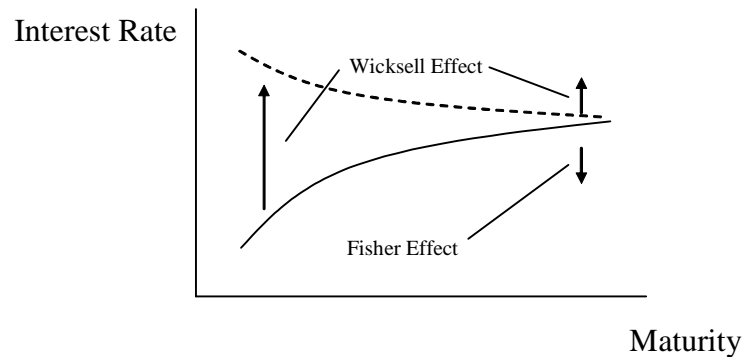


Figure 5: Inverting the Yield Curve with the Wicksell Effect and the Fisher Effect

Hayek (1969) states that in order to maintain the level of malinvestment, the rate of money supply increases must be accelerated even when expectations of future prices remain constant. If there is an expectation of future inflation, the rate of monetary expansion must also outpace expectations of inflation. During periods of increasing price levels, expectations of future inflation are not constant. In the neo-classical model of the Long-Run/Short-Run Phillips curves, the economy is on a point to the left of the Long-Run curve. Such a point is inherently unstable and the only manner in which the economy can maintain its level of output is through accelerating rates of inflation.

With monetary expansion, price levels increase for two reasons: the previous expansions of the money supply drive prices higher in an uneven manner and the instability of the malinvestments induces entrepreneurs to bid up input prices. Malinvestments are projects for which there is not enough savings to support them. During the monetary expansion phase of the boom, new investments are encouraged and consumers increase their levels of consumption and decrease their level of savings.⁴³ As a result, there is a shortage of real resources at existing prices. Assuming that prices are not sticky upwards, the consequence is an increasing input-price level at an accelerating rate. The

⁴³ Again the significance of using the modified Preferred-Habitat theory becomes important. The fact that there is a divergence between entrepreneur's plans and that of consumers is based upon the inclusion of Böhm-Bawerk's formulation of time-preference. Since consumers have not changed their time-preferences, when new money is injected into the economy, they decrease their rate of savings. A Preferred-Habitat theory using Fisher as its foundation could not support such a claim.

inflationary effects of the earlier monetary expansion are compounded due to the need of entrepreneurs to finance their malinvestments. Only with a disaggregated approach, such as with the capital-based approach, can it be shown that inflation must not only accelerate to account for expectations, but that it must expand at a higher multiple to accommodate increasing input prices in order to maintain output levels.⁴⁴

Kashyap, Stein and Wilcox (1993) demonstrate that when the monetary authorities engage in a policy of monetary contraction, there is an immediate effect on the money stock. The first consequence is a reduction of new loans made to entrepreneurs. As input prices increase, there is a need for new financial capital to complete or maintain the malinvestments. The firms with investment-grade bonds have access to credit markets, but firms without this rating scramble for financial capital. They drive up short-term rates in order to finance their projects. Cantor and Wenninger (1993) illustrate how, in a time of credit tightness, funds flow away from low-grade investment instruments (in their analysis, away from the junk bond market) and into bonds with at least a grade of Baa. Long-term rates do not change due to two factors: expectations for future inflation have not changed and firms with investment-grade bonds are able to borrow long-term by tapping the funds flowing out of the low-grade investment instruments. Romer and Romer (1993) show that “[S]mall firms are particularly dependent on banks for finance. . . .”⁴⁵ They also conclude that during the periods of monetary contraction where the Federal Reserve is able to increase short-term interest rates, banks are able to maintain the levels of loans. However, banks do not increase their loan levels that would be required to maintain the malinvestments.

When the monetary authorities believe that the current rate or future rate of inflation is too high, they engage in a policy of monetary tightening. Christiano, Eichenbaum and Evans (1996) show that a contractionary monetary policy increases the federal funds rate. The short-term rates increase relative to the long-term rates. Kashyap, Stein and Wilcox (1993) show that the issuance of commercial paper and bonds jumps relative to bank loans after monetary tightening. Bernanke and Blinder (1992) argue that a tight monetary policy leads to a short-run sell-off of the banks’ security holdings (with little effect on

⁴⁴ If the monetary authorities adopt a policy of accelerating inflation, the consequence is hyperinflation. However, a real resource crunch will usually come about before that point is reached, e.g., the 1980 recession.

⁴⁵ See Romer and Romer (1993) p. 39.

loans), therefore reducing the capital value of these assets and making it more difficult for firms to borrow against their assets. They demonstrate that, over time, banks terminate old loans and refuse to make new ones. In other words, the monetary shock hits securities first. After the securities are sold off, banks rebuild their portfolios and loans start to fall off. After an average period of 2 years, securities return to their previous levels and loans reflect the entire decline. When there is a monetary contraction, according to the results of Kashyap, Stein and Wilcox (1993), a reduction of the supply of loans and the effects on production will not begin to materialize until 6-9 months later. Furthermore, they find evidence that output corresponds with loans. Christiano, Eichenbaum and Evans (1996) argue that households do not adjust their financial assets and liabilities for several quarters after a contractionary monetary shock. The authors also support the findings that the net funds raised by businesses are able to increase for up to one year after the policy shock, after that period, these funds decline. This delay explains the timing issue—the fact that the yield curve tends to invert approximately one year before a recession.

Bernanke and Gertler (1995) argue that interest rates initially spike after monetary contraction and return to their trends after 9 months. This evidence corresponds to the data that show that the yield curve tends to return to its normally positive slope just prior to a recession. This phenomenon was observed in 1957, 1960, 1967, 1989-90 and 2001. Furthermore, Bernanke and Gertler (1995) argue that with a monetary contraction, final demand falls off before production does, implying that inventories rise in the short-run.⁴⁶ According to their results, durable spending displays the largest response to monetary policy shocks, which corresponds to the arguments presented in section 4.

Owens and Schreft (1995) argue that the recessions of 1953-4, 1957-8 and 1960-1 were caused by credit crunches. They report that a credit crunch occurred in the spring of 1953 and the recession began in Q2:1953. While the yield curve did not become invert or humped, it flattened throughout the preceding period.

The next US recession began in Q3:1957. Based on the Minutes of the FOMC Meetings, Romer and Romer (1993) identify a contractionary monetary shock in September 1955. The yield curve displayed the effects of a credit crunch when it became humped in

⁴⁶ Dimelis (2001) demonstrates that business inventories are procyclical and are more volatile than sales. She also points out that EU swings are larger than US swings. She attributes this characteristic to better inventory practices in the US.

December 1956, but it did not invert over the course of the recession.⁴⁷

Owens and Schreft (1995) find evidence of a credit crunch in last third of 1959. The recession began in Q2:1960, and while the yield curve became humped in September 1959, it did not invert.

As noted above, the economic downturns of 1966, 1969-70, 1973-5, 1981-2 and 2001 were also caused by a credit crunch. In February 1966, President Johnson publicly stated that he feared an approaching inflation and said that he was counting on the Federal Reserve to prevent it. Owens and Schreft (1995) report that the Federal Reserve met with bankers and imposed quantitative limits on certain types of lending. The yield curve inverted in September 1966 and the economy experienced negative growth for Q2:1967.

In late 1968, the fear of inflation arose again. Romer and Romer (1993), using the Minutes of the FOMC Meetings, identify the contractionary monetary shock in December 1968. Owens and Schreft (1995) identify the January 14, 1969 meeting of the FOMC where a shift toward tighter monetary policy took place. The recession began in Q4:1969 and lasted through Q4:1970. In any case, the yield curve became humped in November 1968 and inverted in briefly in January 1969 and then again inverted between July 1969 and August 1969. It inverted once more between November 1969 and January 1970.⁴⁸

For the 1973-5 recession (which began in Q4:1973), the fear of inflation led the Federal Reserve to raise discount rate 4.5% to 5% on January 15th, 1973. The yield curve became humped in February 1973, inverted in June of the same year, and remained inverted until September 1974. However Romer and Romer (1993), again using the Minutes of the FOMC Meetings, identify the contractionary monetary shock in April 1974.

Preceding the recession of 1981-2, which began in Q3:1981, Owens and Schreft (1995) argue that the Fed maintained tight credit policy throughout 1981. The yield curve became humped in September 1980 and inverted in November 1980.

On December 5th, 1996, Chairman Greenspan used the phrase “irrational exuberance,” sending the first warning that inflationary pressures were on the horizon. How-

⁴⁷ However, Owens and Schreft (1995) do not find evidence of a credit crunch until the fall of 1957.

⁴⁸ It is interesting to note that the yield curve became humped before the contractionary monetary policy was put into place. A possible reason for this is that there was a real resource crunch just starting to take effect at this time as well. The real resource crunch is explained below.

ever, after a series of rate cuts (cutting the discount rate by 50 basis points to 4.50% by December 1998) the Federal Reserve did not increase the discount rate until August 1999. Beginning in that month, the Federal Reserve began a series of discount rate increases, which culminated in a discount rate of 6.00% in June 2000. The stated reason for the change in policy is found in the FOMC Press Release August 24th, 1999: "Today's increase in the federal funds rate, together with the policy action in June and the firming of conditions more generally in U.S. financial markets over recent months, should markedly diminish the risk of rising inflation going forward." The November 16th, 1999 FOMC Press Release stated that the Federal Reserve was increasing the fed funds rate and the discount rate to check "inflationary imbalances." However the annualized rate of inflation, according to the Consumer Price Index (CPI), for August and November 1999 were merely 1.48% and 1.47% respectively.

As a consequence of the policy of monetary tightening, the yield curve became humped in April 2000 and inverted in August 2000. It and remained inverted through December 2000. The NBER dates the beginning of the recession in March 2001.

While the cause of the 2001 recession may be claimed to be the monetary policy, the Federal Reserve was actually reacting to significant inflationary pressures in the producers markets.⁴⁹ Between April 2000 and January 2001, the Producer Price Index (PPI) for industrial commodities increased over 8.09% and the PPI for all commodities increased over 7.11%. Also during this period, the CPI increased at an approximate rate of only 2.21%. These inflationary pressures are accounted for by a real resource crunch where malinvestments that have built up in the economy can no longer be supported without an accelerating rate of inflation. In other words, if the Federal Reserve had not intervened with a contractionary monetary policy, the economy would have experienced an inverted yield curve and recession because of the impending real resource crunch. The action of the Federal Reserve only hastened the outcome, but did not substantively change the result.

In the surveyed downturns, the monetary authority actively followed a policy of monetary contraction. However, not all recessions are caused by such policies. The recessions of 1980 and 1990-1 were caused by a real resource crunch. The existence of

⁴⁹ This case is the opposite scenario of the 1980 recession, which was a recession caused by a real resource crunch and enhanced by a credit crunch.

malinvestments can be analytically identified only with a capital-based approach. The subsequent need to liquidate these malinvestments is significant because this need causes the yield curve to invert and sets the economy down a path toward recession even without a policy of monetary contraction.

Real Resource Crunch

Unlike more aggregated models, the capital-based approach can also account for the upper-turning point of a business cycle even when the monetary authorities do not engage in monetary tightening. During the malinvestment boom, entrepreneurs are given false signals to undertake malinvestments. Also during the boom phase, consumers rebalance their portfolios so that they increase their spending on consumer goods and reduce their level of savings. These malinvestments are unstable because there are not enough resources to complete and maintain each of these projects. As Robertson described above, windfall profits disappear, wages and input prices rise and “no longer lag appreciably behind the rise in prices. . . .”⁵⁰ Consequently, there is a scramble for financial capital by entrepreneurs to prevent the liquidation of their projects. They bid up short-term rates and the yield curve inverts due to a real resource crunch.

As described above, the price level is driven upward during the malinvestment boom because of two factors: the expectation of future inflation and the bidding up of input prices by entrepreneurs to prevent the liquidation of their projects. Even when the monetary authorities engage in a policy of monetary expansion to meet the increasing expectations of inflation, the total amount of stable investments is limited at any one point in time by the level of savings in the economy. Savings provide the wherewithal for entrepreneurs to build, complete and maintain their projects.

Monetary injections falsify the price signals to the entrepreneurs, causing them to begin more projects than are actually tenable at that point in time. Additionally, the steepening of the yield curve signals to consumers that short-term credit for consumer purchases are less expensive. As described in section 4, a decrease in short-term rates indicates that the cost of financing short-term purchases falls. Furthermore, with ever increasing rates of inflation, consumers will rebalance their portfolios away from savings

⁵⁰ See Robertson (1959) pp. 97 and fn. 41 above.

and checking accounts (and cash holdings) and into tangible assets. In the aggregate, the demand for consumer goods increases and consumers save less. With fewer savings in the economy, the total amount of possible stable investment projects diminishes.

The initial effect of the monetary expansion is that entrepreneurs are able to bid resources to their projects because they are able to cheaply borrow and use the new credit. However, the act of embarking on these investment projects bids up these resource prices. At first the effects of the misalignment of the social structure of production is not apparent and the dueling structure of production emerges (as seen in Figure 2 above).

To better illustrate this process, the following example is provided.⁵¹ Suppose that a builder exists who has enough bricks to finish four houses, yet he starts to build five. With a decrease in the normal rate of profit, he sees the additional house as a potential windfall (economic) profit. He figures that he will be able to purchase the bricks necessary for the completion of the project in the future when he needs them. Suppose further that he borrows \$100,000 to finance the project. Competing entrepreneurs also follow this pattern due to the false market signals. As the entrepreneur runs out of bricks, he starts to purchase more to complete the project. However, other entrepreneurs are also bidding for more bricks. The price of the bricks increases with the increasing demand. The \$100,000 initially borrowed to complete the project is no longer enough. The entrepreneur must find additional financial capital to complete all five houses.

To state the situation more generally, the amount of funds previously borrowed to complete projects (across all lengths) is now insufficient. There is an immediate need for funds to complete and maintain the malinvestments. The scramble for additional funds may be more intense with short-term projects. All entrepreneurs are faced with the alternative of liquidation or of finding supplementary funds to complete their project, but the intensity in demand for funds may be much higher for projects that are almost complete. In other words, an entrepreneur may be more highly motivated to secure funds to finish a project that will produce output next month than he would be to secure funds for a project that will not yield a return until next year.

As a result of these actions, short-term rates are bid up quickly, inverting the yield

⁵¹ This example is similar to one presented in Mises (1966) pp. 559-560.

curve.⁵² Tribó (2001) argues that smaller firms must look for short-term credit when faced with output problems.⁵³ He further finds that larger firms are able to tap into the long-term credit markets. However, as noted above, those firms without investment-grade ratings scramble for financial capital.

Over the course of a business cycle, long rates tend to remain relatively stable. The larger firms with investment-grade ratings are able to attract funds for long-term investments from the low-grade investment sectors. The effect from the increases in the short rates is diminished across the yield curve, because long-term lenders take on less risk since they tend to be the mortgage holders, etc. They are the first to collect if the firm enters bankruptcy. There is a liquidity premium to lending long, yet long-term instruments have an inherent hedge against business cycle risk. Thus, the yield curve inverts instead of shifting.

Three of the recessions since the mid-1950s were not caused by a credit crunch. While there is evidence that there was monetary tightening in every recession except for the 1990-1 recession, the tightening for the 1980 downturn did not take place until after the recession was under way.

As shown above, the 1969-70 recession has elements of both a credit crunch and a real resource crunch. The yield curve became humped in November 1968, a month or two before the policy of monetary contraction was agreed upon by the Federal Reserve. These two causes are not mutually exclusive and may work simultaneously, thus this evidence is not surprising.

The recession of 1980 is an example of a recession caused by a real resource crunch and enhanced by a credit crunch.⁵⁴ The recession began in Q1:1980. The yield curve became humped in September 1978, inverted in December 1978, and remained so until

⁵² Summarizing his empirical findings, Keeler (2002) states, "As the aggregate economy expands and firms progress in building capital and expanding output, shortages of resources occur which raise resource prices. Short term interest rates are increased toward long term rates and the yield curve flattens or may invert. The primary mechanism in this endogenous market process is the intertemporal disequilibrium between sources and uses of income; at low interest rates, consumers and investors increase spending, and need to finance an increase in both consumption and investment, but savers decrease the quantity supplied of funds." p. 5.

⁵³ See also Romer and Romer (1993).

⁵⁴ As noted above, the 1980 and 2001 recession are opposite scenarios but both contain the same components: a policy of monetary contraction and a real resource crunch.

April 1980. While Romer and Romer (1993) identify contractionary monetary shocks in August 1978 and October 1979, Owens and Schreft (1995) argue that the credit crunch did not occur until the first half of 1980. Their position is that in order to regain control over inflation and the expectation of high rates of future inflation, the Carter Administration imposed credit controls on March 14th, 1980. Furthermore, the Federal Reserve did not fully enforce these regulations until the Federal Reserve officials met on May 17th, where Chairman Volcker warned the banks that the Federal Reserve would enforce the program.

Producer prices grew at an accelerating rate between September 1978 and January 1980 (the dates where the yield curve became humped and the beginning of the recession). Over this period, the PPI for industrial commodities grew at a rate over 22.64%, and the PPI for all commodities at a rate over 17.19%. The CPI, over the same period, grew at a rate of 16.99%. This evidence corresponds with the scramble by entrepreneurs to find funding to complete their projects.

The first post-war recession that did not experience a contractionary monetary policy was the 1990-1 recession, which began in Q3:1990. Although the yield curve never inverted, it became humped in February 1989 and lasted in this state through September 1989. While Romer and Romer (1993) identify a contractionary monetary shock in December 1988, most analysts doubt that such a shock occurred. Cantor and Wenninger (1993) state that, "One of the most striking features of the recent credit cycle [of the 1990-1 recession] has been the [credit] crisis that never happened." They argue that there was a boom in the credit markets between 1986 and 1991. Bernanke (1993) interprets the 1990-1 recession as the result of a credit crunch without a contractionary monetary policy.

A real resource crunch is evidenced by increasing rates of input prices while output prices fail to keep pace. Unfortunately, the aggregated data from the Federal Reserve Bank of St. Louis is not specific enough to capture this relationship.⁵⁵ Hughes (1997) provides some evidence that corresponds to the real resource crunch scenario. He shows that a malinvestment boom took place, with long-term bank borrowing by manufacturing

⁵⁵ Nevertheless, the data from FRED II is as follows: from the date when the yield curve began to change, January 1989, through the beginning of the recession, June 1990, the PPI for industrial commodities increased at a rate greater than 3.28%, and the PPI for all commodities increased faster than 3.43%.

industries increasing from \$60.5 billion in 1981 to \$197.2 billion in 1991 (in unadjusted dollars). He also demonstrates that early-stage firms (such as primary metals and Iron and Steel industries) greatly expanded their capacity from 1981 to 1985, but their output prices collapsed in 1986. While his arguments tend to support the Austrian Business Cycle theory, he seems to argue that, at least for the early-stage industries, the 1990-1 recession really began in 1986. Thus to find evidence of a real resource crunch for the 1990-1 recession, one must look at the market which most analysts identify as the one which set off the recession—the commercial real estate market.

When viewed from the perspective of the commercial real estate market, one sees that the 1990-1 recession was caused by a real resource crunch. Cantor and Wenninger (1993) demonstrate that there was a large increase in the value of real estate prices prior to the late 1980s. Additionally, the authors argue that deregulation forced non-bank thrift institutions (like insurance companies) to look for ways to increase their rates of return. Thus these institutions extended credit to more risky ventures (like those in the real estate market), but the capital requirements for these thrift institutions remained low and many weak firms were exempted from tough regulatory scrutiny.

Cantor and Wenninger (1993) point out that the real estate market collapsed in the late 1980s, after which regulators increased scrutiny of these types of loans, making it very difficult to obtain funding for real estate ventures. Owens and Schreft (1995), in their paper which also observed decreasing real estate values in the late 1980s, state that there were complaints that regulators were too closely scrutinizing real estate portfolios, making real estate lending extremely difficult. Many new buildings came on the market at the same time, depressing rental and sales prices. Additionally in many cases, the tax breaks that made commercial building profitable were removed. Bernanke (1993) observes that large losses in the real estate market reduced the amount of bank capital, which he labeled a “capital crunch.” However, Bernanke downplays the supply of funds as a major cause of the recession, because as bank loans fell in 1989, commercial paper increased. Cantor and Wenninger (1993) state that during the period between 1986 and 1991, “nondepository credit growth continued to exceed GDP growth by a wide margin (4.5 to 5.5 percentage points). Depository credit, on the other hand, decelerated sharply as thrift credit went into an outright decline in the 1989-1991 period.”⁵⁶ There was a

⁵⁶ See Cantor and Wenninger (1993) p. 5.

shift from financing through banks, etc. to self-financing in the commercial paper and commercial bond markets. Those firms with investment-grade securities were able to obtain financing, while those without such a rating were not. Cantor and Wenninger (1993) further argue that just prior to the recession, (1989-90) money stopped flowing into “junk” bonds and instead went into investment-grade corporate bonds. Despite an easy monetary policy,⁵⁷ those borrowers without direct access to the financial markets (i.e., those without investment-grade ratings) did not benefit from this policy. Their scramble for financial capital caused the yield curve to become humped. Below investment-grade borrowers were shut out of the long and short-term money markets and eventually were forced to liquidate their projects, while those with investment-grade ratings benefited from the easy credit policy.

Bernanke and Lown (1991) support the conclusion that the decrease in bank loans was not supply-side related. Thus, they are skeptical that a credit crunch caused the recession. The lack of contractionary monetary policy explains the appearance of a humped yield curve instead of an inverted yield curve. While Bernanke and Lown argue that the demand for loans was a more important cause of the 1990-1 slowdown, the evidence they provide is that all forms of credit (including commercial paper) decreased during the 1990-1 recession—indicating a decrease in demand for credit. There seems to be a timing discrepancy in their analysis. In 1989 commercial paper increased and then, when the recession began, all forms of credit decreased. Their evidence supports the real resource crunch scenario, instead of the scenario where a fall in the demand for credit caused the recession. There was a scramble for credit in the late 1980s, which is seen in the increase in commercial paper issuances and the change in the shape of the yield curve. When businesses started to fail in 1990, the demand for all credit fell and the recession was underway.

6 Summary

The capital-based approach compares favorably with CCAPM and Estrella models because it is able to explain why the yield curve tends to invert before a recession. Unlike these other models, this approach centers its focus on the malinvestments built up during

⁵⁷ Owens and Schreft (1995) report that the Fed eased monetary policy in spring of 1990.

the malinvestment boom. These other models, by over aggregating, are unable recognize that the root cause of the inversion of the yield curve is the malinvestments.

In this paper, the assumption was made that the initial monetary injection was short-term working capital. It has been shown that this capital is transformed into fixed capital, long-term projects and early-stage malinvestments. To the extent that these projects are inconvertible, the liquidation process becomes more severe.

The modified Preferred-Habitat theory is an essential component to the model used, because it is able to illustrate how monetary injections lead to a disequilibrium between consumption/savings horizons of households and the investment projects of entrepreneurs. Monetary injections cause the yield curve to steepen which falsely signals entrepreneurs to begin new investments and encourages households to increase their demand for final goods and services.

The unstable malinvestments force input prices to rise and lead to a credit crunch, a real resource crunch, or a combination of both. In their attempts to prevent their individual projects from being liquidated, entrepreneurs will cause the yield curve to flatten, become humped or even invert as they scramble for financial capital (even when the monetary authority adopts a policy of easy credit). Thus in every recession since the mid-1950s, an inverted or humped yield curve occurred no more than 5 quarters prior to the upper-turning point of the business cycle.

Bibliography

- [1] Bernanke, Ben, 1993. "How Important is the Credit Channel in the Transmission of Monetary Policy?: A Comment," *Carnegie-Rochester Conference Series on Public Policy*, 39, 47-52.
- [2] Bernanke, Ben, and Alan Blinder, 1992. "The Federal Funds Rate and the Channels of Monetary Transmission," *American Economic Review*, 82(4), 901-921.
- [3] Bernanke, Ben, and Mark Gertler, 1995. "Inside the Black Box: The Credit Channel of Monetary Policy Transmission," *The Journal of Economic Perspectives*, 9(4), 27-48.

- [4] Bernanke, Ben, and Cara Lown, 1991. "The Credit Crunch," *Brookings Papers on Economic Activity*, 2, 205-247.
- [5] Brown, William, and Douglas Goodman, 1991. "A Yield Curve Model for Predicting Turning Points in Industrial Production," *Business Economics*, July, 55-58.
- [6] Cantor, Richard, and John Wenninger, 1993. "Perspective on the Credit Slowdown," *Federal Reserve Bank of New York, Quarterly Review*, 18(1), 3-36.
- [7] Cecchetti, Stephan, 1987. "The Case of the Negative Nominal Interest Rates: New Estimates of the Term Structure of Interest Rates During the Great Depression," NBER, Working Paper no. 2472.
- [8] Christiano, Lawrence, Martin Eichenbaum, and Charles Evans 1996. "The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds," *The Review of Economics and Statistics*, 78(1), 16-34.
- [9] Cochran, John, 2001. "Capital-Based Macroeconomics," *Quarterly Journal of Austrian Economics*, 4(3), 17-25.
- [10] Cwik, Paul, 2004. "An Investigation of Inverted Yield Curves and Economic Downturns," forthcoming Dissertation, Auburn University.
- [11] Cwik, Paul, 1998. "The Recession of 1990: A Comment," *Quarterly Journal of Austrian Economics*, 1(2), 85-88.
- [12] Davis, Lance, 1971. "Capital Mobility and American Growth," in: Fogel, Robert, and Stanley Engerman (Eds.), *The Reinterpretation of American Economic History*, New York, NY: Harper & Row, pp. 285-300.
- [13] Dimelis, Sophia, 2001. "Inventory Investment over the Business Cycle in the EU and the US," *International Journal of Production Economics*, 71(1-3), 1-8.
- [14] Estrella, Arturo, 1998. "Monetary Policy and the Predictive Power of the Term Structure of Interest Rates," *Federal Reserve Bank of New York, Working Paper*, November.

- [15] Estrella, Arturo, and Frederic Mishkin, 1998. "Predicting U.S. Recessions: Financial Variables as Leading Indicators," *The Review of Economics and Statistics*, 80(1), 45-61.
- [16] Evans, Martin, 1987. "A Macroeconomic Model of the Term Structure of Interest Rates," New York University Salomon Brothers Center, Working Paper no. 445, 29.
- [17] Garrison, Roger, 2001. *Time and Money: The Macroeconomics of Capital Structure*, New York, NY: Routledge.
- [18] Hayek, Friedrich von, 1969. "Three Elucidations of the Ricardo Effect," *Journal of Political Economy*, 77(2), 274-285.
- [19] Hayek, Friedrich von, 1945. "The Use of Knowledge in Society," *American Economic Review*, 35(4), 519-530.
- [20] Hughes, Arthur, 1997. *The Recession of 1990: An Austrian Explanation*, *Review of Austrian Economics*, 10(1), 107-123.
- [21] Kashyap, Anil, Jeremy Stein and David Wilcox, 1993. "Monetary Policy and Credit Conditions: Evidence from the Composition of External Finance," *American Economic Review*, 83(1), 79-98.
- [22] Keen, Howard, 1989. "The Yield Curve as a Predictor of Business Cycle Turning Points," *Business Economics*, 24(4), 37-43.
- [23] Keeler, James, 2002. "Relative Prices and the Business Cycle," paper presented at the Southern Economic Association Annual Meetings in Tampa, the Society for the Development of Austrian Economics, Session 42C: Research in Empirical Austrian Economics.
- [24] Keeler, James, 2001. "Empirical Evidence on the Austrian Business Cycle Theory," *Review of Austrian Economics*, 14(4), 331-351.
- [25] Kessel, Reuben A, 1965. "The Cyclical Behavior of the Term Structure of Interest Rates," NBER, Occasional Paper no. 91.

- [26] Machlup, Fritz, 1935. "The Rate of Interest as Cost Factor and Capitalization Factor," *American Economic Review*, 25(3), 459-465.
- [27] Machlup, Fritz, 1932. "The Liquidity of Short-Term Capital," *Economica*, 12(35-38), 271-284.
- [28] McCulloch, J. Huston, 1990. "Appendix B: U.S. Term Structure Data, 1946-87," in: Friedman, B.M., and F.H. Hahn (Eds.), *Handbook of Monetary Economics*, North-Holland, Amsterdam, pp. 672-687.
- [29] Miller, Roger LeRoy and David VanHoose, 2001. *Money, Banking, & Financial Markets*, 2001 edition, Australia: South-Western / Thomson Learning.
- [30] Mises, Ludwig von, 1980. "Profit and Loss," in *Planning for Freedom*, Fourth Edition, South Holland, IL: Libertarian Press, 108-150.
- [31] Mises, Ludwig von, 1966. *Human Action: A Treatise on Economics*, Third Revised Edition, Chicago, IL: Contemporary Books, Inc.
- [32] Mishkin, Frederic, 2001. *The Economics of Money, Banking, and Financial Markets*, Sixth edition, New York, NY: Addison Wesley Longman.
- [33] Owens, Raymond, and Stacey Schreft, 1995. "Identifying Credit Crunches," *Contemporary Economic Policy*, 13(2), 63-76.
- [34] Robertson, Dennis, 1959. *Lectures On Economic Principles*, Volume 3, London: Staples Press.
- [35] Romer, Christina and David Romer, 1993. "Credit Channels or Credit Actions?" NBER, Working Paper no. 4485.
- [36] Schmitz, Stefan, 2003. "Uncertainty in the Austrian Theory of Capital," unpublished manuscript.
- [37] Skousen, Mark, 1990. *The Structure of Production*, New York, NY: New York University Press.
- [38] Tribó, Josep, 2001. "Inventories, Financial Structure and Market Structure," *International Journal of Production Economics*, 71(1-3), 79-89.

New Perspectives on Political Economy

Volume 1, Number 1, 2005, pp. 38 – 60

Institutions and Instability in Africa: Nigeria, Sudan, and Reflections from Mises's Nation, State and Economy

Aleksi Ylönen*

JEL Classification: D63, I34, R58, N27

Abstract: In recent years as part of development economics discourse a new emphasis has emerged dealing with institutions and economic development. One path of the institutional analysis has drawn on historical evidence in order to explain the root causes of welfare differences between low-income and high-income countries. As a result, increasing evidence has surfaced linking implementation and persistence of types of institutions to the level of economic development and accumulation in many of the former colonies. This paper argues that in Nigeria and Sudan extractive colonial institutions were imposed and their legacy endured to the period of independence. By creating poverty and inequality as control mechanisms in favor of the colonizer, these institutions led to political and socio-economic marginalization of large segments of the population and therefore also to weak, politically unstable, and conflict torn post-colonial states.

* Aleksi Ylönen, Department of Economics, Universitat Jaume I, Spain, aylnen@guest.uji.es / aeylo-nen@hotmail.com.

1 Introduction

Renewed interest has emerged in recent years within development economics regarding institutional approaches to social and economic organization in the developing countries. This has been partly due to the recognition of failures of a large part of the World Bank structural adjustment programs destined to Africa, which have generally paid little attention to the role of institutions in economic readjustment and creation of markets.¹ However, although a variety of case studies that deal with institutions and economic organization in Africa have been undertaken, very few analyses have focused on the colonial roots explaining contemporary resource distribution and their effects on political stability.

In Africa, the colonial institutions were imposed to extract resources and create poverty among the colonial subjects for enhanced control over them.² In most occasions, these governance structures had to be blended with the traditional institutions and informal social control mechanisms in order to function appropriately in each particular cultural setting. Hence it is important to recognize the cultural differences of institutions and how designing development policies for Africa based exclusively on neoclassical understanding of institutions is likely to lead to failure.

This paper goes beyond the interpretation of institutions in the neoclassical framework and aggregates them as formal and customary governance structures and practices. This is important in order to observe how institutions dictate the national resource distribution and affect the political stability in African states such as Nigeria and Sudan. The observations provided in this paper show evidence of how poverty and inequality (especially horizontal inequality) that emerge from political, economic, and social group marginalization may contribute to political instability and civil violence. However, due to the absence of reliable numerical data, predominantly descriptive accounts have been interpreted in order to examine the role of institutional practices in the emergence of political instability and civil strife in Nigeria and Sudan.

¹ See i.e. Howard Stein, "Theories of Institutions and Economic Reform." *World Development* Vol. 22, No. 12 (1994), pp. 1833-1849.

² See i.e. Macharia Munene, "Culture and the Economy: The Creation of Poverty", Kenyatta University Culture Week Seminar, September 20, 2001.

The article is divided into nine sections. Section two provides the view of how the paper defines institutions and how they are linked to political control. The third section discusses Mises's perception of imperialism and nationalism related to marginalization based on *Nation, State and Economy* (1919). Fourth section introduces a framework that seeks to link colonial institutional legacy to political instability, while sections five, six and seven attempt to establish causal linkages between institutions, extraction, poverty and inequality, and ethnic mobilization. Section eight provides brief case studies of Nigeria and Sudan, and section nine concludes.

2 Institutions as Controlling Mechanisms

Institutions are often viewed as arrangements governing societal organization. According to *The New Shorter Oxford English Dictionary* institution can be "An established law, custom, or practice".³ They are normally defined depending on the focus of the investigation and academic perspective within each discipline. Therefore, there does not appear a generally accepted definition of institutions among scholars. For instance, according to Stein (1994) in economics the neoclassical Structural Adjustment Programs (SAP), the New Institutional Economics (NIE), and the Old Institutional Economics (OIE) traditions all have their particular views of institutions that govern a society and organize its economy. Diverging views of institutions by the eminent scholars of the Austrian School of Economics, the NIE, and the OIE traditions demonstrates the problem of definition. While Mises (1949) argues that institutions impose rules that prevent "... an attempt to substitute more satisfactory state of affairs for a less satisfactory one"⁴, North (1991) perceives them as "... the humanly devised constraints that structure political, economic and social interaction"⁵ pointing out the common view among the NIE scholars. On the other hand, Veblen (1919), an eminent scholar within the OIE tradition, views institutions as less instrumental and more natural "... as settled habits of thought common to the generality of man".⁶

³ *The New Shorter Oxford English Dictionary*. 1993. Volume 1. (Oxford: Oxford University Press), p. 1383.

⁴ Ludwig von Mises. *Human Action, A Treatise on Economics* (New Haven: Yale University Press, 1949), p. 97.

⁵ Douglass North. "Institutions", *Journal of Economic Perspectives* 5, 1 (1991), p. 97.

⁶ Thorstein Veblen. *The Place of Science in Modern Civilization and Other Essays* (New York: Huebsch, 1919), p. 239.

This paper adopts a definition of institutions as governance structures, socially established laws, customs or practices. It does not differentiate between informal constraints and formal rules, but aggregates institutions that affect political, economic and social interaction.⁷ According to Acemoglu et. al. (2003) institutions can be viewed as social arrangements including "... constitutional and social limits on politicians' and elites' power, the rule of law, provisions for mediating social cleavages, strong property rights enforcement, a minimum amount of equal opportunity and relatively broad-based access to education, etc."⁸ They can be aggregated into broad categories such as political and economic institutions that dictate the economic performance and the distribution of resources within a society.

Acemoglu et. al. (2004) construct a framework that places a political elite or a ruler in charge of political institutions, which determine their *de jure* political power, while the politically powerful groups that enjoy the largest share of national resources possess *de facto* political power. It explains how concentration of political power to a small elite, as in the case of many former colonies, results in political and economic institutions that predominantly serve interests of that particular group in expense of the general population. The framework introduced in this paper follows this line of thought but is extended further in an attempt to understand how institutions are linked to inequality and political instability in post-colonial states, such as Nigeria and Sudan.

3 Of Imperialism, Authoritarian State, and Marginalization

Ludwig von Mises views authoritarian state structures that were also present in most African colonial states as obstruction to equality and liberty of individuals and minority populations. In his *Nation, State and Economy* (1919), Mises argues that

The absolute ruler not only regards every other community between his subjects as dangerous, so that he tries to dissolve all traditional comradely relations between them that do not derive their origin from state laws enacted by him and is hostile to

⁷ See North (1981, 1991) on the evidence of differences between institutions as informal constraints or formal rules.

⁸ Daron Acemoglu, Simon Johnson, James Robinson and Yunyong Thaicharoen. "Institutional causes, macroeconomic symptoms: volatility, crises and growth", *Journal of Monetary Economics* 50 (2003), p. 52.

every new formation of community, perhaps through clubs; he also will not allow the subjects of his different territories to begin to feel themselves comrades in their role as subjects.⁹

This reflects how the same divide and rule tactics that were introduced in Europe were later used throughout Africa during colonization and created disintegrated societies in which ethnic or local affinities overtook the sense of nationhood. Mises further suggests that, “The imperialistic people’s state scarcely differs from the old princely state . . . It wants to hear nothing of the right of peoples”¹⁰. This links imperialism and colonization to oppression of civil liberties and dispossession of the right to self-determination.

What Mises further aptly points out is that imperialism does not serve common good since it is neither cost effective for the imperial power nor does it recognize liberty of the oppressed populations.¹¹ He therefore foresaw the European exodus from Africa due to rising costs of colonial administrations, and the effects of the authoritarian governance structures that focused on extraction, repression and marginalization of the local populations through the emergence of disparate levels of poverty and inequality.

The colonial institutional legacy endured in Africa to a large extent to the post-colonial period. As van de Walle (2001) suggests, the colonial dynamics of exploitation endured as

African elites inherited state structures and a style of governance from the colonial era that were illiberal and geared toward enforcing law and order rather than the promotion of citizen welfare. Traditions of authoritarian rule, paternalism, and dirigisme were embedded in the institutions the new leaders inherited and largely kept.¹²

Similarly, Mises recognizes that minorities, such as the African elites, often oppose democratic transition in order to avoid power sharing that would threaten their position in controlling the state apparatus.¹³

⁹ Ludwig von Mises. *Nation, State and Economy*. Translated English edition by Leland B. Yeager (New York University Press, 1983), p. 59.

¹⁰ *Ibid.*, 108.

¹¹ *Ibid.*, 106-114.

¹² Nicolas van de Walle. *African Economies and the Politics of Permanent Crisis, 1979-1999*. (Cambridge: Cambridge University Press, 2001), pp. 116-117.

¹³ Mises, *Nation, State and Economy*. p. 77.

He is also concerned about oppression and inequality among individuals. Mises points out that authoritarian state structures create political marginalization and inequality by suggesting that

He who is compelled to obey laws on whose enactment he has no influence, he who must endure a government ruling over him in whose formation he can take no part, is, in the political sense, unfree and is politically without rights, even though his personal rights may be protected by law.¹⁴

This has also been the case in many post-colonial societies in Africa, where minorities, or sometimes majority populations, have been marginalized by elites that emerged to dominate the state structures after colonialism.

Finally, although Mises views proportional political representation as insufficient solution for recognition of minority rights in multiethnic states, he also recognizes the ongoing competition between different groups and condemns the assimilation efforts by majorities. He concludes that in this way “Liberal nationalism gives way to militant antidemocratic imperialism”.¹⁵

Of course, all of the above contradicts the liberal tradition and remains tied to political instability in the form of emerging challenges to the colonial and the later post-colonial repression in Africa. The next section introduces the framework constructed to interpret the linkages between the persisting authoritarian colonial institutional tradition and the political instability in Africa.

4 The Framework

Institutions put forward in various tropical colonies were designed to facilitate effective extraction of resources. This was especially true in Africa and in the cases of Nigeria and Sudan that were both believed to be rich in natural resources.¹⁶ Since the extraction had to be undertaken in a politically unstable environment, the institutions were designed to impoverish the local population particularly in the less controlled marginal areas of

¹⁴ *Ibid.*, 73.

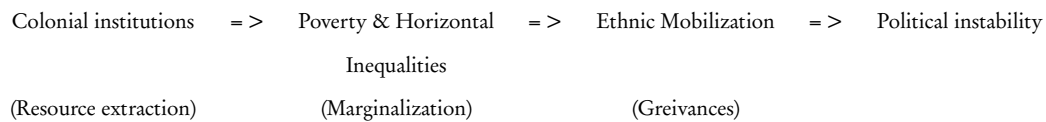
¹⁵ *Ibid.*, 84.

¹⁶ See more on Nigeria and Sudan in section eight.

the colonial authority. This guaranteed the colonial *status quo*. As a result, the peripheral regions that were often inhabited by large minority populations grew increasingly politically, economically, and socially marginalized.

After colonialism the marginal regions remained in a similar status within the post-colonial polities that were now governed by local elite put in place by the former colonial masters. Although nominal structural changes often took place to reform the colonial institutions in preparation for independence, they did not change *in practice* and reproduced the colonial society in the self-governed states through the persistence of institutional traditions dictating dynamics of the national resource distribution. While the post-colonial states often resembled their colonial counterparts, they grew to be weaker entities since the authority of the central government declined and became increasingly challenged by the marginalized groups of the national periphery.

The framework introduced here can be simplified in the following manner:



The following sections five, six, and seven discuss in more detail the evidence of linkages between institutions, extraction of resources, poverty, inequality, grievances, ethnic mobilization, and political instability.

5 Institutions and Extraction of Resources

A commonly accepted characteristic of institutions also recognized in this paper is that they tend to evolve slowly. It has also been established in the political economy literature that colonialism explains the foundation of capitalist institutional structure in post-colonial societies and that institutions dictate economic outcomes. According to Acemoglu et. al. (2002), the Europeans implemented various types of institutions depending on the geographical characteristics of colonies. Acemoglu et. al. (2003; 2004) divide these into two main categories. On the one hand, they argue that institutions that ensured the property rights for a large part of the society emerged in the European settler colonies with large settler communities and insignificant amount of local influence. Such

colonies tended to be situated in geographically cooler parts of the globe, such as those located in North America.

On the other hand, institutions designed to extract wealth tended to emerge in colonies that were not predominantly European settler communities. These regions were mostly located in the tropical part of the planet and largely consisted of native populations.¹⁷ Therefore, these tropical colonies had to be controlled through institutions that exercised repressive authoritarianism to establish control since the colonial administrators were often few.

European imperialism was extended to the tropical colonies partly because of resources. In order to ensure an effective flow of the local resources to feed the colonizing state, one solution was to implement what Acemoglu et al. (2003; 2004) call the “extractive” institutions. These institutions served two primary functions. First, they had to secure sufficient control of the territory in order to guarantee effective exploitation of local resources. Second, they had to be structured so that they could be easily adapted to changing circumstances by being firmly controlled and manipulated by the colonial authorities.

Similarly to the institutional perception of the OIE tradition, it was found out early on that imposition of institutions that were culturally embedded on the European model would not work effectively in most of the tropical colonies. Therefore, the extractive institutions needed to be disguised in order to be acceptable enough to be implemented in the local societies, and achieve meaning and credibility among the local population while ensuring efficient exploitation. The best way to impose these institutions was by finding local societal structures that could be put to serve the colonial interests.

Consequently, institutions that did not effectively limit the rulers’ and elites’ personal interests emerged in tropical colonies where the Europeans were a small minority, governing large local populations. In addition, since institutional practices are slow to change, they endured throughout the period of preparation of colonies for independence although superficial changes were often made to the governing structures.¹⁸

Similarly, after independence the institutional norms, practices, and customs imposed

¹⁷ See i.e. Easterly and Levine (2003), and Acemoglu et al.

¹⁸ See section eight for Nigeria and for Sudan.

during the colonial period persisted and often dominated post-colonial societies in Africa that had experienced short but the most intensive colonization. In other words, it is therefore plausible to argue that colonial institutional tradition is largely responsible of how political power and governance structures are perceived in a number of contemporary African states. Hence, the prevailing governance structures and practices are associated with marginalization of minority (or at times majority) populations and the problem of poverty and inequality, which are briefly discussed next.

6 Colonial Institutions: Creation of Poverty and Inequality

Creation of Poverty

As was noted in the previous section, one of the main functions of the extractive institutions implemented in African colonies was to create a sphere of control that facilitated effective exploitation of resources. This control was to be established through minimization of challenges to the colonial order through creation of poverty. Macharia Munene argues that

... the reason some people go out of their way to create poverty is political, because poverty is a controlling mechanism ... Those who resisted colonialism had to be made poor and all the resources available to colonial authorities including missionaries were deployed to impoverish the African.¹⁹

In general terms, this seems to be the case in colonial Africa. However, as the cases of Nigeria and Sudan later demonstrate, the greatest amount of poverty and exploitation was created in the marginal areas of the colonial authority, while the central regions where most political and economic power was concentrated often enjoyed proportionately larger share of development.

The regional development disparities that emerged during the colonial period often persisted after independence due to the domination of colonial tradition in the post-colonial state. This lent itself to the emergence of a perception of group-based inequalities

¹⁹ Macharia Munene. "Culture and the Economy", p. 2.

that manifested themselves in growing political instability as can be seen in the cases of Nigeria and Sudan.

Of Inequality

Equally to the institutions, it is difficult to establish globally accepted definition for inequality. According to *The New Shorter Oxford English Dictionary*, inequality can be defined as “Inconsistency in treatment of people or distribution of things . . . superiority or inferiority in relation to something . . . [or] social or economic disparity”.²⁰ In economics literature, there is an overwhelming emphasis of measuring inequality through individual income. This is also the case in the emerging economic literature on causes of civil wars. Most work in this field related to inequalities and conflict has focused on economic inequalities often measured through income distribution.²¹

However, Stewart (2000) points out that income distribution is a vertical measure and often becomes insignificant or significant in reducing rather than provoking conflict. Hence, Stewart (2000, 2002) and Cramer (2003) propose that horizontal inequality, which measures inter group inequality within a society, is more meaningful way of measuring the emergence of conflict between groups. This measurement should be undertaken through classification of groups based on religion, ethnicity, or class. According to them, these group identities are important, if not the main factor creating political instability and mobilization for civil conflict.²²

Stewart (2000) further suggests that political, economic and social inequalities matter in the overall measurement of horizontal inequality. She develops a framework that has four categories measuring political participation, economic assets, employment and income, and social access and situation. Another study, Stewart (2002), illustrates empirical evidence on horizontal inequalities in Mexico, Fiji, Uganda, Sri Lanka, South Africa, Northern Ireland and Brazil, concluding that they are highly conflict provoking. She also goes further in demonstrating that both political and economic group differentiation is

²⁰ *The New Shorter Oxford English Dictionary*, p. 1357.

²¹ For recent political economy literature on civil wars, see i.e. Collier (2000), Collier and Hoeffler (1998), Keen (1997; 2000), De Soysa (2000).

²² Large literature that examines identity and mobilization for civil violence exists. See section seven for some references.

particularly important in the attempt of ethnic elites to mobilize groups for conflict. This is briefly discussed in section seven.

Furthermore, it is important to note that data on horizontal inequalities is usually hard to find. It is so because inequalities between groups are often not reported due to their contested nature. In addition, the most severe horizontal inequalities tend to exist within states that have non-democratic regimes or otherwise highly concentrated political power to one particular group. It is also important to note that the regimes governing states with high level of horizontal inequality are often unstable, challenged, vulnerable or lacking popular support. Although data on horizontal inequalities is difficult to encounter, in Stewart (2000, 2002) such efforts have been undertaken but remain severely incomplete. More recent attempt, UNRISD (2004) portrays evidence on horizontal group inequalities in the public sector in 16 countries in a comparative study. Finally, since most compelling evidence on horizontal inequalities can often be found mainly in descriptive historical accounts, the analysis of Nigeria and Sudan undertaken in this paper relies on such evidence.

Institutions and Inequality

How are institutions linked to inequality? Although the relation between institutions and growth (or lack of) has been widely documented in the political economy literature, less work has been done relating institutions to inequality. Similarly to what Mises suggests, North (1981) argues that the politically powerful groups often do not develop efficient institutions because these may prevent them from maximizing private revenues, and Robinson (1999) affirms that the political elite may use predatory strategies to protect its political power. This is largely why peculiar political and economic arrangements and decisions that reinforce the elite's grip to political power, rather than advancing the well being of larger population within a society, have been widespread in numerous African states.

Similarly to what has been suggested earlier regarding how post-colonial institutions in Africa function *in practice*, Glaeser et. al. (2003) argue that the political elites are often able to subvert legal, political and regulatory institutions for their own benefit in the expense of large cross section of the society. This takes place in societies with institutions

that are not sufficiently independent of the manipulation by the political elite. Moreover, Acemoglu and Robinson (1999) develop a theory of social conflict that focuses on the role of elite economic opportunities and constraints shaping the institutional structures, while Acemoglu et. al. (2003; 2004) point out how institutions may prolong the largely non-democratic and unpopular regimes' ability to remain in power, constrain economic development, and restrain the access to national resources by groups that could emerge to threaten them. This, of course, is a plausible argument considering the possible prevalence of horizontal inequalities and seems to have been the case in Africa in places such as Nigeria, and Sudan. Finally, van de Walle (2001) argues that the political elite remains relatively isolated from societal pressures for institutional change in the case of most of Africa, since political elites are able to manipulate or subvert institutions for their liking in order to maintain themselves in power. This gives the political elite an advantage in maintaining the *status quo* over the societal pressures. Van de Walle's argument coincides with an earlier groundbreaking work of Alavi (1972), which suggests that postcolonial states are almost immune to the pressure of social classes and rather largely respond to external interest.

Acemoglu et. al. (2004) establish that extractive institutions that still remain intact in some states correlate with lower level of economic development than do so called *good* institutions. They are often also present in weak states with small privileged political elite and strong patrimonial networks that control the national resources.²³ Finally, van de Walle (2001) argues that the dynamics of maintaining political power in Africa sustain the role of national wealth in serving to enrich the political elite (and hence affirm its power) in expense of a wholehearted investment in development that might give leverage to groups contesting the authoritarian political order.

Finally, it seems to be that good institutional quality is associated with more equitable income distribution. Chong and Gradstein (2004) show evidence that states with poor institutions are likely to have high level of income inequality, and how interaction of political and income inequality may result in constrains of institutional reforms as in the case of Russia. Similarly, Sonin (2003) presents evidence on how poor quality of institutions is associated with low growth due to a disproportionate resource allocation favouring the rich.

²³ See i.e. Reno (2002; 2003; 2004) how this relates to violent political strife in weak or collapsed states.

7 Evidence on Inequality and Mobilization for Violence

Grievances related to ethnic mobilization for civil war have received much attention in recent years especially in the literature that concentrates on economic agendas in civil conflict. The greed versus grievance debate, which emerged during joint International Peace Academy and World Bank research project on causes of civil wars demonstrated the prevalence of economic incentives in the contemporary civil war formations. This resulted in studies that also explored the link between inequality and violent conflict. A number of them have considered income inequality insignificant to civil war formation. However, as noted earlier, other studies have proposed different way of measuring inequality and have since come to distinct conclusions.²⁴

How are groups mobilized for inter-group violence? It has been widely documented that in the circumstances of insecurity and state failure to adequately protect its population, individuals tend to find refuge in their socially defined groups. According to Stewart (2000) in Central Africa, these groups are often defined based on ethnic identity while in Latin America the defining factor has largely been a mixture of class and ethnic elements. Badru (1998) suggests that the ethnic affinities play more important role than class in Nigeria, while Deng (1995), Lesch (1998), and Johnson (2003) make similar conclusions on Sudan.

Leaders of ethnic groups sometimes manipulate them with self-interested political and economic rhetoric that serves also as powerful motivational factor for group mobilization. Therefore, the economic, the political, and the cultural become intertwined as these motivational forces guide the ethnically defined groups. According to Mozaffar et. al. (2003) these groups then become mobilized as *ethnopolitical* entities. Furthermore, according to Fearon and Laitin (2000) overwhelming evidence exists claiming that ethnic identities are predominantly constructed rather than primordial, unchanging, or everlasting violent conditions between groups. Scholars of the constructivist school also argue that ethnic groups are constructed and manipulated by their respective political leaders with sometimes devastating consequences such as in Burundi, Rwanda, or Yugoslavia.²⁵

²⁴ For instance, Cramer (2003) suggests that economic inequality inseparably bound to social, political, cultural and historical inequality and agrees with Stewart (2000) in that tackling horizontal inequalities reduces the preconditions for violent conflict in general terms.

²⁵ See Uvin (2000) for Burundi, Gaffney (2000) for Rwanda and Ignatieff (1997) for Yugoslavia.

Construction and manipulation of ethnic identities is often undertaken through political rhetoric and symbolism and cultivation of fear of the relative other. According to Ignatieff (1997), rhetoric that captures a suitable version of group's history tends to be particularly useful since most ethnic groups have engaged in inter-ethnic violence in the past. In order to aggravate the group consciousness further, leaders of the marginalized ethnic groups tend to use rhetoric based on marginalization, inequality, and poverty, at times provoking sufficient level of grievances to invoke mobilization for violence.²⁶ In other words, ethnic mobilization for violence needs inter group grievances that are best provided by the perceived political, economic or social inequalities.

8 The Framework in the Context of Nigeria and Sudan

Institutions and Domination in Nigeria

Nigeria was integrated in the sphere of British imperialism after the Berlin Conference of partition of Africa in 1884. Consequently, the British moved in to exploit the Nigerian resources by dividing it administratively into three protectorates. By the time the WWI broke out, British authorities unified the three regions into a colonial state called Nigeria. However, the administrative unification did not overcome the wide regional and ethnic differences that became dominant ways of group identity expression in the absence of any kind of national identity building attempt by the colonial authorities.

In contrast, the colonial masters exploited ethnic diversity in order to consolidate their rule. As a result, governance structures were designed with an objective to extract local resources and were supplemented by the colonial trading houses, such as the United African Company. In order to run the colonial economy efficiently, African labor was extracted by playing ethnic affinities against each other, resulting in efficient, loyal, and obedient work force.²⁷ During this period most development took place in the southern and coastal Nigeria where colonial administrative and economic centers were located, while peripheral regions became economically marginalized.

When the international and internal pressure for the British colonial masters to leave

²⁶ In case of Sudan this has been particularly clear in the case of a number of insurgencies.

²⁷ Pade Badru. *Imperialism and Ethnic Politics in Nigeria, 1960-96*. (Asmara: Africa World Press, 1998), p. 6.

Nigeria became more evident in the 1950s, hasty preparations for independence largely ignored the institutional practices and dynamics of resource distribution formed throughout the colonial period. According to Badru (1998),

In desperation, the British imperial state had succumbed to pressure from powerful northern economic interests and designed a constitution that gave serious concessions to northern elites. The parliamentary model that was proposed was modulated by a system of proportional representation in which the federal parliament would be dominated by the ethnic group with the largest population.²⁸

This shows how regional elites became hungry for political power upon British withdrawal since domination of state apparatus meant economic prosperity. Since the northern elites became dominant in the first independent government, other regional elites turned increasingly restless. This led to the first post-colonial crisis.

The ethnic regionalism had also another effect. It prevented the northern elite from effectively controlling the country because the constitutional structure was not suitable to consider regional interests in a just manner. In addition, the legacy of the colonial institutions produced a post-colonial society in which the flow of resources was similar to that of the colonial state. However, since the state itself was now weaker than during the colonial era due to ethnic rivalry, the fear of increased ethnic domination resulted in a fertile ground for mobilization for violence. This was supplemented by regional economic and political inequality that favored the southern region over the oil rich Igbo dominated eastern Nigeria.

Large-scale violence erupted in the mid 1960s, when the Igbo elite rallied its followers to challenge the state. It seems that the emergence of civil war has to do with interests of regional elites as Badru (1998) suggests, "The answer is simple, [I]t was the elite's war of greed fought over the private distribution of petro-dollars".²⁹ Since the civil war, serious coups have taken place in Nigeria together with ongoing demands for additional states to accommodate interests of the local elites.³⁰ In addition, according to Ashwe (1986) the

²⁸ Ibid., 8.

²⁹ Ibid., 85.

³⁰ See Rotimi T. Suberu, "The Struggle for New States in Nigeria, 1976-1990", *African Affairs* 90, 1991, pp. 499-522.

decades-long conflict over oil revenue sharing shows how just and transparent institutions managing natural resource rents might reduce political instability in Nigeria.

These crises illustrate two important characteristics of the weakness of the Nigerian state. First, they demonstrate the limitations of the governance structures in place to overcome challenges to the national government before they occur. Second, the crises portray the narrowness of the institutional structure's ability to accommodate other groups than the ruling elite that tend to have more difficulties reaping private benefit from the national resources.

Since the institutional structure and the actual institutional practices in Nigeria concentrate the political power and the economic prosperity to the small ruling elite, groups with low level of political and economic power, such as minority populations, are overwhelmingly marginalized. This, together with fears of ethnic domination has contributed to political instability at the local level.³¹ Although since the civil war large-scale ethnic mobilization to challenge the state has not occurred, civil violence and crime have been widespread.³²

The Legacy of Colonial Institutions and Marginalization in the Sudan

In the 1820s the Sudan was invaded by the Turco-Egyptian forces of Muhammad Ali, partly because of the myths of Sudanese riches that might have fueled the expansionary campaigns of the Egyptian Crown. It took the Egyptian overlords more than a decade to effectively control the Nile River valley in the central Sudan, but after that expansion took place mainly to the west to Darfur and southwards to what was to be the southern Sudan.

Egyptian authorities integrated the Sudanese borderlands to their sphere of control through submitting them to violent extraction of slaves and resources.³³ However, this policy was not undertaken coherently enough and corruption together with misadministration resulted in growing challenges to the colonial state. As a result, Islamic nationalist

³¹ On ethnic domination see *Badru, Imperialism and Ethnic Politics in Nigeria, 1960-1996*, pp. 9-11.

³² See i.e. "Nigeria Worries U.S.", *This Day*, February 17, 2005 available at <http://allafrica.com/stories/200502170077.html>.

³³ This has been widely documented in a number of historical accounts. See i.e. Hassan (2000) and Johnson (2003).

movement of the Mahdi was able to bring down the colonial state in the 1880s and Sudan enjoyed a short period of self-rule until British invasion that led to its annexation to the British colonial system by the turn of the century.

The British colonization of Sudan was more systematically undertaken in order to avoid the Turco-Egyptian failures that had resulted in the demise of previous imperial rule. Some of the colonial institutions implemented were to be assimilated to the local customs such as taxation of pastoral populations in the South and implementation of common law legislation that could be adopted in the local legal traditions.³⁴

Unlike the earlier Egyptian masters, the British authorities were able to pacify most of the colony by 1920s. This was done by granting some regions such as Darfur, large degree of autonomy while others were subdued by force.³⁵ Once relative stability had been achieved, the colonial economic ventures were implemented in order to feed the British economy. Cotton became the main crop grown in the central Sudan to be exported to Britain. This made the central riverain Sudan the richest and most developed region in expense of the marginal lands that enjoyed only scattered development during the British period.³⁶

Once the British prepared for departure, they handed over the state control exclusively to the Arab Muslim elite of the central Sudan in expense of other populations of the highly diverse state. Although the political structures were nominally changed to accommodate a more representative form of government, the dynamics of the institutional practices and the lack of common national identity prevented the widespread participation to the Sudanese politics since it became largely Arab-Muslim dominated and predominantly sectarian.³⁷ As a result, the control of the state was left firmly to the hands of the northern Arab-Muslim elite that continues to control the state apparatus today.

By the end of the colonial rule, violence broke out in the south, a region that had been largely marginalized both in colonial and independent Sudan. Challenge to the national

³⁴ Douglas H. Johnson, *The Root Causes of Sudan's Civil Wars*, (Indiana University Press: Bloomington, 2003), pp. 11-14.

³⁵ See i.e. de Chand (2000), Hassan (2000), and Johnson (2003).

³⁶ See i.e. Markakis (1998) and Johnson (2003).

³⁷ de Chand (2000), Johnson (2003), and Melvill (2002) provide excellent overviews on political marginalization in the Sudan.

government further intensified during the 1960s before the large scale fighting finally ended in 1972, resulting in Southern autonomy. However, the regional grievances were not comprehensively addressed and after serious tampering with the southern autonomy by the Nimeiri regime, the war broke out again in 1983.

As Johnson (2003) points out, the structural root causes to civil violence in the Sudan seem to lay in the deep injustices that were created historically during the two waves of colonization. Institutions related imperialism and exploitation hence created a particular type of social hierarchy in which Arab and Muslim identity dominates other groups that are marginalized in the peripheral regions.³⁸ This social hierarchy is still reproduced in the contemporary Sudan and contributes to political, economic and social marginalization, the main characteristics of poverty and horizontal inequality.

In sum, it is plausible to argue that the institutions created in Sudan during colonialism were designed largely to secure the extraction of resources first by the Egyptians and later by the British. As a result, the patterns of how resources were extracted and distributed corresponded to the interests of the colonizing powers. The colonial authorities were therefore exclusively in control of the political power and the resource flows. Consequently, political power, wealth, and economic development within Sudan all became concentrated on the central Nile River region that became the heartland of the administrative control of the colonial authority. The periphery remained largely frontier land where large part of resource extraction and deepest poverty took place. Regional and group based inequalities were implanted during this period as historically the groups in the Sudanese periphery were violently exploited (slavery) and did not receive proportional share of neither political power nor national resources. Finally, this is at least partly, if not primarily, responsible for the political instability and ongoing challenges to the Sudanese state for the most of the post-independence period.

9 Concluding Remarks

This paper has dealt with obstacles to development in Africa. It has suggested that institutional traditions implemented in the colonial era have largely endured to the post-colonial

³⁸ See Deng (1995) and Lesch (1998) for detailed accounts on the formation of social hierarchy, identity differences and conflict in the Sudan.

period and contribute to the widespread poverty, inequality and political instability. This has been the case in Nigeria and Sudan, which are both controlled by post-colonial elite through authoritarian governance structures that create a particular perception of political power and guarantee political elite domination of national resources.

In his *Nation, State, and Economy* (1919), Ludwig von Mises has given valuable insights to political repression, inequality and minority rights. He views authoritarian state structures as impediment to liberal nationalism and individual rights, while noting that they are often controlled by privileged minorities unwilling to let go of political power for common good. He perceived this sort of repression and political marginalization largely unjust and economically inefficient, and foresaw the European exodus from Africa due to rising costs of colonial administration. Finally, although Mises does not view proportional representation as adequate political strategy to secure minority rights, he condemns the repression by one group over others and denounces assimilation efforts often undertaken by majorities to homogenize the society.

Finally, the framework suggested in this paper attempts to link colonial institutional legacy with poverty and inequality, ethnic mobilization, and political instability. It argues that understanding African underdevelopment today requires close examination of colonial institutions and their enduring tradition dictating how political power is perceived and how it affects the distribution of national resources. This seems to be the case in Nigeria and the Sudan, where colonial roots have grown into post-colonial crises.

Bibliography

- [1] Acemoglu, D., S. Johnson, J. Robinson (2004). "Institutions as the Fundamental Cause of Long-Run Growth" NBER Working Paper No. W10481.
- [2] Acemoglu, D., S. Johnson, J. Robinson and Y. Thaicharoen (2003). "Institutional causes, macroeconomic symptoms: volatility, crises and growth", *Journal of Monetary Economics* 50, 49-123.
- [3] Acemoglu, D., S. Johnson, and J.A. Robinson (2002). "Reversal of Fortune: Geography and Institutions in Making of the Modern World Income Distribution." *Quarterly Journal of Economics*, 117, 1231-1294.

- [4] Acemoglu, D. and J. Robinson (1999). "Political Origins of Poverty and Prosperity", unpublished manuscript, Massachusetts Institute of Technology.
- [5] Alavi, H. (1972). "The Post-Colonial State", *New Left Review* 74, 59-82.
- [6] Ashwe, C. (1986). *Fiscal Federalism in Nigeria*. Centre for Research on Federal Financial Relations (Canberra: Australian National University).
- [7] Badru, P. (1998). *Imperialism and Ethnic Politics in Nigeria, 1960-96*. (Asmara: Africa World Press).
- [8] Chong, A. and M. Gradstein (2004). "Inequality and Institutions", Working Paper 506, Inter-American Development Bank, Washington DC.
- [9] Collier, P. (2000). "Economic Causes of Civil Conflict and their Implications for Policy", The World Bank, Washington DC.
- [10] Collier, P. and A. Hoeffler. (1998). "On Economic Causes of Civil War", *Oxford Economic Papers* 50, 563-73.
- [11] Cramer, C. (2003). "Does Inequality Cause Conflict?" *Journal of International Development* 15, 397-412.
- [12] De Soysa, I. (2000). "The Resource Curse: Are Civil Wars Driven by Rapacity or Paucity", in Mats Berdal and David M. Malone (eds), *Greed & Grievance: Economic Agendas in Civil Wars*. Boulder: Lynne Rienner.
- [13] De Chand, D. (2000). "The Sources of Conflict Between the North and the South in Sudan", conference paper, presented at the Fifth International Conference of Sudan Studies, University of Durham, August 30-September 1.
- [14] Deng, F. (1995). *War of Visions: Conflict Identities in the Sudan*. (Washington, DC: The Brookings Institution).
- [15] Easterly, W. and Levine, R. (2003). "Tropics, germs, and crops: how endowments influence economic development", *Journal of Monetary Economics* 50, 3-39.
- [16] Fearon, J. D. and D. D. Laitin. (2000). "Violence and the Social Construction of Ethnic Identity", *International Organization* 54, 845-877.

- [17] Gaffney, P. (2000). "Burundi: The Long Sombre Shadow of Ethnic Instability", in E. W. Nafziger, F. Stewart and R. Väyrynen (eds) *Weak States and Vulnerable Economies: Humanitarian Emergencies in the Third World* (Oxford: Oxford University Press).
- [18] Glaeser, E., J. Scheinkman and A. Shleifer (2003). "The Injustice of Inequality", *Journal of Monetary Economics* 50, 199-222.
- [19] Hassan, A. I. (2000). "The Strategy, Responses and Legacy of the First Imperialist Era in the Sudan 1820-1885", conference paper, presented at the Fifth International Conference of Sudan Studies, University of Durham, August 30- September 1.
- [20] Ignatieff, M. (1997). *The Warrior's Honour: Ethnic War and the Modern Conscience*. (New York: Henry Holt).
- [21] Johnson, D. H. (2003). *The Root Causes of Sudan's Civil Wars*. (Bloomington: Indiana University Press).
- [22] Keen, D. (2000). "Incentives and Disincentives for Violence", in Mats Berdal and David
- [23] M. Malone (eds), *Greed & Grievance: Economic Agendas in Civil Wars*. Boulder: Lynne Rienner.
- [24] Keen, D. (1997). "A Rational Kind of Madness", *Oxford Development Studies*, 25, 1, 67-76.
- [25] Lesch, A. (1998). *The Sudan: Contested National Identities*. (Bloomington: Indiana University Press).
- [26] Markakis, J. (1998). *Resource Conflict in the Horn of Africa*. (London: SAGE Publications).
- [27] Melvill, D. (2002). "Restoring Peace and Democracy in Sudan: Limited Choices for African Leadership", IGD Occasional Paper No. 34, November.
- [28] Mises, L., von. (1949). *Human Action, A Treatise on Economics*. (New Haven: Yale University Press).

- [29] Mises, L., von. (1919). *Nation, State and Economy*. Translated English edition by Leland B. Yeager (New York University Press, 1983). Available online at <http://www.mises.org/nsande/nse.pdf>
- [30] Mozaffar, S., J. Scarritt and G. Galaich (2003). "Electoral Institutions, Ethnopolitical Cleavages, and Party Systems in Africa's Emerging Democracies", *American Political Science Review* 97, 3, 379-390.
- [31] Munene, M. "Culture and the Economy: The Creation of Poverty", Kenyatta University Culture Week Seminar, September 20, 2001.
- [32] North, D. (1991). "Institutions", *Journal of Economic Perspectives* 5, 1, 97-112.
- [33] North, D. (1981). *Structure and Change in Economic History*. (New York: W. W. Norton).
- [34] Reno, W. (2004). "The Empirical Challenges to Economic Analyses of the Conflicts", conference paper presented at *The Economic Analysis of Conflict: Problems and Prospects*, Washington DC, April 19-20.
- [35] Reno, W. (2003). "'Resource Wars' in the Shadow of State Collapse", conference paper presented at *Resource Politics and Security in a Global Age*, University of Sheffield, June 26-28.
- [36] Reno, W. (2002). "Economies of War and Their Transformation: Sudan and the Variable Impact of Natural Resources on Conflict", conference paper presented at *Money Makes the War Go Round: Transforming the Economy of War in Sudan*, Brussels, June 12-13.
- [37] Robinson, J. (1999). "A Political Theory of Underdevelopment", unpublished manuscript, University of California, Berkeley.
- [38] Sonin, K. (2003). "Why the Rich May Prefer Poor Protection of Property Rights", *Journal of Comparative Economics* 31, 715-731.
- [39] Stein, H. (1994). "Theories of Institutions and Economic Reform", *World Development* 22, 12, 1833-1849.

- [40] Stewart, F. (2002). "Horizontal Inequalities: A Neglected Dimension of Development", Working Paper Number 81, Queen Elizabeth House, Oxford University.
- [41] Stewart, F. (2000). "Crisis Prevention: Tackling Horizontal Inequalities", *Oxford Development Studies* 28, 3, pp. 245-262.
- [42] Suberu, R. T. "The Struggle for New States in Nigeria, 1976-1990", *African Affairs* 90, 1991, 499-522.
- [43] UNRISD (2004). "Ethnicity, Inequality and the Public Sector: A Comparative Study".
- [44] United Nations Research Institute for Social Development (UNRISD) report, April.
- [45] Uvin, P. (2000). "Rwanda: The Social Roots of Genocide", in E. W. Nafziger, F. Stewart and R. Väyrynen (eds) *Weak States and Vulnerable Economies: Humanitarian Emergencies in the Third World* (Oxford: Oxford University Press).
- [46] Van de Walle, N. (2001). *African Economies and the Politics of Permanent Crisis, 1979-1999*. (Cambridge: Cambridge University Press).
- [47] Veblen, T. (1919). *The Place of Science in Modern Civilization and Other Essays* (New York: Huebsch).
- [48] *The New Shorter Oxford English Dictionary*. (1993). Volume 1. Edited by L. Brown (Oxford: Oxford University Press).

New Perspectives on Political Economy
Volume 1, Number 1, 2005, pp. 61 – 82

End Of The Warriors

Christian Michel*

Abstract: A primitive society is characterised by the absence of a state. As soon as the state emerges, human societies become divided into castes. Anthropologist Georges Dumézil discerned three such castes in all Indo-European societies: priests, warriors and producers. This paper seeks to understand why the producer caste was always considered the lowest of the three in terms of prestige, despite being the most numerous and arguably the most useful. Producers embody the values of life and nature; warriors are on the side of culture, they must resist the natural urge to flee in the face of mortal danger. The producer acts out of self-interest, the warrior does what is right. The debt owed by society to those who accept to lay their lives for its protection is infinite. It cannot be repaid in the producers' currency (money), but only in terms of prestige and power. But in accomplishing their mission, warriors must resort to all the methods forbidden to producers, killing, deceiving, coercing. Warriors were kept outside of society, even physically, in barracks and camps, so that their values would not infect the producers caste, nor would the bourgeois values of comfort, family life, and legitimate fear of death diminish the warriors' morale. The state bureaucracy today has usurped the debt owed by society to its warriors. Albeit bureaucrats are hardly at risk of their lives, they claim to have become our protectors (against unemployment, illness, old age...) and they have found new wars to wage against drugs, poverty, crime and terror. They claim the moral high ground over producers, continuing the division of society into castes that primitives resisted for so long.

* www.liberalia.com, e-mail: cmichel@cmichel.com

1 Priests, warriors, producers

Primitive societies have no experience of the state.¹ They have no use for it. This absence does not mean, as naïve propagandists for democracy would have it, that these societies live under the thumb of a despot, an all-powerful chief who could well proclaim, following the French king Louis XIV, “The state is me”. From the Inuits of the far North to the Aborigines of Australia, primitive societies generally do not appoint a chief. When they do, the chief reigns but does not govern. He only symbolizes the group’s unity and its independence from other communities. The chief does not even exercise power in hunting and war. A primitive society’s army is not made up of orders and counter orders; it is a group of irregulars. Insubordination is characteristic of such societies. If the chief wants to play chief, he is ostracized. If he persists, he is killed.²

Primitives are fiercely attached to the idea of political and economic equality, according to anthropologist Pierre Clastres, who has devoted a great deal of his work to this subject. Every individual has special qualities – one is a skilled hunter, another a fearsome fighter – but no type of prowess, even if it bestows prestige, can ever confer power. Primitive society rejects the rift between dominant and dominated, governing and governed, master and subject. Now, the state is the instrument of this fracture. It is the locus of power par excellence. All societies structured by the state find themselves irreparably and deeply divided between those who control its apparatus and those who are its subjects.

¹ Pierre Clastres (*La Société contre l'État*, Paris, Plon, 1974; English trans. *Society Against the State*, Basil Blackwell, Oxford, 1977), even believes this can serve as a definition: The “primitive” society is one that has no experience with the state. “Primitive” can refer to groups of humans that died out during prehistoric times as well as those that lead the same existence today as they did 30,000 years ago.

² Pierre Clastres, “Liberté, Malencontre, Innommable” in *Recherches d'anthropologie politique*, Paris, le Seuil, 1980.

Primitives want nothing to do with such division.³

How then did the state come about? How did we move from the political equality of primitive societies to relationships characterized by class and dominance? The best known theory attributes the state's origins to economic factors, a theory that was evidently popularised by Marxists.⁴ The development of agriculture required deforestation, irrigation and the construction of lofts to store the harvest. In order to protect their investments from pillage, agricultural societies used part of their surplus production to maintain a corps of professional warriors. Such an approach never fails to be dangerous: those guarding against external aggression eventually came to guard an imprisoned, servile population.

Another theory views politics, rather than economics, as the founder of the state.⁵ Within certain societies, a group of priests developed what modern language calls an "ideology". This ideology no longer identified the group solely as a descendant of a mythic ancestor or in terms of totemic membership. Such a group could then subjugate

³ If I were to offer my own explanation for this refusal, I would start with the small size of the Primitives societies. While an ethnic group may easily number in the tens of thousands, the social organization of primitive peoples is characterized by clans numbering 200 to 300 members, sometimes only a few dozen. However, envy is no different among Primitives than among our contemporaries. The existence of wealthy people in Malibu or Monte Carlo is an abstract fact, we can rationalize the resentment we feel about it with "social justice" concepts. But if our cousin or co-worker gets rich, we have a much stronger emotional reaction. It is only when societies become large enough to create impersonal social relations that individual situations can become differentiated. Certain people institutionalise their power and accumulate wealth without encountering too much hostility. Many theoreticians of the state, such as Aristotle and Rousseau, view small societies as the ideal political unit. That notion, however, ignores the fact that conformism and resentment are often the natural complement to warm, personal relationships. "Keeping up with the Joneses" refers to the Joneses living down the street. We do not identify as well with the other Joneses. They become statistics that do not arouse the same feelings of envy and resentment. The obvious reference here is to René Girard's work, devoted to "mimetic rivalry," particularly *Mensonge romantique et Vérité romanesque*, Paris, Pluriel, 1978, and *La Violence et le sacré*, Paris, Grasset, 1972; English trans. *Violence and the Sacred*, Baltimore, Johns Hopkins University Press, 1979.

⁴ Friedrich Engels, *Origins Of The Family, Private Property And State*, International Publishers Company, 1990, available at <http://www.marxists.org/archive/marx/works/1884-fam/index.htm>. For examples of Marxist anthropology, see also V. Gordon Childe, *Man Makes Himself*, written after his enthusiastic trip to the Soviet Union in 1936 and Morton Herbert Fried, *The Evolution Of Political Society*, New York, McGraw Hill Higher Education, 1967.

⁵ Fustel de Coulanges, *La Cité antique*, Paris, Champs Flammarion, 1998, Volume 3, chapter 3. English trans. *The Ancient City*, Trans. William Small, Garden City, Doubleday Anchor Books, 1964. Elman R. Service, *Origin of the State and Civilization*, New York, W.W. Norton, 1975. See also Elman R. Service, *Political Power and the Origin of Social Complexity*, in *Configurations of Power*, edited by John S. Henderson and Patricia J. Netherly, Cornell University Press, Ithaca and London, 1993.

others, not for the purpose of reducing them to slavery, but to assimilate them – given that filiation was no longer a criterion for belonging to the group.⁶ This acquisition-based, rather than endogenous, growth conferred a decisive numerical advantage both in war and the construction of infrastructure. The same ideology served not only to determine the society’s “foreign policy,” but also to justify the power of its creators – the priests and warriors – in their dominant caste positions.

2 The three orders

Whatever their relevance, these theories focus on three categories of players:

- Priests, vested with spiritual duties, whether they are magicians, shamans or prophets.
- Warriors
- Producers of wealth, i.e. all those who do not belong to the two previous classes and who carry out the work required by society.

Since George Dumézil published his works, we now know that this classification existed in all Indo-European societies.⁷ From Ossetians in the Caucasus to Vikings, from Greeks and Romans to Irish Celts and the many societies in India, Iran and the Slavic lands, all these societies were based on this tripartite model: priests, warriors and producers. Dumézil explains that this model does not claim to describe each society’s reality, but the way in which the society represents its own reality through myths, legends and epics. This fact only strengthens the central paradox: Why would producers place themselves

⁶ This required a true paradigm shift. Totemism is as universal among primitive peoples as the absence of the state. Claude Lévi-Strauss explained its significance: “Saying that clan A descends’ from bears and that clan B descends’ from eagles is nothing more than a concrete and abbreviated way of viewing the relationship between A and B as analogous to the relationship between species”. (Claude Lévi-Strauss, *Le Totémisme aujourd’hui*, Paris, PUF, 1962). Agreeing to assimilate individuals from another clan thus marks the emergence of a new, broader conception of the human being, which would continue to expand in waves up until the universalism of today.

⁷ George Dumézil, *Heur et malheur du guerrier*, Paris, Flammarion, 1985 (English trans. *Destiny Of Warrior*, Chicago University Press, 1971) and *Mythes et dieux indo-européens*, Paris, Flammarion, 1992. For an introduction to the work of Georges Dumézil, see Wouter W. Belier, *Decayed Gods, Origin And Development Of Georges Dumézil’s “Idéologie Tripartite”*, Brill Academic Publishers, 1991.

at the bottom of the social ladder when representing their world? After all, they are far greater in number than the other orders, especially if women are counted. Women rarely serve as priests and almost never as warriors. And they symbolize fertility, which by definition characterizes producers. In terms of what is useful for society, the occupations of farmer, carpenter, sailor or banker are essential. Shouldn't those who practice these occupations be honoured in literature and art? We know that not to be the case. Heroes are usually warriors and sometimes saints and artists. Villains are in business. What could explain such disrepute?

Free-market proponents, even less so than others, do not have the answer. They believe that all human behaviour is motivated by self-interest. While they are not wrong in principle, they have the tendency to measure self-interest only in terms of monetary profit and loss. Very few people, however, make their decisions solely on the basis of this one criterion.⁸ Many other considerations are at play. The following flowery lines, which come from a book published in the United States in 1995, *Ethics and Public Service*, effectively illustrate the bankruptcy of the *Homo oeconomicus* model when it attempts to explain human behaviour:

“Man’s feet may wallow in the bog of self-interest, but his eyes and ears are strangely attuned to the call of the mountaintop. There is a distinction between “I want this because it is in my self-interest” and “I want this because it is right.” Man’s self-respect is in large part determined by his capacity to make himself and others believe that self is an inadequate referent for decisional morality. This capacity of man to transcend, to sublimate and to transform narrowly vested compulsions is at the heart of all civilized morality”.⁹

The author contrasts taking action “because it’s the right thing to do” with acting out of self-interest. It is commonly agreed that personal gain motivates producers, merchants and capitalists, while doing what is right is the *raison d’être* of public service. And the heart of public service is the army, the warrior class.

⁸ “Intelligent beings never base their goals mainly on economic factors. In the proper sense of the term, our actions are not ruled by economic motives’. There are simply economic factors that influence our efforts to satisfy other goals”. Friedrich Hayek, *The Road To Serfdom*, University of Chicago Press, 1994; chap. VII (with an introduction by Milton Friedman).

⁹ Stephen Bailey, *Ethics and Public Service*. The quotation is found in James H. Tower, *Truth, Faith and Allegiance*, University Press of Kentucky, 1995.

3 Death and the warrior

At the core of the power relationship is the debt relationship. Producers get paid. They get paid, in money, the full cost of their work; if that were not the case, they would find other work. That is the basic premise of free-market economic theory. Once service is rendered and payment made, the legal relationship ends and no one has the right to demand anything from the other party. But how do we pay the warrior? Here is a man willing to sacrifice his own life to save yours, save the lives of your loved ones and protect your property from plunder and destruction. Should you pay for his services in dollars or ounces of gold? How can one ever completely repay him for such sacrifice? Homo oeconomicus goes bust. The debt never ends. What we owe the warrior can only be expressed in the intangible currency of prestige and power.

Their relationship to death, therefore, seals the respective social status of the warrior and producer. The producer walks on the riverbank of life, on the side of nature and biology. Like all creatures, he is driven by “the force through which things persevere in their being”, as Spinoza said.¹⁰ The warrior, however, strides on the other side, that of culture. He has chosen the riverbank of death. Biology programs us to beget children and live old enough to see them reproduce in turn. But culture can replace this instruction by convincing some of us that it is glorious to be killed in battle at the peak of youth.¹¹ Thus opens the fracture separating the warrior from the rest of society. The warrior will always scorn the producer, the bourgeois, because the bourgeois fears death.¹²

¹⁰ “Life is the force through which things persevere in their being”. Spinoza, *Descartes’ Principles of Philosophy*, in *Collected Works of Spinoza*, vol. 1, collected and translated by E. Curley, Princeton University Press, 1985.

¹¹ The Romans taught: “*Dulce et decorum est pro patria mori*”. The barbarian is the man who boasts about killing. He crows about piercing his enemies with his spear and carries their scalp on his belt. (Our modern pilots paint a roundel on the cabin of their airplane, but they do so to display the *planes* they have brought down, not the enemies they have killed. This distancing is not mere hypocrisy because the enemy pilot could have escaped by parachuting out). In societies civilized by Christianity, the soldier embodies the sacrificial victim giving up his life for the king, the fatherland or a cause rather than an idealization of the killer. The paradoxical result is that our civilized commanders did not hesitate to send far more men to a certain death than the barbarian chieftains would ever have dared. This conception of the sacrificial soldier peaked during the period between the Napoleonic wars and the Korean War with the butchery of the American Civil War and the two world wars. (We, Westerners, are less accepting of sacrifice; what cause today could be important enough to die for?). As it becomes more individualistic, more liberal and more capitalistic, our society becomes that of Eros gradually triumphing over Thanatos.

¹² See my essay, *How Should We Think About Economics Today?*, www.liberalia.com

The producer and the bourgeois are on the side of biology, which the warrior has apparently left behind. But isn't it precisely biology that dictates certain altruistic behaviours and sacrifices? Sociobiologists believe so; they argue that giving up one's life is sometimes the only way to ensure the survival of one's descendents and the perpetuation of one's genetic heritage.¹³ This theory, however, only applies to warriors from clans and from tribes related by blood. It cannot explain the accepted deaths of those defending the diverse populations of historical empires and modern states.

Nor does it seem credible to reduce soldiering to an expression of the aggressiveness that males apparently carry in their genes. War is not a series of fistfights. When projectiles (throwing sticks, blowguns, bows and arrows) made their appearance very early in prehistoric times, the hasty rage that made the hand tremble became a handicap. Those who felt such rage in combat did not live long. Modern warfare, which mobilizes all the resources of advanced technology, is even more dependent on cool heads and methodical and deliberate action. It is difficult to imagine an activity that is more rationalised and socialized than war.

We all die, of course, both producers and warriors, but we do not die the same death.¹⁴ Peasants, the bourgeoisie, women, you and I all die *from* something; we die *from* old age, *from* accidents, *from* illness. But the warrior dies *for* something; he dies *for* his king, *for* the fatherland, *for* the revolution. Bourgeois death is a simple link in the chain of events caused by biology and the whims of nature. It is history. It is history's shapeless, monochromatic pattern. The warrior's death, however, *makes* history. It offers history the adventurous and unexpected. Francis Fukuyama could thus write that the disappearance of warlike empires would bring about the end of history.¹⁵ The military function, which defies nature, reflects the human conscience's deliberate stand against biological evolution – which, in the final analysis, is the gist of history. For that reason, the death of warriors always takes on a grandiose and tragic dimension.

¹³ R. Dawkins, *The Selfish Gene*, Oxford, Oxford University Press, 1976.

¹⁴ Oswald Spengler, *Decline Of The West*, Oxford University Press, 1990.

¹⁵ Francis Fukuyama, *The End Of History And The Last Man*, London, Penguin Books, 1992.

4 The warrior's curse

From the beginning of history, the warrior class had special status in society. Plato demanded it, and chivalry exemplified it.¹⁶ The isolation of the warrior class from the rest of society was so strict that not only was access restricted – it was hereditary in most cases – but the warriors themselves could not assume any other role. Such a prohibition should surprise us. It is consistent that a privileged caste, living entirely off levies on the country's economic activity, would restrict membership. But why should they have been forbidden from lowering themselves by choosing another occupation if they had been insane enough to do so? Graduates of prestigious universities generally do not aspire to be refuse collectors and doctors typically have no desire to become nurses, but no law forbids them. In fact, if more college graduates were to work at jobs below their skill level, competition at the top would be less intense.

Historians offer many reasons for this segregation of the warrior class. One such reason, which may seem obvious, appears not to have been recognized: The moral values of warriors differ from those of producers. Physical courage brings honour to the warrior, but should be completely useless to a producer in a well-ordered society. The audacity to kill without remorse is required of the warrior, but obviously forbidden to the producer. Warriors are loyalists; producers are loyal. Soldiers are praised for using the types of ruses and traps for which capitalists are so reproached. All the great generals have won their laurels and gained respect and admiration for their ability to kill, abuse and deceive.¹⁷

Of course, they always claim a good cause, which results in the warrior's curse. To be good, he must be bad. To accomplish his mission of defending society, he must resort to

¹⁶ Plato, *The Republic*, trans. Desmond Lee, London, Penguin, 1976 and *The Laws*, trans. Trevor Saunders, London, Penguin 1972.

¹⁷ Mo Ti, a Chinese man of letters who wrote the following around 400 B.C., condemned this inverted warrior morality: "If a man kills an innocent person and steals his clothes, spear and sword, he commits a more serious crime than if he entered a stable to steal an ox or horse. The wrong is greater, the offence more serious and the crime blacker... But we see nothing wrong with committing a murder when attacking a country; we applaud and speak of justice... When a man kills another, he is guilty and sentenced to death. Therefore, according to the same criterion... , he who kills 100 men should suffer far greater punishment... Similarly, if a simple homicide is considered a crime, but multiple homicide, such as occurs when another country is attacked, is praised as good, can that be called knowing good from bad?" (quoted by S.B. Griffiths in his introduction to Sun Tzu, *The Art Of War*, Oxford University Press, 1972).

all the methods society condemns. We can therefore contrast, item by item, the morality of warriors with the morality of producers. Producers act out of self-interest; there are no higher values for them than biology and nature (their life and the lives of their offspring). However, the time-tested method to achieve these values is cooperation with others, and cooperation's golden rule is, "Do unto others as you would have others do unto you".¹⁸ Warriors, however, do not defend their own interests (what type of self-interest could possibly motivate a person who has accepted death?). They expect no cooperation from others, "others" being whoever happens to be the enemy of the day. They do not gain what they want through negotiation and cooperation, but through conquest. And they especially do not want the enemy to do to them what they are doing to the enemy.

5 The "raison d'État"

This complete reversal of values within a society would be impossible without the construct, however cultural and artificial it may be, that we call the State. It is solely the *raison d'État* that makes existence possible according to the warrior's moral code.¹⁹ If this reverse morality were to spread beyond their closed caste, if murder, trickery and deception were to become the values of producers, who represent the vast majority of society, the very process of civilization would fail. Historically, the isolation of such unnatural behaviours was thus necessary to the community as a whole. Limiting them to a specific group, a closed caste like the hereditary nobility, was in the interest of the producers themselves. So the caste system did not function solely by restricting access to the ranks of the military aristocracy, but also by prohibiting these same nobles from

¹⁸ In a seminal book that has been justly acclaimed, Robert Axelrod, (*The Evolution of Cooperation*, Basic Books, New York, 1984) demonstrates that the most beneficial long-term behavioural model consists of treating others as they treat us; he bases his argument on game theory and the famous prisoner's dilemma. If it seems judicious to apply this model to everyday circumstances, it can be contrasted with René Girard's analyses in situations of serious conflict. (op. cit. and *Des Choses cachées depuis la fondation du monde*, Paris, Grasset, 1978, English trans. *Things Hidden Since The Foundation Of The World*, Stanford University Press, 1987).

¹⁹ There were warriors well before the appearance of the state. Primitive societies are the most warlike of all. But the very basis of my argument throughout this text is the following: In primitive societies, *all the men* are warriors and producers in turn. Violence and predation are committed against the outside enemy, not the members of the tribe. In modern societies, however, the dominant class of soldiers and civil servants exploits its fellow citizens rather than foreigners.

doing productive work – literally making work *ig-noble*.²⁰ Their “criminal” morality ran too great a risk of corrupting the entire social fabric.

But the reverse is also true. The warrior’s existence is so contrary to biology that he is always at risk of life-affirming values taking the upper hand. Nothing in our genes compels us to leave our home and family and go off to kill other members of our species. On the contrary, our genetic programming instructs us to flee danger. The soldier therefore must be subjected to constant discipline and kept apart from the rest of society to switch off these biological instructions. Let’s not forget that throughout history, the soldier’s daily lot was not war, with its rushes of adrenaline, but *preparation* for war. As a result, the endlessly repeated military exercises, the routine manoeuvres, the marches in quick time, the shared meals, the chants, the drills, the hard discipline of the barracks were nothing other than interminable training – similar to what an animal must be subjected to when made to act in unnatural ways.

Isn’t it the same for the other dominant caste, the priests? In a completely different setting, but for the same reasons, priests accept discipline that restrains, if not breaks, their natural impulses. With its rules and rituals, the convent is not that different from the barracks. Their common stated goal is to distance themselves from the world’s temptations. And what is more corrupting than bourgeois life – sex, family, comfort, money...? Any abbot is well aware of it, and any conqueror knows that the “delights of Capua” represent the greatest danger on the path to his triumphs.

6 The intellectuals

A society in which warriors attempt to maintain the power of their caste and their ability to monopolize wealth must limit the influence of producers and of women. That is the task traditionally given to intellectuals, formerly the clergy and today the masses of

²⁰ Philippe du Puy de Clinchamps, *La Noblesse*, Paris, Que Sais-Je?, 1962. The occupations not considered suitable for French nobles included all the mechanical trades, even at the management level; farming leased land (except for land belonging to the king and royal princes); all types of trade; and low-level offices like notary, bailiff and prosecutor. After Colbert, the temptations of wealth began to corrupt the nobility and sea trading and the growing foundry industry were exempted from these restrictions. But it was only after World War I that the nobility finally felt free to embrace any type of profession.

teachers, scientists, artists and journalists that receive salaries from the State.²¹ With the influence conferred on them by a microphone or university chair, they make every effort to discredit the bourgeois values of productive cooperation and glorify predation. They believe that living off taxes is more honest than being paid by satisfied customers; if the cause is just, all means must be made available to ensure its victory.

As long as the entire society is measured against these perverse values, the warrior takes pride of place. He does “what is right” while others drag their feet “in the mud of self-interest”. Such acknowledged moral superiority leads to the exploitation of discredited producers while saving the dominant class from constant and costly recourse to violence. If society’s values should change and society lets itself be guided by the values of production, the warrior will lose the image of saintliness, of one who has “renounced self-interest”. How could he then justify continuing to exploit producers?

7 The warriors’ usurped heritage

This might be all very interesting – at least, the author hopes it is – but how does it concern us? Well, in more than one way, but the greatest source of our concern should be that modern governments and their bureaucracy base their moral superiority on the prestigious heritage of the military class.²²

Gradually, in the 1930s, with the New Deal in the United States, the Beveridge plan in Great Britain and the rise of social democracies throughout Europe, governments succeeded in portraying their image as that of *protector*. According to their slogans, they would save us from the scourges of unemployment and social inequality, the rapacity of multinational corporations and mafias, and the encroachment of foreign cultures. But

²¹ Ludwig von Mises, *The Anti-Capitalistic Mentality*, Princeton, New Jersey, D. van Nostrand Company, 1956. See also in Robert Nozick, *Socratic Puzzles*, Harvard University Press, 1997, the chapter entitled *Why Do Intellectuals Oppose Capitalism?* How many brilliant minds of that time challenged the divine right of kings, aristocratic privilege and serfdom? How many dare today to attack the privileges of the state bureaucracy, who will seem as abusive to our descendants as Ancien Régime class structure does to us? Intellectuals do not care much for freedom; they traditionally align themselves with whoever is in power, whoever pampers them or from whom they expect even greater favours.

²² Saying “my country” brings to mind the memory of landscapes and the sounds of a language, but if this country is a state or aspires to be one, we also recall the piously learned names of bloodied fields where young men who fought for it lie buried. Barbara Ehrenreich, *Blood Rites, Origins And History Of The Passions Of War*, Metropolitan Books, 1997.

isn't protection the soldier's duty par excellence? Now that the threat of military invasion has disappeared in the West, the soldier has permanently yielded the function of Great Protector to the bureaucracy. It is fascinating to observe how State employees cloak themselves in the quasi-religious mantle of prestige and respect that society has always conferred on its soldiers.

What official ceremony would do without a military parade, a band leading the way? The inauguration of heads of State, the unveiling of monuments, the celebration of national holidays and visits by foreign dignitaries all take place in front of an honour guard. Flags and national anthems irresistibly evoke the military history of the country. A modern state is an institution of living pseudo-soldiers governing in the name of real dead soldiers. Every State institution, crowned with the glory of its heroes, hijacks the debt owed by society for the blood shed by its soldiers. As repayment, government employees demand the right to act according to inverted military values. They always find a war to fight, 'war on drugs', 'war on poverty', 'war on crime', 'war on terror', that exonerates them from breaking their public commitments (electoral or other), stealing money through taxes, spying, cheating, censoring and using armed violence – aggressions directed not at a foreign enemy but their fellow citizens.

Producers are part of a web of cooperation; their peers forcefully call them to order as soon as they stray. But government employees do what producers dare not do, giving themselves permission in the name of cultural values: "public service", the "common good", the "higher interest of the state", "social justice", etc. If these causes are not strong enough, they invent wars'; witness this military vocabulary applied to all sorts of situations: war on drugs'; war on crime'; war on poverty' even. After all, the bureaucrat reasons, I'm not acting out of self-interest (right?); the means are just because I'm serving a just cause. How can petty considerations like individual rights and respect for privacy and property be raised against this state that I represent and for which people have faced death?

Government has grown to manage all aspects of our existence. As a result, humanity faces the danger of seeing the inverse morality of the military infect society as a whole. It was precisely this danger that the caste division, present throughout history, was designed to prevent. Our government officials forget that the warrior was being consistent in his refusal of life-affirming values. To him, the test of devotion to "public service" was

the supreme sacrifice. His heroic death, he thought, would atone for his violations of common morality. While claiming their right to be predators, state bureaucrats and pen pushers take no risks – not even the risk of losing their cushy jobs. They are valets dressed up in the clothes of their masters.

8 Tribute

A society in which a corps of soldiers holds the legal monopoly on violence, and is financially maintained by the masses of producers, is the social organization that we call “civilization”, or “political society”. The two words derive from the same root – the concept of “citizen”: *civis* in Latin and *polites* in Greek. Societies that reject politics and the division it creates between dominant and dominated, in which each man is a warrior and no man a chief, are called “savage” and “primitive”.²³ We have erected a deplorable epistemological barrier, as if civilization’s benefits, which distinguish us from “savages”, would be unimaginable without this social fracture. It is as if exploitation were a prerequisite for prosperity, extortion for justice, police power for establishing peace and letting art flourish.

At the core of the power relationship is the debt relationship, as we have noted in reference to Pierre Clastres. But the nature of society changes as the “direction” of debt changes. If the debt relationship trends toward society and away from the chief, as in primitive societies, society remains undivided. Those who enjoy the prestige gained from chieftainship – including the regalia of office, distinctive tattoos, special finery, not to mention women’s esteem expressed in sexual favours – must pay. This reciprocity does not shatter society’s homogeneity, nor does it involve any submission or breakdown into classes. “You want prestige? How much are you willing to pay us for our show of respect?” Political power, on the other hand, is established when the debt relationship is reversed, when payment originates in society and moves upward towards government. At first, political power was exercised over those who were *outside* society, those who had been conquered. Subjugating the alien meant imposing tribute. Then the state emerged. And the first act of the state was to raise taxes. Raising taxes is a bizarre philosophical transmutation in which armed robbery is no longer considered a crime but an act of civic

²³ This is the essence of the entire gun control debate, which so captivates Americans.

virtue.²⁴ The alien, subjugated and subject to taxes, is now located *within* society. The reverse morality of the warrior infects the social fabric itself, now split between dominant and dominated, exploiters and exploited. The state institution opens the type of social rift that was so fiercely rejected by primitive peoples and turns power against society itself.

9 Political society, warrior society

The strict egalitarianism valued by primitive societies prevented any type of progress.²⁵ Weakened by this immobility, most such societies died out. Progress involves continually adapting to evolution. If there is such a thing as evolution – and this does seem to be an accepted fact – humans have two ways to respond. They can decide to continually reform their collective organization or determine that it is the best possible and that nothing must change. If they choose the latter, their society will gradually clash with its environment until it calls its immobility into question, belatedly and at high cost. Without change, the society will disappear, which was the fate of the Primitives and the

²⁴ When I was a high school student in Paris, during the Algerian war of independence, I knew a National Liberation Front (FLN) militant whose mission was to impose a “revolutionary tax” on Algerian students and merchants in the neighbourhood. The French police arrested him for extortion. After the Evian accords, this zealous militant raised taxes for the new Algerian state from neighbourhood students and merchants, but now the French police supported him. His extortion scheme was exactly the same, but this time it was for a recognized state. Murder, coercion and theft are not forbidden in society, but they are reserved for the class of state employees.

²⁵ The law of primitive peoples was the law of their ancestors, and therefore immutable (the dead do not change their minds). Progress began when the chief’s desires became law, and his successor then felt free to want something else.

Soviet Union, among others.²⁶

When human beings give each other the right to innovate, each person may consider the various ways of living in the world both individually and collectively. Each is free to cut a new path or to follow those who seem to be moving in the direction of his or her goals. Philosophy, science, economics and spirituality are nothing but various disciplines whose value lies in leading us to more just relations with others and nature.

In the political organization of the world, however, this relationship to nature is not built individually, but society by society. Each society imagines an overall way of functioning and its citizens must comply. Within each state, detailed laws regulate how to live, marry, raise children and care for oneself, what to produce, consume, read and view, what rules should apply to business, under what conditions people should work, how much to save, etc. The models, therefore, do not differ within each state, but only *between* states. The distressing result is that instead of having tens or hundreds of thousands of ways of living together, developed by people who have come together voluntarily in a community based on affinity, ethnicity, culture or interests, we have been reduced to comparing life in a handful of political societies, i.e. organized according to military

²⁶ Primitive peoples strived to create an autarky within their clan. Only goods considered essential, in terms of usefulness and prestige, would be traded. The Soviet Union had the same policy and for the same reasons: the “open society” of producers is incompatible with the hierarchical order of the military. Europe broke away from other societies during the Renaissance, leaving them far behind; this can be explained by its ability to absorb foreign ideas and methods, which demonstrated its enormous self-confidence. Lévi-Strauss recalls in *Tristes Tropiques* that during the first encounters between the Spanish and Carib Indians, the Indians wondered whether the Europeans were gods or men and the Spaniards wondered whether these “savages” were humans or animals. The reverse would have been unthinkable. Thomas Sowell, in his book *Conquests and Culture* (New York, Basic Books, 1998), subtly notes that the great American empires were not conquered by a handful of Spaniards, but by all the technologies that the entire Old World had developed and traded at the time: Italian ships, Arab compasses, steel from Toledo, Chinese gunpowder, English cannons. The Aztecs cannot be reproached for not participating in this wave of innovation. However, opponents of globalisation might usefully ponder the cause of the Aztecs’ collapse.

values.²⁷

The competition between political societies during the historical process of humanity's evolution has thus hinged on one single criterion for success – the power of the state, projected both internally and externally. In other words, our societies chose an organizational model that was based not on the producer's values (the most appealing, the least expensive) but on the warrior's values (the most powerful). And it was at the cost of horrifying and bloody power conflicts that today's dominant political model, social democracy, was imposed (first, colonization to eliminate primitive societies, then two centuries of all-out war over rival political systems – monarchy, fascism and popular democracies).

A multitude of social organizations would allow more people to find one that fit their values. If one such way of life harms nature, the effects would be limited by other less damaging practices. The social democratic monopoly, however, only affords us one single chance: it's make or break. And if it works this time, it has to work the next time and the next time. The more power is centralized (*a fortiori* under a world government), the more serious the consequences for humanity if one bad decision is taken. The first results do not seem to indicate, to say the least, that the emerging global political model is satisfying all human aspirations or blending in harmoniously with nature.²⁸

This begs the historically novel question: How can we pursue the evolutionary process within a social organization that is the only type authorized? The way we live in society and our adaptation to the environment are not hardwired in our genes. Like other kinds of knowledge, they must be discovered through trial and error. If governments forbid such experimentation, if they do not step aside for communities offering

²⁷ It would be more accurate to emphasize that throughout history, societies that were the most effective at waging war were those that gave the greatest respect to producers. This is only a superficial paradox. There is no doubt that danger and war taught human beings cooperation. Those who knew how to cooperate through good communications, advanced language and the acceptance of responsibility physically eliminated or at least drove out less developed tribes from lands well-stocked with game. During this historical period, societies in which confidence inspired investment, respect for other people's word favoured trade, and the state's predation did not totally destroy wealth, were able to equip the most fearsome armies. The more producers' values are respected *within* society, the better warriors' values can be expressed on the outside.

²⁸ In the absence of rival models, the caste of state employees has the wherewithal to hide its mistakes for a very long time. It suppresses the incriminating information, including by use of taxation and subsidies, which are nothing more than a form of censorship applied to the market.

other ways to live, aren't we just repeating the lethal immobility of primitive societies?

10 The end of political societies

An ideology cannot be refuted. A political view of the world, handed down by warriors and taken up by the entire state apparatus, gets too deeply rooted in the culture and the collective unconscious. Those who have the ability to discredit it – the intellectuals – are its direct beneficiaries. An ideology goes out of fashion. It is gradually abandoned by a growing number of people when it proves unable to make sense of reality.

Aren't we today witnessing an historic failure of the state?²⁹ The increase in the number of states should not foster any illusions; it does not demonstrate their greater value, but rather the loss of their individual importance. The break-up of European and Russian colonial empires over the last half-century has brought the number of states to nearly 200, and who could tell who is more of a puppet, thief, beggar, killer or simply laughingstock than the others? As we have noted, the legitimacy of governments derives from their supposed protection of citizens. As the warriors' heirs, they need enemies just as the doctor needs patients.³⁰ In a desperate attempt to hide the fact that they *themselves* are the major danger citizens face, they must invent even greater dangers. Their imagination, however, is no longer up to the task.

²⁹ The states have done a great deal of killing, but they have also played a positive role in the evolutionary process, as impenetrable and shocking as this may be. They helped integrate and pacify their territories, and served as vehicles for education and culture. We realize today that these same benefits could have been obtained without government oppression, but were human beings ready to skip the state stage as evolution progressed? As Marx demonstrated, it is technological progress that brings about a maturing of individuals' consciences, and not the reverse. We can assert at the beginning of the 21st century that technological advances have made it impractical to manage societies according to the political models of the 18th and 19th centuries. We should soon expect a new awareness of the exploitation of producers by the state bureaucracy and its protégés.

³⁰ Humans have not always been at the top of the food chain. When faced with danger, they tend to follow the herd instinct. Danger brings together the herd, leaving the most vulnerable on the sidelines (for example, those who do not enjoy a privileged status, job and retirement security, etc.). "War is the health of the state," wrote the American poet Randolph Bourne in 1917. "It automatically sets in motion throughout society those irresistible forces for uniformity, for passionate co-operation with the Government in coercing into obedience the minority groups and individuals which lack the larger herd sense." R. Bourne, *The Radical Will: Selected Writings 1911-1918* (New York, Urizen Books, 1977). And R.L. Mencken added, demonstrating that politics is the continuation of war by other means: "The whole aim of practical politics is to keep the populace alarmed – and thus clamorous to be led to safety – by menacing it with an endless series of hobgoblins, all of them imaginary."

Drugs, mafias, terrorism, ecological and economic crises – sometimes the threat is imaginary and sometimes it is caused by the governments themselves. And when the outside threat is real, they are powerless to avert it. The modern avatar of the state, social democracy, has carved out an almost limitless market for itself – the legalization of theft, sold under the name “social justice”. What could be more appealing than a political agenda that promises: “If you vote for us, we will make sure other people pay for benefits that will not cost you any contribution, work, or worry”? Alas, Ponzi schemes do not last forever.³¹ Welfare states, like obsolete firms with saturated markets and rising costs, merge or join together in cartels, such as the European Union and NAFTA.

Such mergers, however, do not delay evolution, even though they are promoted as harbingers of a new world order. They do not protect us from reality. Human societies evolved very slowly during primitive times, when change was measured in millennia. It has now accelerated. All it took was a few decades to abolish the warrior’s ideology of the *line*: the “line” as a boundary between exploiters and exploited, public and private, national and foreign, as a social rift tearing apart all societies that are not “primitive”. Today technology surges up in unforeseen ways, upsetting hierarchies – the military organizations’ characteristic feature – and forcing the decentralization of power. Our transnational world is no longer based on “us and them”, the only structure that makes sense to the warrior. Instead, it favours mixed relations – cooperation in some areas and competition in others. For that reason, the new order no longer depends on the “line”, but on *networks*, which now organize societies of producers rather than warrior societies. That is why the time has come for “peaceful societies” to overtake “political societies”.

11 The peaceful society

Throughout our life, we come to create relationships with a few hundred people, and often far fewer. These are family members, friends, neighbours and colleagues. In a peaceful society, they are the only people we wish to know, based on love and shared

³¹ Charles Ponzi was an infamous American con man who managed to wipe out the savings of 40,000 investors in less than eight months in 1928. He promised them spectacular returns, which were, of course, paid to the first depositors with the money contributed by later investors. Others have imitated Ponzi (especially in Russia in the early 1990s), but none could reach the scale of the European states’ social “security” and pension schemes.

interests. We buy products from the rest of the world, we hear about it in the media. If we have a reason to meet one or another of the world's six billion individuals, we initiate contact. Many of us will willingly demonstrate solidarity with strangers when necessary. Most of the time, however, all we expect is non-interference; we want others to respect our property rights and to let us live in peace with the people we have chosen.

Politics on the other hand consists of prohibiting us from choosing our relationships. Governments superimpose another dimension onto the ties of friendship and common interest – that of citizenship. Relationships between citizens are not voluntary; they are forced upon them by the authorities. Citizens in a democracy do not engage in dialogues like friends or in negotiations like producers, that leave each party free to agree or to break the talks. The democratic way of interaction between citizens is through elections. Voting means adopting a method of resolving conflict that, like war, subjects losers to the will of winners. (The non-political solution consists of letting individuals do as they wish as long as they do not physically harm others).³²

Power therefore corrupts not only those who exercise it, but the entire social fabric. Obedience precludes trust amongst subjects. Each person is required to become an accomplice of the Master, an informer on his neighbour.³³ That is called civic duty. As citizens we have no other counterpart than the government. Citizens qua citizens have

³² The Nambikwara, like all non-political societies, have a good solution for preventing exploitation by the powerful: "If the chief appears too demanding, if he claims too many women for himself or if he is incapable of providing a satisfactory solution to the food problem in times of scarcity, discontent becomes manifest. Individuals or whole families will leave the group and go off to join some other with a better reputation" (Claude Lévi Strauss, *Tristes Tropiques*, London, Jonathan Cape, 1973). Another relevant observation, valid for primitive societies in general but specifically referring to the Nuer of Kenya, is the following: "This lack of centralized, coercive power allowed people in primitive society to move on and move out when they found themselves unhappy with their circumstances. In our present state system, citizenship is not voluntary; one may leave the state one lives in, but only with the compliance of another state. Among the Nuer, if a whole community fought with its neighbour and was discontent with the outcome, it had the option of moving to a different section or a different tribe and taking up residence there. An individual had the same option." Eli Sagan, *At the Dawn of Tyranny, The Origins of Individualism, Political Oppression and the State* (Boston, Faber and Faber, 1986). Any desire to secede is brutally suppressed in political societies. Secession, which was costly when wealth consisted of agricultural land that could hardly be carried off, is once again becoming a threat to states with the development of the information society. When all wealth is information, which remains stored between the ears of its producers, it becomes a highly portable commodity. See my *Libertarianism and the Information Revolution* at www.liberalia.com.

³³ The laws on money laundering serve as confirmation. They were enthusiastically passed by the democracies which, like Switzerland, came across as defenders of individual freedom for a while.

no reason to engage in social intercourse other than to establish whether they are for or against the government. They support and strengthen it, in the hope of gaining individual benefit by imposing their own political choices.³⁴ If this strategy of conquest fails, if the other party wins, the victims do not have the innocence of slaves or serfs. Dominated and exploited, they are treated as they wanted to treat others. There is no innocent citizen.

12 Reconciliation

And what if citizens freed themselves? If society decided to reject domination and rid itself of the resulting financial and human cost, what a new “breakaway” this society would enjoy! What an example it would set for others – an example that would no longer be measured in terms of military power, but in terms of creativity, the range of opportunities offered to everyone, and adaptability to the environment.³⁵

The three “orders” described by Dumézil are not imaginary. We cannot conceive of a society without an openness to spirituality, without systems of defence and without production. But this tripartite structure is more constituent of the individual than of society. Each of us needs to be a priest, warrior and producer *all at the same time*. There is no need to divide society into classes, as was done at the dawn of civilization in the way described by anthropologists and historians. There is even less reason to organize these classes into a hierarchy of powers.

Human beings belong to nature, spring from it, and assess nature’s constraints as they act upon it. They transform nature with their work. But this very ability to transform nature makes us unique and results in something that is no longer nature in its pure form. Thus arise the two great functions – producer in the realm of nature and priest in the realm of the supernatural. Only the warrior is artifice.

³⁴ One can never quote Frédéric Bastiat’s definition often enough: “The state is the great fiction through which everyone strives to live at the expense of everyone else”. (*L’État*, essay published in the *Journal des Débats*, 25 September 1848 edition, available on the excellent Web site devoted to the great French economist: www.bastiat.org)

³⁵ Not counting the stocks of chemical, bacteriological and nuclear weapons accumulated by governments, the greatest threat to the environment results from the fact that the world’s single model of development is based on the search for power.

But just as the warrior is tempted to return to nature, as we have seen above, the priest is tempted to turn away from it. Nature is dethroned, becoming a source of affliction and corruption of the soul, a “valley of tears”. As for the producer, the temptation is materialistic: the thing produced acquires greater importance than those for whom it is intended. This materialistic temptation is greatest amongst producers whose activity is pre-capitalist, like farmers and craftsmen, who lose sight of the fact that the real goal is serving others rather than producing things. Production is only the means for providing this service.³⁶

In human evolution, the purpose of the capitalist market is to integrate warrior values while reversing their goal – no longer protecting some, but serving all. The producer’s world is divided between those who buy his products and those who have yet to buy them. That is the producer’s only boundary. No “*raison d’État*” exists for him; as a result, there is no reason to invent an enemy. One does not make the supreme sacrifice for the sake of a balance sheet.³⁷ The way of the warrior becomes the way of business when the total devotion demanded of him is no longer related to his role as citizen, when he stops being loyal to a state in order to meet the needs of all human beings. Unifying rather than dividing humanity now becomes his calling. A new kind of warrior with no weapons other than imagination and the will to persuade, he rejects war and political artifice to embrace the values of reciprocity and life. A new kind of priest, he does not turn away from nature, but knows that his work as a producer can serve to humanize him. As a result, his eyes and ears are forever in harmony with high ideals.

The emergence of the state has exacerbated the conflicts amongst societies and di-

³⁶ See my *How Should We Think About Economics Today?*, www.liberalia.com.

³⁷ Freud believed that every society eternally acts out the conflict between Éros and Thanatos, the life and death instincts. The sublimation of this conflict on behalf of some higher cause provided consolation to its victims and subjects for the great plays of 17th century writers. Producers refuse to accept this conflict between happiness and collective values. The Romantics mock their bourgeois and feminine morality, which exalts life and productivity, as “petty” and “selfish”. In a book with the titillating name, *Fellatio, Masochism, Politics & Love*, Leo Abse puts the following words in the mouth of one of the last individuals nostalgic for the warrior epics: “I have a profound sense of envy: Why should these men, simply through a chronological accident, have had the chance to serve their country so courageously whereas our own generation has been consigned to live in an era that history will quickly forget? It is impossible not to feel a sense of nihilism about an age in which the biggest issues seem to be about interest rates and Ron Davies’ nocturnal activities, rather than life and death, war and peace. . . It is hard to find much purpose in the modern era, shorn as it has been of sacrifice and danger.” One thinks also of Ernst Junger and the exacting warrior morality he defended during the two world wars.

vided them internally into antagonistic classes. This division reflects the three functions that each of us can and should exercise ourselves: spirituality; service to others (the warrior's true function); and the transformation of matter. When they are reconciled in the economy, these three functions can no longer cause the division of the social sphere. Humanity, reaching a more advanced stage of evolution, could then achieve the primitives' anarchist ideal.

New Perspectives on Political Economy

Volume 1, Number 1, 2005, pp. 83 – 119

Monetární teorie v českém ekonomickém myšlení 90. let 20. století

J. Koderová*

JEL Classification: E50, E42, B22

Abstract: The article deals with the evolution of Czech monetary theory during the nineties of the last century. At that time the Czech monetary economists were interested in two following problems: the practical realisation of monetary policy and exchange rate regime. The first mentioned problem is analysed in more details in this article. However the main leaders of the Czech transformation process knew the Austrian School's theories very well, especially F. A. Hayek's Law, Legislation and Liberty, the monetary policy of the Czech National Bank has been based on the ideas of monetarism.

1 Úvodní poznámka

Období od počátku ekonomické transformace je obdobím relativně velice krátkým. Pro vývoj monetární teorie a politiky v bývalém Československu a potažmo i v České republice je však obdobím převratným, znamenajícím přechod od pasivního pojetí měnové politiky, v němž byla značně podceněna úloha peněz, k měnové politice v ekonomice tržní, spojené s důrazem na racionalizaci úlohy peněz v hospodářském životě společnosti. Zdaleka nešlo o proces jednoduchý, ale o záležitost velice komplikovanou spojenou s řadou nepříznivě působících faktorů.

* katedra měnové teorie a politiky, Vysoká škola ekonomická v Praze, nám. W. Churchilla 4, Praha 3

V první řadě to byla skutečnost, že nešlo o koncipování a formování měnové politiky ve fungující tržní ekonomice (což samo o sobě představuje složitý problém), ale o hledání podoby pro měnovou politiku, která měla napomáhat samotnému procesu utváření tohoto typu ekonomiky a být jeho aktivní účastníci.

Druhým neméně významným faktorem byla skutečnost, že po celou dobu si československá a posléze česká ekonomika zachovávala charakter malé otevřené ekonomiky, procházející navíc významnou změnou, a to jak pokud jde o komoditní, tak i pokud jde o teritoriální zahraničně obchodní orientaci. Přitom bylo třeba počítat i s nezbytností účasti zahraničního kapitálu v procesu transformace a se všemi s tím spojenými okolnostmi a efekty. Velice důležitou složkou celého procesu bylo narovnání deformovaných cenových relací uvnitř ekonomiky a zrealnění kurzu československé a posléze české měny.

V této situaci chyběla nejen praktická erudice realizátorů nezbytných opatření v oblasti národohospodářské praxe a možnost studia zkušeností z uskutečňování analogických procesů v jiných zemích, ale nezřídka i hlubší a zažitě teoretické znalosti některých makro i mikroekonomických vztahů a souvislostí.

Proto se nezřídka objevovaly práce čistě teoretické zaměřené na výklad právě diskutovaných problémů, které měly ozřejmit podstatu jejich fungování a vazeb k základním makroekonomickým proměnným. Významným počinem bylo vydávání překladů renomovaných zahraničních učebnic a publikací a posléze i vlastních učebnic a monografií českých monetárních ekonomů.

Česká ekonomická věda splácela v tomto období také dluh vůči vývoji českého ekonomického myšlení, v jehož rámci se nezřídka objevovaly fundované exkurzy do období mezi oběma světovými válkami a bezprostředně po druhé světové válce, tj. období, v nichž byly právě otázky spojené s koncipováním a realizací měnové politiky velice aktuální.

Celý proces formování nového typu měnové politiky v 90. letech byl provázen četnými diskusemi, kterých se účastnili přední čeští národohospodáři, členové akademické obce i politici. Základním problémem při jejich mapování je jejich četnost a šíře problémů, kterým byly věnovány, jakož i velice rozsáhlá škála jejich účastníků.

Dalším komplikujícím faktorem skutečnost, že tyto diskuse probíhaly při různých

příležitostech a na různých místech, přičemž některé z nich nebyly zdokumentovány. Publikované příspěvky zase nacházely širokou možnost umístění v nejrůznějších novinách a časopisech, a to nejen odborně zaměřených.

V souvislosti s tím vyvstává *nutnost redukce* celé existující fakticity, a to hned v několika směrech. V první řadě je to typ diskusí, tj. jejich zaměření na řešení určitého problému, který se profiloval v jisté době jako problém stěžejní, svým významem dané období přesahoval, měl vliv na budoucí vývoj měnových podmínek a v návaznosti na to i na vývoj reálné ekonomiky.

Dále je to „prostor“, v němž diskuse probíhaly, přičemž za nejvýznamnější z nich je možno podle mého názoru považovat diskuse probíhající v rámci Československé/České společnosti ekonomické, Státní banky československé, později České národní banky a jejich Institutu ekonomie (1992-1998), diskusních seminářů pořádaných Katedrou hospodářské politiky Fakulty národohospodářské VŠE v Praze a na ně navazujících seminářů pořádaných Institutem pro ekonomickou a ekologickou politiku téže fakulty. Řadě závažných měnově politických témat byly věnovány také semináře Centra pro ekonomiku a politiku, opominout nelze ani některé diskuse, které proběhly v Liberálním institutu.

Z hlediska publikace jednotlivých příspěvků se jako rozhodující jeví časopisy Finance a úvěr (zejména pak některá jeho monotematická čísla) a Politická ekonomie, dále potom České banky/Bankovníctví, Hospodářské noviny, případně Ekonom a Euro. Významným zdrojem jsou bezesporu výzkumné zprávy a materiály publikované SBČS resp. ČNB, Českou společností ekonomickou, Komerční bankou (Hospodářské trendy), Centrem pro ekonomiku a politiku a Nadáním Josefa, Marie a Zdeňky Hlávkových.

Stranou nelze podle mého názoru ponechat ani výzkum v oblasti monetární teorie a její výuku realizovanou na prestižních českých vysokých školách s ekonomickým zaměřením - Vysoká škola ekonomická v Praze, Fakulta sociálních věd Univerzity Karlovy, Fakulta ekonomická Vysoké školy báňské - Technické univerzity v Ostravě, Fakulta ekonomicko-správní Masarykovy university v Brně a Fakulta obchodně-podnikatelská Slezské univerzity v Opavě. Významnou úlohu při formování názorů českých monetárních ekonomů sehrály bezesporu i překlady prestižních zahraničních publikací týkajících se dané oblasti.

Za jistý handicap lze označit práci se „živým materiálem“, tj. s názory a polemikami

našich současníků, které ještě neprošly „sítím“ historického vývoje a prověřením časem.

2 Vymezení hlavních problémových okruhů zkoumání

V podstatě lze konstatovat, že celému sledovanému období dominovaly *dvě základní linie diskusí*, které se v procesu svého vývoje pod vlivem různých faktorů dále rozvíjely a aktualizovaly. První z nich představuje vlastní proces koncipování a realizace měnové politiky, druhá je spojena s volbou systému a s nastavením měnového kurzu československé/české koruny a s procesem nastolování vnější ekonomické rovnováhy. Je ovšem zřejmé, že obě linie spolu navzájem velice úzce souvisejí a existují mezi nimi velice těsné vazby.

Diskuse o koncipování a realizaci měnové politiky, kterým bude věnován tento příspěvek, zahrnují celou škálu problémových okruhů, které se objevovaly v různých dobách s různou intenzitou. V první řadě sem patří diskuse o postavení a úloze centrální banky, dále pak diskuse o aplikaci konkrétních měnově politických nástrojů a o jejich účinnosti, a zejména pak diskuse o transmisním mechanismu měnové politiky jako takovém.

Tyto diskuse probíhaly od samého počátku koncipování měnové politiky a měly za následek přechod od poloadministrativního řízení přes strukturu bankovních aktiv (limity úvěrů a úrokové stropy) k aplikaci monetaristického (měnového) transmisního mechanismu s měnovou bází jako operativním a peněžní zásobou jako zprostředkujícím kritériem. Již v průběhu jeho aplikace došlo k některým modifikacím a posléze k nahrazení cílování peněžní zásoby cílováním inflace. V souvislosti s tím se rozvinula rozsáhlá diskuse o perspektivách vývoje a potenciálních důsledcích této změny, která patří k nejvýznamnějším diskusím v současném vývoji české monetární ekonomie.

Kromě toho byly velice zajímavé i diskuse o endogenitě peněžní zásoby v ČR vyvolané zejména neplněním záměrů ČNB pokud jde o růst peněžní zásoby v době využívání monetaristického transmisního mechanismu a „podstřelování“ intervalu cílované inflace v pozdější době. Dále zde mají místo diskuse o měnově politických příčinách poklesu ekonomiky 1997-1999 a diskuse o postavení centrální banky.

Dalším významným a široce diskutovaným problémem byla problematika nastavení

a tvorby režimu měnového kurzu československé/české koruny. Zpočátku se odehrávala především mezi stoupenci V. Komárka a V. Klause, výrazně zintenzívnila v souvislosti s měnovou krizí české koruny v roce 1997 a s na ni navazující změnou režimu měnového kurzu. K této diskusi se váže ještě další okruh diskutovaných problémů - diskuse o problematice vyrovnanosti platební bilance a vztahu vnější a vnitřní rovnováhy.

Vzhledem k závažnosti této problematiky a k rozsahu publikované literatury, která se k ní váže, by si její výklad vyžádal samostatnou práci. Z tohoto důvodu je předkládaný příspěvek věnován pouze diskusím o problematice postavení centrální banky a o koncipování a realizaci měnové politiky ČNB v 90. letech minulého století.

3 Teoretická východiska

Z hlediska zaměření PCPE se jako velice významná jeví odpověď na otázku, do jaké míry ovlivnila znalost alternativních pohledů na měnové uspořádání pocházejících zejména z učení příslušníků rakouské školy podobu československého a posléze českého peněžního systému a procesu realizace měnové politiky.

Je všeobecně známo, že vídeňská subjektivně psychologická škola v počátečním stadiu svého vývoje zcela zásadním způsobem ovlivňovala české ekonomické myšlení, a to nejen díky tomu, že české země byly v období vytváření základů jejího učení, které spadá do doby od 70. let 19. století do 20. let století dvacátého, součástí Rakouska-Uherska. Všichni tři hlavní protagonisté jejího formativního období byli nějakým způsobem spjati s českým územím. Duchovní otec a dodnes kultovní osobnost celé školy, vídeňský univerzitní profesor Carl Menger studoval práva na univerzitě ve Vídni a na německé části Karlo-Ferdinandovy univerzity v Praze, innsbrucký a posléze vídeňský univerzitní profesor a několikanásobný ministr financí rakousko-uherského mocnářství E. von Böhm-Bawerk se narodil v Brně a pozdější Mengerův nástupce na místě profesora vídeňské univerzity F. von Wieser působil předtím 20 let jako profesor na německé části pražské Karlo-Ferdinandovy univerzity a byl také jejím děkanem.

Byli to právě tito ekonomové, kteří ve své době nejsilněji působili na české ekonomické myšlení, a to zejména svou teorií hodnoty založenou na teorii mezního užítku, hledáním vztahu mezi teorií hodnoty a teorií ceny a úroku a svým metodologickým

principem hospodářského individualismu a subjektivismu. Silný vliv měla i Mengerova a Wieserova teorie subjektivní hodnoty peněz se svým důrazem na vliv peněžních důchodů na ceny, silně kontrastujícím s kvantitativní teorií peněz.¹

Tato tradice byla sice v totalitním období silně potlačena, ale nikdy nebyla zcela zlikvidována. Základní informace o učení příslušníků rakouské školy (převážně jejích starších generací) dostávali studenti VŠE v Praze v rámci kurzů dějin ekonomického myšlení, jeho dalším vývojem se zabývali vedle učitelů těchto kurzů také pracovníci ve vědecko-výzkumných institucích.

Šlo zejména o mladé iniciativní pracovníky, které vedle alternativního výkladu základních ekonomických kategorií (teorie cen, teorie peněz, teorie výroby a kapitálu, problematika vývoje trhu práce a teorie ekonomického růstu a hospodářských cyklů) s největší pravděpodobností okouzila také metodologie rakouské školy a důsledný anti-tatistický liberalistický přístup² jejích pozdějších představitelů L. v. Misesa a F. A. v. Hayeka a jejich následovníků.

Je tudíž logické, že se mezi hlavními architekty transformace československé a posléze české ekonomiky, kteří se rekrutovali převážně z řad teoretických ekonomů,³ objevily osobnosti, které se k rakouské škole explicitně hlásily, mezi nimi v první řadě V. Klaus a T. Ježek, ale i někteří další. Řada indikátorů vede k závěru, že klíčoví protagonisté transformace byli obeznámeni s fundamentální prací F. A. v. Hayeka *Právo zákonodárství a svoboda*,⁴ která vyšla také v českém překladu v roce 1991, jakož i s řadou dalších stěžejních prací příslušníků rakouské školy.

¹ Blíže viz Vencovský, F.: *Dějiny českého ekonomického myšlení do roku 1948*. Nadace Universitas Masarykiana. Edice Heureka. Brno, 1997.

² Liberalismus je zde zákonitě vyvozován ze dvou základních teoretických závěrů ekonomického učení rakouské školy, a to že jakékoli vnější omezování svobodné konkurence vede jednoznačně ke snižování efektivnosti tržního hospodářství a že každé podněcování hospodářské aktivity, které nevychází ze svobodného rozhodnutí hospodářských subjektů omezit přítomnou spotřebu na základě substituce za větší spotřebu v budoucnosti, deformuje strukturu společenské výroby a vede k chybným investicím se všemi dalšími negativními důsledky v podobě konečného poklesu agregátní produkce. (Blíže viz Kindlová, E.: *Od C. Mengera k F. A. Hayekovi*. Habilitační práce. VŠE Praha, 2002.)

³ Teoretičtí ekonomové měli paradoxně nejbližší k formování představ o možnostech přechodu k tržní ekonomice a k posílení úlohy peněz, neboť domácí praktické zkušenosti z tržního hospodářství de facto vůbec nebyly k dispozici a zkušenosti ze standardních tržních ekonomik v tomto případě nebyly zcela relevantní.

⁴ Havel, J.: *Ekonomika a právo v transformaci*. Politická ekonomie 1998, č. 6, s. 869

Právě Hayekův spontánně vznikající systém lidských interakcí se stal spolu s koncepcí neviditelné ruky trhu podstatou filosofie transformace československé a posléze české ekonomiky. V souladu s touto filosofií bylo cílem vládní politiky dovedení transformace do fáze víceméně samovolného vývoje, čímž měl být také celý proces transformace ukončen.

Václav Klaus tuto skutečnost charakterizoval slovy: „V souladu se svou zásadní ekonomickou filozofií věřím v neviditelnou ruku Adama Smithe a v Hayekův spontánně vznikající systém lidských interakcí. Současně jsem si vědom toho, že je třeba - zejména při prvotní konstrukci takto pojímaného systému - dodržovat určité netriviální zákonitosti a jistá sekvenční pravidla a že je proto třeba ponechat v této fázi nemalou „konstrukční“ úlohu i ekonomickému centru, tedy instituci, jejíž úloha v období normálního chodu takto koncipovaného systému musí být naopak minimální.“⁵

V oblasti monetární politiky se však z řady důvodů prosazovala spíše monetaristická koncepce a její pojetí liberalismu.

4 Diskuse o postavení centrální banky

Ekonomická věda nabízí celou řadu pohledů na vznik a vývoj peněz a na jejich postavení a úlohu v hospodářském životě společnosti, které se mohou navzájem i dosti podstatně odlišovat. Díky tomu se setkáváme se širokou škálou teorií a teoretických koncepcí, které leží mezi dvěma krajními póly a tu úžeji tu volněji se přiklánějí k některému z nich. Na jedné straně jsou to teorie a teoretické koncepce, které úlohu peněz v ekonomice bagatelizují a redukují ji na plnění základních peněžních funkcí, tj. vyjadřování cen ostatních zboží a zprostředkování směny, přičemž vliv na vývoj reálných hospodářských veličin buď vůbec neuvažují či jej připouštějí pouze v případě, kdy dojde k šokům v monetární oblasti.

Na druhé straně jsou to teorie, v jejichž pojetí jsou peníze rozhodujícím fenoménem vývoje rozvinutých tržních ekonomik, takže zajišťování rovnovážného vývoje v monetární oblasti v interakci s ostatními oblastmi hospodářského života společnosti představuje podle nich jeden ze základních předpokladů bezporuchového fungování ekonomiky.

⁵ Klaus, V.: *Ekonomická teorie a realita transformačních procesů*. Management Press, Praha 1995

Hlavním předmětem zkoumání těchto teorií je v souvislosti s tím hledání cest ke zjišťování potřebného množství peněz v ekonomice (neboli predikování poptávky po penězích) na jedné straně a zkoumání nástrojů, s jejichž pomocí by bylo možno odpovídající množství peněz emitovat na straně druhé, jinými slovy řečeno hledání cest k zajišťování rovnovážného vývoje na peněžním trhu.

Česká monetární teorie zaznamenala ve svém vývoji oba extrémní přístupy. Snaha o posílení úlohy peněz a racionalizaci jejich fungování v procesu transformace ekonomiky vedla české ekonomy k upřednostňování teorií zdůrazňujících význam peněz. Řada okolností vedla k tomu, že se základním východiskem našich monetárních ekonomů koncipujících zásady pro měnovou politiku stalo učení monetarismu.

V první řadě to byly základní charakteristické rysy monetaristické doktríny: hlásání zásady hospodářského liberalismu⁶ a představa, že centrální banka je schopna pouze na základě dodržování pravidla dopředu oznamovaného stálého růstu peněžní zásoby odpovídajícího předpokládanému růstu hrubého domácího produktu zajistit bezporuchové fungování ekonomiky a nízkou míru inflace.

Právě obavy z hrozících inflačních tlaků, tradice vytváření silné měny z období první republiky, všeobecný respekt k pevné západoněmecké marce a potažmo i k měnové politice Bundesbanky, a do jisté míry i možnost, byť značně omezená, využití zkušeností z praxe existujících centrálních bank představují další neméně důležité faktory hovořící ve prospěch přijetí základních myšlenek monetarismu, resp. systému s relativně silnou nezávislou centrální bankou.

K tomu, že velice sporadická upozornění na minimálně teoretickou možnost existence systémů bez centrální banky zůstala zcela okrajovou záležitostí,⁷ přispěla beze-

⁶ Zde však již nejde o antietatistický liberalistický přístup příslušníků rakouské školy charakterizovaný v předcházející části příspěvku, ale o systém do jisté míry připouštějící stávající status quo v podobě státu a centrální banky, zároveň však usilující o minimalizaci státního ekonomického intervencionismu a o podřízení působnosti jinak zcela nezávislé - a de facto současně také nikomu nezodpovědné - centrální banky relativně jednoduchému pravidlu adekvátnosti tempa růstu zvoleného peněžního agregátu předpokládanému tempu růstu nominálního agregátního výstupu.

⁷ viz např. Bulíř, A.: Volné bankovníctví: znovunalezené řešení? *Finance a úvěr*, 1992, č. 8; Zahradník, P.: Volné bankovníctví a slučitelnost s realitou. *Finance a úvěr*, 1992, č. 8. Pozdější stat' Čihák, M. a Holub, T.: Měnový výbor: cesta z dilemat centrálního bankovníctví v ČR? *Bankovníctví 2000*, č. 6, již měla zcela jiný charakter a spíše než otázku uspořádání bankovní soustavy se snažila najít odpověď na otázku související s kurzem koruny

sporu také skutečnost, že nezávisle na listopadových událostech roku 1989 došlo k některým změnám v peněžním sektoru ekonomiky již na počátku roku 1990, a to na základě Zákona č. 130/189 Sb. o Státní bance československé, který vstoupil v platnost 1. ledna 1990. Zákon upravoval postavení SBČS v souvislosti s plánovanou přestavbou hospodářského mechanismu, jejímž cílem bylo posílení úlohy peněz v ekonomice.

Tento zákon se měl stát spolu se zákonem o bankách a spořitelnách z téhož roku základem bankovní reformy, jejímž hlavním cílem bylo vytvoření dvoustupňového bankovního systému. Ač byly oba zákony zpracovány a schváleny v době administrativně direktivního systému řízení, došlo k jejich naplňování ve zcela nových a nepředpokládaných politických, společenských a hospodářských podmínkách. Přesto vytvářely alespoň v počátečním období základní institucionální rámec pro vznik dvoustupňového bankovního systému a pro koncipování a realizaci nového typu měnové politiky.

Pokud jde o zákon o SBČS, některá jeho ustanovení lze na tehdejší dobu bezesporu označit jako progresivní, nicméně je třeba říci, že zákon by zcela určitě vypadal jinak, kdyby byl připravován a schválen po listopadových událostech roku 1989.

Dne 20. prosince 1991 byl přijat Federálním shromážděním České a Slovenské federativní republiky zákon o Státní bance československé č. 22/1992 Sb., který jí přisoudil postavení a funkce standardních emisních bank ve fungujících tržních ekonomikách, a zákon o bankách č. 21/1992 Sb., který stanovil základní pravidla pro fungování a pro činnost obchodních bank.

Přípravy zákona se zúčastnila řada expertů z evropských centrálních bank, Mezinárodního měnového fondu a Světové banky. Byl také využit zákon o německé Bundesbank a zohledněn návrh zákona o Evropském systému centrálních bank a Evropské centrální bance. Cílem bylo vytvořit právní podklad pro působení SBČS jako moderní centrální banky.

SBČS v něm byla definována jako ústřední banka ČSFR, jako právnická osoba působící jako federální ústřední orgán, který může vydávat právní předpisy a opatření. Jako hlavní cíl byla stanovena stabilita československé měny a bance byly přisouzeny prakticky všechny funkce typické pro centrální banky ve vyspělých tržních ekonomikách.⁸

⁸ Mataj, J. - Vojtíšek, P.: Právní a ekonomické aspekty nového Zákona o SBČS. FÚ 1992, č. 3. Vencovský, F.: Vzestupy a propady československé koruny. Oeconomica, Praha 2003

Pro splnění stanoveného cíle byl SBČS poskytnut vysoký stupeň nezávislosti na federální i republikových vládách⁹, který zásadním způsobem změnil její postavení oproti předchozímu zákonu. Došlo i ke striktnímu vymezení vztahu ke státnímu rozpočtu.¹⁰ Nezávislému postavení SBČS odpovídala i právní úprava jejího hospodaření, především pak její nezainteresovanost na zisku.

Z ekonomického hlediska lze říci, že SBČS dostala v zákoně pro plnění svého cíle takové možnosti, jaké byly obvyklé pro centrální banky ve vyspělých tržních ekonomikách. V některých aspektech byl zákon dokonce progresivnější než tehdejší zákonné úpravy zemí Evropského společenství, neboť již vycházel z ustanovení návrhu zákona o Evropské centrální bance. Na druhé straně ale také respektoval tehdejší nižší úroveň československého bankovníctví a ač perspektivně počítal především s uplatňováním tržních nástrojů, byla v něm připuštěna i možnost využití nástrojů přímých v případě nepříznivého vývoje.

Již v průběhu roku 1992 však bylo připravováno rozdělení Československa na dva samostatné státy¹¹ a v souvislosti s tím i republikové dělení SBČS. Federální shromáždění přijalo 8. října 1992 novelu státního zákona o České a Slovenské federativní republice, která oběma republikám umožnila založit si vlastní centrální banku.

K 1. lednu 1993 byla založena Česká národní banka jako nástupce SBČS na území České republiky. Její postavení, cíle, rámec její činnosti a funkce vymezoval zákon o ČNB č. 6/1993 Sb. přijatý 17. prosince 1992. Tento zákon vycházel ze Zákona o Státní bance Československé č. 22/1992 Sb. Jako hlavní cíl ČNB stanovil zabezpečování stability české měny.¹²

Z hlediska obvykle používaných kritérií se ČNB zařadila mezi centrální banky s vy-

⁹ Vztahy mezi centrální bankou a vládami probíhaly pouze na základě konzultační a poradní činnosti a prostřednictvím zaujímání stanovisek k materiálům, které se týkaly působnosti SBČS. S tím souvisela možnost účasti členů vlády na zasedáních bankovní rady a guvernéra a viceguvernéřů pověřených řízením republikových ústředí na schůzích vlád. Dvakrát ročně měla banka předkládat Federálnímu shromáždění zprávu o měnovém vývoji a informovat o něm Českou národní radu a Slovenskou národní radu. Federálnímu shromáždění měla také předkládat ke schválení roční zprávu o výsledku svého hospodaření.

¹⁰ Podle §35 mohla SBČS poskytnout České a Slovenské Federativní Republice, České republice a Slovenské republice krátkodobý úvěr nákupem pokladničních poukázek splatných do tří měsíců od jejich nákupu, přičemž celkový stav těchto úvěrů nesměla překročit 5 % příjmů příslušného rozpočtu v uplynulém roce.

¹¹ Zákon č. 542/1992 Sb. Federálního shromáždění o zániku České a Slovenské federativní republiky

¹² V zákoně nebylo explicitě uvedeno, zda se jedná zároveň o stabilitu vnitřní i vnější.

sokou mírou nezávislosti a mohla tak nezávisle formulovat a provádět svou měnovou politiku. Zatímco počáteční legislativní úpravy postavení centrální banky probíhaly spíše formou přejímání zahraničních zkušeností a doporučení, nesrovnatelně širší diskuse o postavení a úloze centrální banky se rozpoutala jako reakce na hospodářský pokles v letech 1997 až 1999, který byl nezřídka označován jako důsledek příliš razantního restriktivního postupu ČNB.

Tato diskuse, která měla spíše charakter politické debaty,¹³ se týkala nejen měnové politiky České národní banky, ale i jejího postavení, a to nejen z hlediska ekonomického, ale i ústavního. Svou úlohu v rozpoutání této diskuse, sehrálo bezesporu také nabytí účinnosti „malé“ (6. 2. 1998) a „velké“ (1. 9. 1998) novely zákona o bankách.

S jistou dávkou zjednodušení lze říci, že v rámci této diskuse vykrytalizovaly zhruba tři skupiny názorů. První, i když nikoli nejvýznamnější, interpretace hovořila o transformační recesi, jejíž hlavní příčinu spatřovala ve špatně provedené privatizaci, která nevytvořila dostatečně silnou vrstvu vlastníků a neumožnila podnikům získat dostatečné kapitálové zdroje, a tím výrazně snížila vyhlídky na silnější hospodářský růst.

Druhá skupina spojovala vznik recese s poměrně rychlým hospodářským růstem v letech 1993 - 1996, který měl za následek přehřátí ekonomiky a vznik vnější nerovnováhy, jež musela být řešena zpomalením růstu. Zde však došlo k další názorové diferenciaci. Odpůrci řešení, které v této souvislosti zvolila ČNB, kritizovali, že setrvala u režimu fixního měnového kurzu po té, co v říjnu 1995 provedla rozsáhlou liberalizaci toků na finančním účtu platební bilance a umožnila obchodování s českou korunou na mezinárodních měnových trzích při zachování režimu fixního měnového kurzu.

Dále pak, že po rozšíření flukтуаčního pásma bez konzultace s vládou přistoupila v červenci 1996 ke zvýšení povinných minimálních rezerv a úrokových sazeb a že na rozpočtové balíčky v roce 1997 nereagovala razantním uvolněním měnové politiky. Takto vytvořený brzdící impuls zesílený dále stanovením ambiciózního inflačního cíle pro rok

¹³ Také sociální demokracie označovala měnovou politiku za jednu z nejdůležitějších příčin recese a požadovala větší vstřícnost bankovní rady vůči potřebám hospodářského vývoje a intenzivnější spolupráci s vládou. (srv. např. Zeman chce mít ve vedení ČNB své lidi. Lidové noviny 26. 9. 1998, s. 1; ČNB je příliš nezávislá a nemá zpětnou vazbu. Lidové noviny 1. 10. 1998, s. 16). V této otázce tak došlo dokonce k určitému konsenzu mezi SDS a ODS, který vyústil do rozsáhlé výměny názorů o postavení, nezávislosti, zodpovědnosti a cílech centrální banky. (srv. např. Na mušce politiků. Euro, 2. 11. 1998; Ve stínu politiky. Euro 30. 11. 1998, s. 12)

1998 a s jeho značným „podstřelením“ pak podle nich výrazně prohloubil recesi.¹⁴

Podle stoupenců měnové politiky ČNB naopak monetární restrikce nemusela být tak silná a bylo možno předejít řadě problémů, pokud by vláda včas korigovala fiskální a mzdový vývoj a pokud by více spolupracovala s centrální bankou.¹⁵

Nezávisle na této debatě, i když její vliv nelze zcela vyloučit, se v české ekonomické obci zrodila již v průběhu roku 1998 myšlenka na uspořádání diskuse o postavení centrální banky. Pro její realizaci byly uváděny především dva důvody. Jednak to byly výrazné změny v postavení centrálních bank ve vyspělých tržních ekonomikách v průběhu 90. let v souvislosti se selháním politiků pokud jde o udržování stability měny, k němuž došlo v 70. a na počátku 80. let, jednak to byl samotný vývoj ekonomické teorie.¹⁶

Myšlenka na uspořádání diskuse nabyla konkrétnější podoby na jaře roku 1999 a zrealizovala se jednak na stránkách časopisu *Finance* a úvěr, kde jí byla věnována dvě monotematická čísla (9 a 10), jednak na semináři České společnosti ekonomické (v říjnu téhož roku), který měl zcela netradičně podobu panelové diskuse.¹⁷

¹⁴ srv. např. Holman, R.: *Transformace české ekonomiky*. CEP, Praha 2000. Klaus, V.: *Deficit, restrikce, dezinflace a stav ekonomiky*. Vystoupení na VŠE v Praze dne 26. 11. 1998).

¹⁵ Tošovský, J.: přednáška na Manažerském fóru v Praze. Červen 1998)

¹⁶ Od počátku 80. let začaly v ekonomické teorii získávat popularitu modely časové nekonzistence, které doporučovaly, aby vláda svěřila měnovou politiku nezávislé centrální bance, která bude schopna zajistit nižší inflaci než centrální banka podřízená vládě, a to při dlouhodobě nezměněném výstupu. Novější literatura (v rámci tzv. nové politické ekonomie) pak jako jeden z pragmatických důvodů, proč vlády realizují výše uvedené opatření, uvádí přenos zodpovědnosti na centrální banku. (podrobněji viz Čihák, M. - Holub, T.: *Co říká ekonomická teorie o nezávislosti centrální banky?* FÚ 1999, č. 9)

¹⁷ Panelové diskuse se zúčastnilo 9 ekonomů. Zastoupeni zde byli jak členové akademické obce a analytici, tak i ekonomové se zkušenostmi z centrálně-bankovní praxe (M. Čihák, J. Havel, J. Jonáš, V. Kotlán, P. Kysilka, M. Mandel, Z. Tůma, P. Zahradník, M. Žák). Diskusi moderoval O. Schneider. Šlo jednak o autory či spoluautory statí publikovaných v uvedeném monotematickém dvojčíslu časopisu *Finance* a úvěr, jednak o další významné domácí odborníky. Diskuse na semináři se týkala řady předem vytipovaných témat a postupovala od obecných otázek (co rozumět pod nezávislostí centrální banky a zda, a pokud ano, tak do jaké míry má být centrální banka nezávislá; zda má zákon o centrální bance obsahovat cíl, a pokud ano, co by jím mělo být a jak by měl být vymezen; jak a komu má být centrální banka za plnění svého cíle zodpovědná) k otázkám velice konkrétním (kdo by měl jmenovat a odvolávat členy bankovní rady apod.) Většina účastníků diskuse dospěla ke shodě v řadě otázek, např. pokud jde o doporučení oddělit měnový orgán od orgánu pro bankovní regulaci a dohled, prosadit větší transparentnost pokud jde o fungování centrální banky, či učinit stanovování cílů předmětem společného jednání vlády a centrální banky. (Blíže viz *Postavení centrální banky*, čtrnáctý seminář České společnosti ekonomické v řadě „*Ekonomické teorie a česká ekonomika*“. Praha, 12. října 1999)

Polemiky probíhající mimo tuto platformu¹⁸ měly většinou silně kontroverzní charakter a ve většině případů i politický náboj. Probíhaly na stránkách většiny významných domácích deníků a týdeníků, zejména pak Lidových novin, Mladé fronty DNES, Práva, Ekonomu a Euro. Není tedy překvapivé, že vyústily do návrhu novely zákona o ČNB, který byl předložen Poslanecké sněmovně v polovině roku 2000. Podle O. Dědka se mělo jednat pouze o jednoduchou harmonizační novelu, která měla naši právní úpravu týkající se centrální banky učinit plně konformní s právní úpravou přijatou v rámci EU.¹⁹

Podle V. Klause šlo naopak o logický důsledek praktické zkušenosti, která ukázala, že míra nezávislosti přisouzená ČNB na počátku transformace se s odstupem času jeví jako přemrštěná a je třeba ji korigovat.²⁰

Přes odpor Senátu a prezidenta bylo nakonec Poslaneckou sněmovnou přijato kontroverzní znění novely, které nabylo účinnosti k 1. 1. 2001. Na základě ústavní stížnosti prezidenta republiky však byla v červnu 2001 Ústavním soudem zrušena její nejspornější ustanovení. Ve svém druhém nálezu pak Ústavní soud zaujal stanovisko, že jmenování guvernéra a viceguvernéřů ČNB spadá výlučně do pravomoci prezidenta a nevyžaduje tudíž konsignaci premiéra či jím pověřeného člena vlády. Na podzim téhož roku byla přijata novela Ústavy měnící ústavně zakotvený cíl ČNB z měnové stability na stabilitu cenovou a byla přijata druhá novela zákona o ČNB, která řešila volná místa v zákoně vzniklá v souvislosti se zrušením některých ustanovení předcházející novely Ústavním soudem.

¹⁸ To samozřejmě neplatí všeobecně. Věcný charakter měla např. také diskuse, která proběhla v Liberálním institutu za účasti Z. Tůmy, Z. Revendy a byla moderována J. Schwarzem.

¹⁹ Sedláček, T.: Nezávislost centrální banky potřebí. *Bankovníctví* 2002, č. 6-7

²⁰ Právní postavení centrální banky v demokratickém státě. Sborník textů č. 3, CEP, Praha 2000. V tomto sborníku jsou obsaženy další příspěvky do diskuse týkající se nezávislosti, kredibility a zodpovědnosti ČNB, novely zákona o ČNB, komparace stávající úpravy naší centrální banky s postavením ECB a německé Bundesbanky, možnosti spolupráce ČNB s vládou a mixu fiskální a monetární politiky a dalších problémů. I když je zde postavení centrální banky nazíráno z různých úhlů pohledu, dospívají všichni autoři v podstatě ke shodnému závěru o nezbytnosti omezení v té době platných pravomocí ČNB.

5 Koncipování a realizace měnové politiky

Již na počátku transformace se ozývaly hlasy volající po realizaci měnové politiky monetaristického typu²¹ vykonávané na principu jednoduchých pravidel a založené na cílování peněžní zásoby. Stoupenci těchto názorů vycházeli z přesvědčení, že centrální banka již disponuje potřebnými měnově politickými nástroji pro realizaci takového typu měnové politiky, připouštěli však možnost výskytu určitých problémů spojených s jejich aplikací v podmínkách transformující se ekonomiky.²²

Bylo tomu tak především proto, že k dispozici nebyla žádná teoretická koncepce, která by zohledňovala atypičnost chování ekonomiky v procesu realizace zásadních systémových změn. Jedinou možnou volbou tak byla standardní ekonomická teorie spojená s úsilím o postupné uvádění transformovaného systému do stavu tržní ekonomiky. Jako hlavní důvody pro upřednostňování monetaristické doktríny lze uvést její liberální orientaci, důraz na úlohu peněz v ekonomice a v neposlední řadě i nabídku zdánlivě velice jednoduchého pravidla pro řízení měnové politiky.

Monetaristický transmisní mechanismus byl do měnové politiky SBČS zaveden po několika měsících přípravy již ve druhém čtvrtletí roku 1992²³, a to v modifikované podobě, v níž úlohu operativního kritéria hrály nevypůjčené rezervy²⁴ a hlavním střednědobým kritériem byla peněžní zásoba (M2).

Ze stejného základu vycházela i koncepce měnové politiky nově založené ČNB, a to i přes to, že v české ekonomice nebyly splněny podmínky pro jeho úspěšnou aplikaci, a to ani v jednom z jeho stěžejních bodů. Velice problematickou záležitostí bylo samo

²¹ Jonáš, J. – Bulíř, A. – Kerouš, M.: Makroekonomická politika v období privatizace – náčrt základních problémů (I). FÚ, 1990, č. 12

²² Zejména byla zdůrazňována existence vysokého inflačního potenciálu a možnost jeho dalšího zvýšení v souvislosti s liberalizací cen a zahraničního obchodu a s odstraňováním strukturálních deformací, což vedlo k obavám, že bude obtížné překonávat nerovnovážné stavy bez výraznějších negativních dopadů na zaměstnanost a tempo růstu produkce. Kromě toho nebyly ekonomické subjekty popsatele stabilními poptávkovými funkcemi, takže bylo velice obtížné předvídat konečné důsledky prováděných měnově politických opatření. (Jonáš, J. - Bulíř, A. - Kerouš, M.: tamtéž)

²³ Ve druhém pololetí téhož roku byly zrušeny úvěrové limity

²⁴ Toto kritérium bylo implicitně formulováno tak, aby volné rezervy obchodních bank byly blízké nule. Protože hlavním nástrojem měnové politiky byly aukce refinančních půjček, umožňovalo to centrální bance kontrolovat výši nevypůjčených rezerv, a tím i měnovou bázi. (Bulíř, A.: Československá monetární politika po roce 1989. FÚ 1993, č. 5)

stanovení potenciálního tempa růstu hrubého domácího produktu, vztahu mezi jeho vývojem a vývojem objemu transakcí, a zejména pak „nezbytné“ míry inflace v procesu ekonomické transformace.

Byly to však i běžné problémy s nastolením podmínek pro úspěšnou aplikaci monetaristického transmisního mechanismu, především problém volby vhodného měnového agregátu,²⁵ operativního kritéria,²⁶ stability peněžního multiplikátoru²⁷ a predikovatelnosti rychlosti obratu peněz,²⁸ které se vyskytovaly i ve vyspělých tržních ekonomikách. V našem případě byly však ještě dále zesíleny nestabilitou podmínek způsobených transformačním procesem.

Silně komplikujícím faktorem byly dále faktory endogenní povahy²⁹, které na jedné straně sice ovlivňovaly množství peněz v ekonomice, na druhé straně se však buď zcela či do značné míry vymykaly kontrole centrální banky. V první řadě to byl silný pří-

²⁵ Byl zvolen měnový agregát M2 obsahující hotovostní oběživo v rukou domácích nebankovních subjektů, netermínované vklady domácích nebankovních subjektů v domácích bankách v domácí měně, termínované a další vklady domácích nebankovních subjektů v domácích bankách v domácí měně a další vklady domácích nebankovních subjektů v domácích bankách v zahraničních měnách.

²⁶ Na místě operativního kritéria byla původně široce definovaná měnová báze, která vedle oběživa a bankovních rezerv zahrnovala také poukázky ČNB. Ta byla v polovině roku 1994 vystřídána operativním řízením pomocí rezerv bank, jež byly počátkem roku 1996 nahrazeny úrokovou sazbou mezibankovního trhu s depozity PRIBOR (1 týden).

²⁷ Ta byla způsobena především dvojnásobnou změnou sazeb povinných minimálních rezerv a nestabilním vývojem poměru mezi oběživem, běžnými vklady a termínovanými vklady. (Mandel, M.: Analýza měnové politiky a měnového vývoje v České republice v letech 1989 - 1995. Bankovníctví, 1997, č. 5) Pro období 1993 - 1996 však bylo stability peněžního multiplikátoru v podmínkách ČR s jistými výchyly téměř dosaženo. (Kodera, J. - Mandel, M.: Transmisní mechanismy měnové politiky v podmínkách české ekonomiky. ČNB, IE, Praha 1997)

²⁸ Ex post počítaná důchodová rychlost peněz pro měnový agregát M2 měla od 2. poloviny roku 1993 klesající tendenci; její pokles se však postupně zpomaloval, přičemž vykazoval i určitý oscilační pohyb. V obdobích, kdy docházelo k růstu rychlosti obratu peněz, bylo možno měnovou politiku ČNB charakterizovat jako restriktivní. V 1. polovině roku 1993 došlo k výraznému přitvrzení měnové politiky v souvislosti s měnovou odlukou. Ve 2. polovině roku 1996 ČNB provedla restriktivní opatření v reakci na nárůst deficitu běžného účtu a po většinu roku 1998 byla prováděna měnová restrikce v souvislosti se zavedením cílování inflace. (Mandel, M.: Vliv změn peněžní zásoby na důchodovou rychlost peněz. in Revenda, Z., Mandel, M., Kodera, J., Musílek, P., Dvořák, P., Brada, J.: Peněžní ekonomie a bankovníctví. 3. vydání. Management Press, Praha 2000.

²⁹ Blíže viz Janáčková, S.: Dilemata současné měnové politiky. Finance a úvěr, 1996, č. 1

liv zahraničního kapitálu,³⁰ který souvisel především s otevřeností české ekonomiky, její relativně dobrou reputací a nabídkou investičních příležitostí, vysokým úrokovým diferencíálem, prakticky nulovým kurzovým rizikem a zavedením směnitelnosti české koruny podle definice MMF pro účely transakcí na běžném účtu na podzim roku 1995,³¹ a problémy s jeho sterilizací.

Dále sem patřil růst mezipodnikové zadluženosti,³² emise euroobligací denominovaných v korunách³³ a problémy spojené s vývojem bankovního sektoru a s náklady na udržení jeho stability.

Prakticky již od roku 1993 tak docházelo k soustavnému překračování v měnovém programu ČNB stanovených i následně korigovaných temp růstu peněžní zásoby.³⁴ Centrální banka tehdy de facto rezignovala na svůj explicitní cíl v oblasti peněžní zásoby a v podstatě implicitně cílovala pevný měnový kurz. Tato situace byla charakterizována jako

³⁰ Příliv zahraničního kapitálu do ČR vyvrcholil po zavedení vnější směnitelnosti české koruny v roce 1995, kdy dosáhl 218,3 mld. korun, tj. 17,4 % podílu na HDP. Od té doby až do zavedení cílování inflace postupně klesal, a to v souvislosti s rozšířením fluktuálního pásma koruny v roce 1996 na 110,5 mld korun, tj. 7,8 % HDP a na svou dosud historicky nejnižší úroveň, tj. 34,3 mld Kč a 2,1 % HDP, v roce 1997, spojenou s květnovými a listopadovými turbulencemi koruny na mezibankovním devizovém trhu, které odrážely vývoj makroekonomických ukazatelů, politickou nejistotu a nestabilní situaci na tuzemském peněžním i kapitálovém trhu. (Údaje čerpány ze zpráv o vývoji platební bilance vydaných ČNB v roce 1997 a 1998).

³¹ Došlo k tomu na základě nového devizového zákona č. 219/1995 Sb. s platností od 1. října 1995. Faktická úroveň směnitelnosti však silně překračovala rámec stanovený článkem VIII. Dohody o MMF. Zákon explicitně vymezoval pouze případy, kdy koruna směnitelná není (šlo především o portfoliové investice pro rezidenty, otevírání účtů rezidentů v zahraničí, poskytování finančních úvěrů nerezidentům a nákup nemovitostí nerezidenty. Na konci roku 1995 pak byl v podstatě liberalizován příliv kapitálu do české ekonomiky a značně pokročilo i odstraňování posledních překážek pro jeho vývoz.

³² Od roku 1993 se sice údaje o mezipodnikové zadluženosti přestaly samostatně vyčleňovat, přesto se uvažovalo o 130 mld Kč v roce 1994. (Blíže viz Janáčková, S.: Měnová politika - úspěchy i hledání nových cest. ČNB, IE, Praha 1995). V letech 1995 - 1999 se pohybovala od 229,5 do 204,1 Kč, přičemž byla ještě posílena o platební nekázeň podniků při úhradách daní a sociálních dávek, která v daném období dosáhla 6,5 % HDP.

³³ První euroobligace denominované v korunách byly emitovány v roce 1995, do února 1997 to bylo 59 emisí v celkovém objemu 78,6 mld Kč.

³⁴ V roce 1994 došlo k úpravě původně programovaného 12 - 15 % růstu peněžní zásoby na 17 %, ale ani tato modifikovaná hodnota nebyla dodržena, neboť skutečné tempo růstu M2 dosáhlo 21 %. Analogická situace nastala i v roce 1995, kdy byla horní hranice původně stanoveného rozmezí 14 - 17 % překročena o 2,4 procentního bodu. V roce 1996 se přírůstky M2 pohybovaly na úrovni 18 - 19 % oproti 13 - 17 % stanoveným měnovým programem. Teprve v roce 1997 byl v průběhu celého roku vývoj peněžní zásoby v souladu se záměry M 2 pro tento rok. Dynamika M2 nevybočila z rámce stanoveného koridoru 7 - 11 % a naopak se pohybovala převážně v jeho dolní polovině.

nesoulad mezi zvyšující se mírou otevřenosti ekonomiky, vývojem reálného efektivního měnového kurzu včetně zvoleného kurzového režimu a měnovou a fiskální politikou.³⁵

Někteří ekonomové tak začali obracet svou pozornost ke koncepcím endogenity peněžní zásoby odmítajícím suverenitu centrální banky v oblasti emise peněz.³⁶ Přesto však nakonec převládla argumentace vycházející z učení hlavního proudu, že v systému fixního měnového kurzu nemá centrální banka prostor pro autonomní³⁷ měnovou politiku. Sílily obavy, že se sterilizace jako nástroj měnové kontroly stane poměrně rychle neúčinnou a že schopnost centrální banky řídit peněžní zásobu při současném zachování režimu fixního měnového kurzu bude omezena v mnohem větším rozsahu, než dosud.³⁸

V souvislosti s přílivem zahraničního kapitálu byla centrální banka v zájmu udržení stabilního měnového kurzu nucena uspokojovat poptávku po korunách, a to i za cenu relativně vysokých temp růstu peněžní zásoby. V rámci boje proti růstu inflace prováděla především sterilizace přílivu kapitálových toků, tj. stahování peněz z oběhu prostřednictvím prodeje domácích cenných papírů. Tato politika měla za následek další příliv zahraničního kapitálu spojený s novými tlaky na změnu úrovně měnového kurzu.

Vznikal zde tzv. začarovaný kruh sterilizační politiky centrální banky, který vyús-

³⁵ Kysilka, P.: Vnější rovnováha jako odraz souladu mezi mírou otevřenosti ekonomiky a parametry makroekonomické politiky. in *Otevřenost české ekonomiky a vnější rovnováha: konec 90. let*. Sborník textů. CEP, Praha 2001

³⁶ Prvním empirickým pokusem o zodpovězení otázky, do jaké míry je vazba mezi penězi, HDP, cenami amzdami v aplikaci na českou ekonomiku v souladu či v rozporu s teorií endogenních peněz rozvíjenou v rámci postkeynesovské ekonomie byla práce Izák, V.: *Endogenous or exogenous money supply? - the Czech case*. ČNB, IE, 1995. VP č. 37. S kritikou Izákových závěrů vystoupil Bulíř, A.: *Exogenita nabídky peněz: některé pochybnosti o předchozím výzkumu*. *Finance a úvěr*, 1996, č. 1. Bulířovy výtky charakterizovali jako zásadní a do značné míry oprávněné Čihák, M. - Janáček, K.: *Inflace v České republice v polovině devadesátých let*. *Politická ekonomie* 1996, č. 1 a připojili další připomínky, které vztáhli i na další Izákovu práci (Izák, V.: *Inflation, Wages and Money*. *Prague Economic Papers*, 1996, č. 3.) Jejich přístup byl podroben kritice in Arlt, J.: *Inflace v České republice v polovině devadesátých let (srovnání alternativních vysvětlení)* *Politická ekonomie*, 1997, č. 5, na niž reagovali in Čihák, M. - Janáček, K.: *Inflace v České republice v polovině devadesátých let: odpověď*. *Politická ekonomie* 1997, č. 5. Do diskuse ze zapojili ještě další autoři: Guba, M. - Stiller, V. - Arlt, J.: *Vztah mezi peněžní zásobou a vývojem inflace v l. 1993 - 1996*. *Finance a úvěr* 1997, č. 3 a 4; Holub, T.: *Analýza inflace v ČR*. *Finance a úvěr*, 1997, č. 12

³⁷ Autonomií zde nelze chápat politickou nezávislost centrální banky na vládě, ale nezávislost ekonomickou ve smyslu schopnosti určovat výše peněžní zásoby a úrokových sazeb bez ohledu na vývoj platební bilance. (Frait, J.: *Autonomie monetární politiky a monetární přístup k platební bilanci /aplikace na ČR v letech 1992 - 1995/*. *FÚ* 1996, č. 5)

³⁸ Frait, J.: *tamtéž*, Jonáš, J.: *Měnová politika a měnový kurz*. *FÚ* 1996, č. 1

til do hledání odpovědi na zásadní otázku, a to zda bude ČNB v podmínkách pevného měnového kurzu a silícího přílivu zahraničního kapitálu i nadále schopna udržet kontrolu nad vývojem peněžní zásoby, nebo zda kompenzující kapitálové toky odsoudí její měnovou politiku k neúčinnosti.

Odpověď na tuto otázku byla hledána v rámci monetárního přístupu k platební bilanci. Jeho empirická verifikace byla spojena s kvantifikací tzv. kompenzačního (offsetového resp. offset) koeficientu,³⁹ jehož empirickými odhady se zabývaly 4 významné studie.⁴⁰ Jejich empirické odhady velikosti kompenzačního koeficientu sice nevykazovaly shodné hodnoty, a to především v důsledku toho, že zvolily odlišné ekonometrické přístupy k jejich odhadu, v žádné z nich však kompenzační koeficient nepřekročil horní hranici hodnot dosahovaných v sousedních zemích, takže se česká měnová politika jevila stejně účinná jako politika okolních centrálních bank.⁴¹

Od konce roku 1995 však kompenzační koeficient⁴² rostl a sterilizace zahraničního kapitálu připlouvajícího do ČR se stávala velice nákladnou a přestávala plnit svou základní funkci, tj. snižování tlaku na růst peněžní zásoby.

³⁹ Podle metodiky ČNB se offset koeficient (b) počítal ze vztahu $PZK_t = PZK_0 - bUCB_t$, kde PZK_t značí příliv zahraničního kapitálu v čase t , PZK_0 je příliv kapitálu nezávislý na operacích centrální banky a UCB_t je rozsah sterilizačních opatření centrální banky (Výroční zpráva ČNB za rok 1995, s. 41). Hodnota offset koeficientu se obvykle pohybuje od 0 do 1, přičemž hodnoty blízké 0 ukazují na téměř úplnou autonomii centrální banky, zatímco hodnoty blízké jedné představují faktickou ztrátu její autonomie, tj. neschopnost centrální banky omezit autonomní růst likvidity na mezibankovním trhu, a tedy i na následný růst peněžní zásoby. Vypočtené hodnoty se pro roky 1994 - 1995 pohybovaly v intervalu 0,3 - 0,6, což zhruba odpovídalo i výpočtům Mezinárodního měnového fondu. V této souvislosti se pak objevovala doporučení na změnu nastavení resp. režimu měnového kurzu koruny (srv. např. Jonáš, J.: Měnová politika a měnový kurz. FÚ 1996, č. 1). S polemikou vůči tomuto názoru vystoupila K. Šmídková, podle jejíž výpočtů se offset koeficient v uvedeném období pohyboval v mezích 0,2 - 0,4, což byly zhruba stejné hodnoty jako u okolních centrálních bank. Z toho vyvozovala, že účinnost nástrojů české měnové politiky není menší než v okolních zemích a že prostor pro případnou sterilizaci přílivu zahraničního kapitálu není ještě zcela vyčerpán. (Šmídková, K.: Účinnost měnové politiky a poptávka po penězích. FÚ, 1996, č. 8)

⁴⁰ Šmídková, K.: Search for the Offset Coefficient. ČNB, IE, Praha 1995; Teja, R.: Offset Coefficient for the Czech Republic. MF ČR, Praha 1995; Frait, J.: Autonomie monetární politiky a monetární přístup k platební bilanci: aplikace na ČR v letech 1992 - 1995. FÚ, 1996, č. 5; Tomšík, V.: Příliv zahraničního kapitálu, režim devizového kurzu, vyvážovací procesy platební bilance a autonomie monetární politiky ČR v letech 1993 - 1996. Studie NHÚ Josefa Hlávky č. 6, Praha 1997

⁴¹ podrobněji viz Tomšík, V.: Výzkum kompenzačního koeficientu v podmínkách ČR v letech 1993 - 1996. FÚ 1998, č. 6

⁴² Tomšík, V.: tamtéž

V dané situaci by mohla vhodně řešení nabídnout restriktivní fiskální politika, pro tu však nebyly vytvořeny politické podmínky, nehledě k tomu, že chyběly zkušenosti s volbou nezbytné míry rozpočtové restrikce. Proto byly cesty k řešení tohoto problému hledány jak v oblasti kurzové politiky, tak i v opuštění stávajícího transmisního mechanismu měnové politiky.

V oblasti kurzové politiky se namísto původních diskusí o udržitelnosti současné úrovně nominálního měnového kurzu a případném způsobu a rozsahu její změny, se stále větší naléhavostí objevovala doporučení hledat cestu v úpravě stávajícího kurzového režimu či dokonce v jeho opuštění.⁴³

Jedním z navrhovaných řešení bylo zvýšení míry kurzového rizika pro krátkodobé investory rozšířením flukтуаčního pásma. To mělo také umožnit zvýšení kontroly centrální banky nad měnovým vývojem, případně pomoci zjistit, které z protisměrných vlivů v platební bilanci působících na vývoj měnového kurzu mají větší váhu.

Rozšíření flukтуаčního pásma na +/- 7,5 %, ke kterému skutečně došlo v únoru 1996, bylo větší, než očekávala většina ekonomů. Spolu s dalšími faktory, především s růstem deficitu běžného účtu a zvýšenou nejistotou na politické scéně, vyvolalo odliv krátkodobého kapitálu. Ten spolu s pokračujícím schodkem běžného účtu sice vyústil do schodku souhrnné platební bilance, nenásledoval však pokles, ale naopak růst peněžní zásoby spojený s růstem peněžního multiplikátoru v první polovině roku 1996.

V souvislosti s tím provedla centrální banka v květnu a v červnu výrazná restriktivní opatření (dvojnásobné zvýšení úrokových sazeb a zvýšení sazeb povinných minimálních rezerv s platností od konce srpna), která vedla k zastavení růstu peněžního multiplikátoru a ke zpomalení růstu peněžní zásoby z 19 % v prvním čtvrtletí 1996 na téměř polovinu.⁴⁴ Následné zpomalení růstu dovozních cen způsobilo, že poprvé od roku 1994 se v roce 1996 zpomalil růst indexu spotřebitelských cen očištěného o ceny potravin a bydlení. Na

⁴³ Srv. např. Dědek, O.: *Currency Convertibility and Exchange Rate Policies in the Czech Republic*. ČNB, IE, 1995. Klaus, V.: *Ekonomická teorie a realita transformačních procesů*. Management Press, Praha 1995. Nešvera, V.: *Dilema úrokové politiky*. FÚ 1995 č. 4. Dědek, O. - Derviz, A.: *Kurzová politika rozšířeného oscilačního pásma*. ČNB, IE, 1996. Janáčková, S.: *Dilemata české měnové politiky*. FÚ 1996, č. 1. Jonáš, J.: *Měnová politika a měnový kurz*. FÚ 1996, č. 1

⁴⁴ Za rok 1996 dosáhl přírůstek peněžní zásoby 9,2 % při očištění o měnově neutralizovaný vklad SPT Telecomu jeho uložením u ČNB.

přelomu let 1996 a 1997 pak došlo i ke zpomalení růstu spotřebitelských cen.⁴⁵

Měnová restrikce však vedla také ke zpomalení růstu produkce na konci roku 1996 a k jejímu poklesu v roce 1997. To zpochybňovalo pozitivní hodnocení rozšíření flukтуаčního pásma.⁴⁶ Ukázalo se, že monetární restrikce, i když její účinky byly v souvislosti s rozšířením flukтуаčního pásma výraznější, není v případě, kdy nahrazuje restrikci fiskální, příliš vhodným nástrojem pro řešení problémů spojených s vnější nerovnováhou a že jediným možným řešením je důsledná koordinace obou typů politik.⁴⁷

Růst peněžní zásoby v první polovině roku 1996, který si vynutil následnou monetární restrikci, byl posléze kvalifikován jako důsledek otálení ČNB s rozšířením flukтуаčního pásma resp. se změnou kurzového režimu, což se později stalo předmětem téměř všeobecné kritiky.⁴⁸

Podle V. Kreidla⁴⁹ by pravděpodobně bývalo lepší, kdyby centrální banka rozšířila flukтуаční pásmo a zvýšila úrokové sazby asi o rok či dva dříve, tedy ještě před tím, než začal příliv krátkodobého kapitálu a problémy s tempem růstu peněžní zásoby a kdy „hrozilo“ zrušení mzdové regulace. Tehdy by zřejmě zhodnocování kurzu nemuselo vyvolávat obavy z dopadu na obchodní bilanci, protože ještě nebyl zcela vyčerpán mzdový polštář.⁵⁰

Někteří odborníci zdůrazňovali, že centrální banka měla přistoupit k rozšíření flukтуаčního pásma či dokonce ke změně kurzového režimu již v souvislosti s uvolněním

⁴⁵ Blíže viz Janáček, K. a kol.: Česká ekonomika začátkem roku 1997. Hospodářské trendy 11, Komerční banka 1997

⁴⁶ srv. např. Hrnčíř, M.: Devizový kurz a transformace české ekonomiky. Sborník vědecké konference k 70. výročí centrálního bankovníctví v ČR. ČNB, Praha 1996. Posléze však byla restriktivní opatření ČNB označena za hlavní příčinu poklesu tempa růstu HDP v roce 1996 a hospodářského poklesu české ekonomiky v letech 1997 - 1998. I když vliv měnové restrikce na vývoj reálné ekonomiky nelze popřít, představitelé ČNB tuto jednoznačnou interpretaci popírali, s tím, že odmítají převzít zodpovědnost za potíže, které s měnovou restrikcí podle jejich názoru v žádném případě nesouvisely.

⁴⁷ Čihák, M.: Ohlédnutí za flukтуаčním pásmem koruny. FÚ 1997, č. 10

⁴⁸ viz předchozí část příspěvku

⁴⁹ Kreidl, V.: Česká měnová politika v roce 1996 a 1997. Bankovníctví 1997, č. 3, s. 6

⁵⁰ Podle názoru M. Mandela a V. Tomšíka se ČNB obávala, že při dřívějším rozšíření pásma oscilace by se kurz ocitl pod silným apreciačním tlakem, který by korunu dlouhodobě „přitlačil“ k dolnímu pásmu oscilace, čímž by nebyl vyvolán efekt zvýšeného rizika. (Vývoj platební bilance ČR v období 1990 - 1996 pohledem modelu IS-LM-BP. PE 1997, č. 4, s. 546

toků na finančním účtu platební bilance v říjnu 1995.⁵¹

Ani pokles míry inflace nebyl přijímán jednoznačně, dokonce lze říci, že jeho hodnocení vyznělo značně rozporně. Jeden z krajních pólů představuje hodnocení V. Klause, podle nějž ČNB měnovou restrikcí provedla v takovém rozsahu zcela záměrně s cílem zabrzdit ekonomiku, přičemž nedokázala odhadnout, jak silný bude její efekt na reálný hospodářský vývoj.⁵²

Skutečnost, že ČNB ještě v dubnu roku 1997 prognózovala růst HDP mezi 3 - 4 %, a to i přes varování tehdejšího ministra financí, však může být interpretována spíše jako důkaz, že si centrální banka nebyla plně vědoma potenciální hloubky svého zásahu.⁵³

Podle P. Kysilky centrální banka spatřovala ve snižování inflace cestu k zastavení reálného zhodnocování kruzu koruny, a tím i k řešení problémů narůstající vnější nerovnováhy ekonomiky. K tomu si však otálením s rozšířením oscilačního pásma nevytvořila dostatečný časový prostor, takže restrikce vedla rychleji k posílení koruny než k omezení inflace, a tím i k přílivu krátkodobého kapitálu a ke zhoršení problému vnější nerovnováhy.⁵⁴

Jiní odborníci⁵⁵ naopak prosazovali výraznější snižování inflace, a to hned ze dvou důvodů. V první řadě to byla nespokojenost s tím, že i když československá a posléze i česká centrální banka dosáhla nesporných úspěchů v boji s inflací, které byly nesrovnatelné s ostatními zeměmi střední a východní Evropy, setrvávání míry inflace na hodno-

⁵¹ srv. např. Klaus, V.: Tři roky po měnové krizi: rekapitulace událostí a jejich souvislostí, aby se na některé věci nezapomnělo. Dyba, K.: Navrhovali jsme změnu kruzového režimu. Holman, R.: Měnová krize 1997 byla trest za nezavedení volného floatingu. Mach, P.: Měnová krize z hlediska změn devizových rezerv. Všechny články in Tři roky od měnové krize. Sborník textů. CEP, Praha 2000

⁵² Klaus, V.: cit. práce s. 13

⁵³ Kočárník, I.: Kočárník, I.: Postřehy jednoho z „obětních beránek“ měnové krize z roku 1997. in Tři roky od měnové krize. Sborník textů, CEP, Praha 2000, s. 31

⁵⁴ Kysilka, P.: Vnější rovnováha jako odraz souladu mezi mírou otevřenosti ekonomiky a parametry makroekonomické politiky. in Otevřenost české ekonomiky a vnější rovnováha: konec 90. let. Sborník textů. CEP, Praha 2001, s. 30

⁵⁵ srv. např. Kreidl, V.: cit. práce, Tůma Z. - Kreidl, V.: Stará kotva opuštěna, kde hledat novou? Ekonom 1996, č. 7; Tůma, Z.: Inflace v České republice a nezávislost centrální banky. Sborník vědecké konference k 70. výročí centrálního bankovníctví v České republice. ČNB, Praha 1996, Kysilka, P.: Inflace je léčitelná nemoc. Bankovníctví 1998, č. 1

tách kolem 10%⁵⁶ trvalo podle jejich názoru již příliš dlouho a hrozilo, že začne sehrávat negativní úlohu při tvorbě inflačních očekávání hospodářských subjektů.

Kromě toho tato - byť relativně stabilní - míra inflace převyšovala míru inflace v zemích našich hlavních obchodních partnerů, takže v důsledku existujícího inflačního diferenciálu docházelo k permanentnímu zhodnocování reálného měnového kurzu české koruny a v souvislosti s tím i k poklesu konkurenceschopnosti domácích výrobců na zahraničních trzích.

V této souvislosti někteří ekonomové trvali na razantnější protiinflační politice. V. Kreidl⁵⁷ doporučoval snížení inflace až na 5 %. Při té příležitosti varoval, že při vyšších mírách inflace může dojít k tlaku na devalvaci koruny a k rozvinutí devalvační spirály.

M. Mandel a V. Tomšík naopak dospěli na základě své analýzy⁵⁸ k závěru, že samotné další přitvrzení měnové politiky bez souběžné fiskální restrikce by bylo neúčinné, resp. by v nejlepším případě vedlo k návratu situace z let 1993 - 1995, a to k politice sterilizace devizových intervencí. Jako minimálně nezbytné řešení označili udržení vývoje peněžní zásoby M2 v tempech růstu do 10 % meziročního srovnání, které ovšem bude v nejlepším případě schopno pouze „zakonzervovat“ současný deficit běžného účtu platební bilance.

K nejradikálnějším odpůrcům myšlenky o nezbytnosti dalšího snižování inflace a zavedení nového transmisního mechanismu měnové politiky a měnově politické strategie v podobě cílování inflace patřil V. Klaus. V této souvislosti zdůrazňoval, že neklesající inflace udržující se na úrovni kolem 10 %, což bylo podstatně méně než v ostatních transformujících se ekonomikách, nebyla v období „klasického“ transformačního oživení s relativně nízkou mírou nezaměstnanosti, kterým roky 1994, 1995 a první polovina roku 1996 bezesporu byly, důvodem k jakémukoli zneklidňování. Na úsilí ČNB o snižování inflace doplatila česká ekonomika nejen ztrátou růstové dynamiky, ale posléze i hospodářským poklesem. Analogicky by mohla podle jeho názoru rychle pokračující

⁵⁶ Míra inflace dosáhla v roce 1992 11,1 %, 1994 = 10,0 %, 1995 = 9,1 %, 1996 = 8,8 % (Výroční zpráva ČNB 1996) a při vyloučení dopadu daňové reformy byla v roce 1993 také těsně nad úrovní 10 % (Tůma, Z.: *Inflace v České republice a nezávislost centrální banky*. Sborník vědecké konference k 70. výročí centrálního bankovníctví v České republice. ČNB, Praha 1996)

⁵⁷ V. Kreidl (*Česká měnová politika v roce 1996 a 1997*. Bankovníctví, 1997, č. 3) doporučoval snížení inflace až na 5 %. V souvislosti s tím varoval, že při vyšších mírách inflace může dojít k tlaku na devalvaci koruny a k rozvinutí devalvační spirály.

⁵⁸ Mandel, M. - Tomšík, V.: cit. práce, tamtéž

dezinflace při nezměněných inflačních očekáváních hospodářských subjektů vyústit do silného omezení hospodářského růstu.⁵⁹

P. Kysilka naproti tomu zdůrazňoval, že vedle politik prováděných na podporu desinflačního vývoje jsou mimořádně důležité zejména politiky zaměřené na zmírňování krátkodobých nákladů spojených s realizací protiinflační strategie jako nedeficitní či ještě lépe přebytkový vývoj veřejných rozpočtů, sladění mzdového vývoje s růstem produktivity práce a jasný návrat k transformační dlouhodobě neinflační prorůstové strategii.⁶⁰

Analogicky bylo i podle K. Šmídkové a M. Hrnčíře možno očekávat, že pokud dojde ke koordinaci hospodářských politik, a to v podobě konsensu o výši inflačního cíle, bude při cílování inflace společná orientace centrální banky a vlády na inflační cíl a jeho zohlednění při mzdových vyjednáváních redukovat krátkodobé náklady dezinflace.⁶¹

Stoupenci razantnějšího postupu vůči inflaci doporučovali změnu měnové strategie centrální banky spojenou s přechodem na cílování inflace prakticky jednoznačně. Střednědobý závazek ČNB, která díky svým výsledkům v úsilí o stabilitu měny nepostrádala potřebnou kredibilitu,⁶² dosáhnout určité míry inflace podle nich na jedné straně mohl sehrát pozitivní úlohu při tvorbě inflačních očekávání hospodářských subjektů a usnadnit tak proces dezinflace, na druhé straně mohl představovat novou nominální kotvu, jejíž úlohu přestal hrát po rozšíření flukтуаčního pásma měnový kurz.⁶³

Nepřímou podporou tohoto přístupu bylo upozornění, že možnosti využití kurzového rizika vytvořeného rozšířením oscilačního pásma k lepší kontrole růstu peněžní zásoby a domácí poptávky jsou omezené a že by se v budoucnosti mohly dokonce stát příčinou vzniku bludného kruhu či dokonce inflační spirály: omezení přílivu krátkodobého kapitálu - domácí úvěrová expanze - restriktivní opatření ČNB vedoucí k růstu úrokových sazeb a inflačního diferenciálu - obnovení přílivu krátkodobého kapitálu a opětná

⁵⁹ Klaus, V.: Tři roky po měnové krizi: rekapitulace událostí a jejich souvislostí, aby se na některé věci nezapomnělo. in Tři roky od měnové krize. Sborník textů. CEP, Praha 2000

⁶⁰ Kysilka, P.: Inflace je léčitelná nemoc. Bankovníctví 1998, č. 1

⁶¹ Šmídková, K. - Hrnčíř, M.: Přechod na strategii cílování inflace. FÚ, 1998, č. 4, s. 207

⁶² V tomto smyslu se vyjadřovala většina ekonomů, srv. např. Janáčková, S.: Dilemata české měnové politiky. FÚ 1996, č. 1. Pospíšil, J.: Inflace a nezávislost centrální banky - máme a můžeme už mít nějaké zkušenosti i v České republice? Sborník vědecké konference k 70. výročí centrálního bankovníctví v České republice. ČNB, Praha 1996. Tůma, Z. cit. práce.

⁶³ Kreidl, V.: obě výše cit. práce. Tůma, Z.: cit. práce. Tůma, Z. - Kreidl, V.: Stará kotva opuštěna, kde hledat novou? Ekonom 1996, č. 7

snaha odradit krátkodobé zahraniční investory spojená nejen s vysokými náklady, ale i s nebezpečím ztráty kredibility pro dlouhodobý kapitál.⁶⁴

Na skutečnost, že ve většině vyspělých tržních ekonomik dochází v průběhu 80. a 90. let k poklesu účinnosti tradičních zprostředkujících kritérií jako je peněžní zásoba, úrokové sazby nebo měnový kurz a že centrální banky přiznávají měnovým kurzům stále větší míru flexibility, takže standardní typ měnové politiky postupně ztrácí své kotvy a je nahrazován cílováním inflace, upozornila i řada dalších ekonomů.⁶⁵

Jak vyplývá z výše uvedeného, nalezlo zavedení cílování inflace⁶⁶, k němuž došlo od 1. 1. 1998,⁶⁷ v ČR řadu obhájců, ale nechyběly ani nesouhlasné postoje. Přechod ČNB na novou měnovou strategii⁶⁸ byl inspirován poznatky řady zemí, které tento krok učinily již dříve, přičemž řada z nich byla před jejím zavedením nucena - stejně jako ČR - opustit režim fixního měnového kurzu.

Kromě toho představovaly snahy o zavádění cílování inflace v řadě zemí určité vyvrcholení dlouhodobějšího nárůstu významu cenové stability, a to jak z národohospodářského hlediska, tak i v rámci měnové politiky. V ČR byl tento trend zachycen již

⁶⁴ Hájek, M. a kol.: Makroekonomická analýza české ekonomiky 1996. PE 1997, č. 3, s. 324

⁶⁵ srv. např. Vencovský, F.: Nové směry ve strategii peněžní politiky v zemích Evropské unie. Profesorská přednáška na VŠE Praha 1996. Frait, J. - Zedníček, R.: Časová nekonzistence a cíle měnové politiky. Bankovníctví 1996 č. 16

⁶⁶ Ke zdůvodnění této volby domácího ekvivalentu pro anglické označení „inflation targeting“ viz Martinová, O.: K terminologii. FÚ, 1998, č. 4

⁶⁷ Tento krok bylo možno dříve či později očekávat, neboť bylo zřejmé, že se k němu centrální banka již delší dobu chystala. Srv. např. ČNB chce čistou inflaci kolem šesti procent. Mladá fronta DNES, 23. 12. 1997, s. 14

⁶⁸ Realizace nové měnové strategie byla zpočátku spojena se stanovením inflačního cíle pro tzv. čistou inflaci (její interval byl pro rok 1998 stanoven v rozmezí 5,5 - 6,5 %, pro rok 2000 v rozmezí 3,5 - 5,5 %) představující pohyb neregulovaných cen (tj. CPI po vyloučení položek regulovaných cen) očištěný dále o vliv nepřímých daní, případně rušení dotací. (Podrobněji viz Zpráva o inflaci. Duben 1998, s. 7 - 8). Při vlastním operativním řízení pak centrální banka na základě prognostického ekonometrického modelu vycházejícího z většího množství indikátorů a respektujícího časové zpoždění mezi změnou operativní úrokové sazby a inflací provádí prognózy vývoje čisté inflace a srovnává je s inflačním cílem. Pokud prognózované hodnoty převyšují stanovený cíl, centrální banka zvyšuje svou operativní úrokovou sazbu, kterou je dvoutýdenní REPO sazba, v opačném případě dochází k jejímu snižování. (Podrobněji viz. Šmídková, K. - Hrnčíř, M.: Přechod na strategii cílování inflace. FÚ, 1998, č. 4) Těžiště měnové politiky se tak přesunulo od orientace na vývoj peněžní zásoby jako hlavního faktoru působícího na změny cenové hladiny k operativnímu řízení krátkodobých úrokových sazeb. Od dubna 2001 přešla ČNB na cílování celkové inflace vyjádřené indexem spotřebitelských cen, přičemž mechanismus i operativní úrokové sazby zůstaly zachovány.

na samotném počátku transformace v rámci institucionální reformy, která stanovila jako základní cíl měnové politiky stabilitu měny a vybavila centrální banku nezávislostí odpovídající novému trendu v pojetí měnové politiky.⁶⁹

V souvislosti se zavedením cílování inflace nešlo tedy o změnu cíle měnové politiky, ale o volbu nové strategie k jeho realizaci, která byla ve shodě s významem, jaký se tomuto pojetí měnové politiky přikládal v zemích Evropské unie, a která znamenala odmítnutí možnosti využívat měnovou politiku k plnění krátkodobých záměrů hospodářské politiky a nadřazení dlouhodobých měnových zájmů krátkodobým hospodářsko politickým cílům.⁷⁰

Mezi hlavními důvody pro zavádění nové měnové strategie v ČR byl uváděn především důraz na stabilitu cenové hladiny, který se od počátku 80. let stával prvořadou prioritou ve většině vyspělých tržních ekonomik, změna režimu měnového kurzu v květnu 1997, dále pak široké a intenzivní propojování národních ekonomik v oblasti financí zneusnadňující orientaci a řízení peněžní zásoby zejména v malých otevřených ekonomikách.

Vedle nejistot spojených s globalizací finančních trhů, masivními toky kapitálu, nestabilitou peněžního multiplikátoru a rychlosti obratu peněz, nevyjasněností vztahů mezi měnovými agregáty a inflací, které způsobovaly řadu problémů při aplikaci monetaristického transmisního mechanismu, byly dále uváděny také pozitivní zkušenosti zahraničních centrálních bank cílujících inflaci (rostoucí předvídatelnost reakční funkce centrální banky v prostředí zvýšených nejistot a zlepšení komunikace mezi centrální bankou a veřejností, větší flexibilita při měnově politickém rozhodování, náhrada za ztrátu nominální kotvy v podobě fixního měnového kurzu, možnost intenzivněji ovlivňovat tvorbu inflačních očekávání a tak i snadněji prosazovat postupnou dezinflaci, zvýšená transparentnost spolu s narůstající předvídatelností a do jisté míry i „automatickostí“ měnově politického rozhodování).⁷¹

Uvedená zdůvodnění se však setkala s celou řadou námitek. Ty se týkaly především samotného přechodu na novou měnovou strategii. V první řadě byl vyvrácen názor, že

⁶⁹ Hrnčíř, M. - Šmídková, K.: Inflační cílení v zahraničí - zdroj informací a inspirace pro českou měnovou politiku. *Bankovníctví* 1998, č. 7

⁷⁰ Vencovský, F.: K nové měnové strategii ČNB. *Bankovníctví* 1998, č. 4

⁷¹ podrobněji viz Hrnčíř, M. - Šmídková, K.: Přechod na strategii cílování inflace. *FÚ* 1998, č. 4, s. 210 a násl.; Šmídková, K. - Hrnčíř, M.: Inflační cílení v zahraničí - zdroj informací a inspirace pro českou měnovou politiku. *Bankovníctví* 1998, č. 7, s. 7 - 8

došlo k selhání monetaristického transmisního mechanismu a že musí být nahrazen mechanismem novým.

Na spornost argumentu týkajícího se poklesu důvěry ČNB ve stabilitu vztahu mezi peněžní zásobou a inflací poukázal M. Čihák s T. Holubem s tím, že empirické analýzy pro ČR zatím vesměs potvrdily statistickou významnost vztahu mezi měnovým agregátem M2 a inflací, i když tento vztah nevycházel pro použité délky časového zpoždění jako proporcionální, jak by tomu mělo být podle monetaristické teorie. Při testování platnosti těchto závěrů pro poslední časové období dospěli oba autoři k výsledkům zhruba odpovídajícím analýzám pro starší období⁷², což potvrzovalo relativní stabilitu vztahu mezi oběma proměnnými. Naproti tomu se jim nepodařilo nalézt stabilní vztah mezi úrokovými sazbami a inflací, který byl nezbytný pro zvolenou variantu cílování inflace, z čehož vyvodili závěr, že současná podoba cílování inflace je méně transparentní než předchozí režim, neboť centrální banka sice oznámila svůj cíl i nástroj, ale nikoli předpokládaný převodový mechanismus.⁷³

Ani některé standardně známé problémy, které provázely cílování peněžní zásoby v praxi ČNB, především nestabilita rychlosti obratu peněz a obtíže při operativním řízení spojené mj. i se způsobem stanovení povinných minimálních rezerv, nebyly akceptovány

⁷² Vedle prací uváděných oběma autory je v této souvislosti třeba zmínit práci M. Guby, V. Stillera a J. Arlta, kteří na základě své analýzy vztahu mezi vývojem peněžní zásoby a vývojem inflace v letech 1993 - 1996 shledali dosavadní přístup ČNB založený na cílování peněžní zásoby jako odůvodněný s tím, že jiné přístupy aplikované v zahraničí jako např. cílování inflace či využívání měnového kurzu jako mezicíle měnové politiky nepředstavují v současných podmínkách pro českou ekonomiku vhodnější alternativu. (Guba, M. - Stiller, V. - Arlt, J.: Vztah mezi vývojem peněžní zásoby a vývojem inflace v l. 1993 - 1996 (2. část) FÚ 1997, č. 4)

⁷³ Čihák, M. - Holub, T.: Cílování inflace v ČR: staré víno v nových lahvích. FÚ 1998, č. 4, s. 225 -226. Neexistenci jednoznačně definovatelného transmisního mechanismu mezi krátkodobou úrokovou sazbou a mírou inflace označil za klíčový problém cílování inflace také V. Kosmata. V této souvislosti označil za „kámen úrazu“ cílování inflace stanovení takové úrovně úrokové sazby PRIBOR, která by vedla ke splnění inflačního cíle. (Kosmata, V.: Přejít na inflační cílení má svá rizika. HN 13. 1. 1998, s. 6)

jako důvod pro jeho opuštění.⁷⁴ Jako zásadní teoretický i praktický problém, který mohl vést ke zpochybnování aplikace monetaristického transmisního mechanismu, byla uváděna kvantifikace monetárního cíle z pohledu vnější rovnováhy s tím, že tento problém není řešen ani v rámci cílování inflace.⁷⁵ V této souvislosti bylo zdůrazňováno, že zatímco praktická realizace monetaristického transmisního mechanismu se dlouhodobě vztahovala k inflaci, a z hlediska krátkého a středního období byla směřována také k nominálnímu důchodu (případně k rovnováze běžného účtu), u cílování inflace dochází k orientaci pouze na jeden střednědobý a dlouhodobý „monocíl“, což v sobě obsahuje riziko větší nestability u ostatních makroekonomických veličin. Cílování inflace tak výrazně omezuje možnost centrální banky ovlivňovat resp. stimulovat jiné ekonomické cíle jako je růst reálného produktu či rovnováha běžného účtu platební bilance.⁷⁶

Zatímco cílování peněžní zásoby dávalo v případě ohrožení vnější ekonomické rovnováhy možnost výraznějšího znehodnocení měnového kurzu, cílování inflace bylo de facto spojeno s implicitním opětovným zavedením fixního měnového kurzu s relativně úzkým oscilačním pásmem, neboť úsilí o dosažení inflačního cíle musí být nutně spojeno s požadavkem určit a respektovat maximálně možnou hranici pro depreciaci domácí měny, která by neakcelerovala růst cen.⁷⁷

Další skupina námitek se týkala vlastní realizace cílování inflace, v první řadě pak

⁷⁴ Nepodloženosť argumentů proti cílování peněžní zásoby, které byly uváděny při přechodu na cílování inflace dokumentovali na základě aplikace P* modelu pro ČR Frait, J. - Komárek L. - Kulhánek, L.: Analýza dynamiky inflace v ČR pomocí P* modelu. FÚ 1998, č. 11. Skutečnost, že ve své první Zprávě o inflaci ČNB uváděla jako faktory inflace na prvním místě peněžní agregáty, na druhém místě úrokové sazby a na třetím indexy měnového kurzu vedla M. Čiháka a T. Holuba k závěru, že se centrální banka reálně přiklání k hybridnímu transmisnímu mechanismu a že výroky o „neprokázaném vztahu mezi peněžní zásobou a inflací“ byly spíše marketingové kampaně při přechodu na cílování inflace než výsledkem empirických analýz. (Čihák, M. - Holub, T.: Cílování inflace: předběžné poznatky a zkušenosti z prvních měsíců. HN 14. 7. 1998, s. 6

⁷⁵ Mandel, M.: Cílování inflace a peněžní zásoby při exogenních šocích (komparace z pohledu modelu AS-AD-NX). FÚ 1998, č. 4

⁷⁶ Kosmata, V.: cit. práce, Mandel, M.: cit. práce, Tomšík, V.: Možné problémy současně měnové politiky. Bankovníctví 1998, č. 13

⁷⁷ Čihák, M. - Holub, T.: cit. práce, Kosmata, V.: cit. práce, Mandel, M.: cit. práce, Tomšík, V.: cit. práce. Z. Tůma již na samém počátku diskuse uváděl, že stanovený inflační cíl je natolik restriktivní, že bude vyžadovat, aby ČNB věnovala měnovému kurzu ještě větší pozornost než dosud. To bylo v rozporu s názorem představitelů centrální banky, kteří se vyjadřovali v tom smyslu, že na devizový trh hodlá ČNB vstupovat až v okamžiku, kdy se kurz bude významně měnit. (srv. ČNB chce čistou inflaci kolem šesti procent. Mladá fronta DNES, 23. 12. 1997, s. 14

stanovení inflačního cíle. Pokud měl inflační cíl potřebným způsobem ukotvit inflační očekávání a měl založit žádoucí disciplinu, musel být dostatečně ambiciózní. Na druhé straně však hrozilo, že pokud nebude nastaven dostatečně citlivě vzhledem k parametrům vývoje ekonomiky,⁷⁸ musí být jeho splnění spojeno s nadměrnou monetární restrikcí, která by mohla vést k prohloubení nerovnováhy, a to jak v oblasti vnější, tak i v oblasti podnikové a bankovní.⁷⁹ Riziko ztráty kredibility resp. neúměrně vysokých nákladů na realizaci nerealisticky ambiciózního cíli si však uvědomovali i představitelé ČNB.⁸⁰

S námitkami se setkala i volba ukazatele čisté inflace, a to hned z několika hledisek. V první řadě z toho důvodu, že šlo o zcela novou kategorii, pro niž neexistovalo historické srovnání⁸¹ a chyběly zkušenosti s citlivostí jejích reakcí na různé inflační situace. Kromě toho se objevovaly námitky týkající se srozumitelnosti ukazatele čisté inflace pro nejširší veřejnost,⁸² neboť se předpokládalo, že jeho prostřednictvím bude možno ovlivňovat inflační očekávání hospodářských subjektů.

ČNB zde však vycházela z přesvědčení, že na základě postupu deregulací a adaptace cenových relací se bude rozdíl mezi celkovou a čistou inflací postupně snižovat, takže střednědobý cíl pro konec roku 2000 již předpokládal konvergenci indexu čisté inflace k indexu inflace celkové.⁸³

⁷⁸ Stanovený inflační cíl v rozmezí 5,5 - 6,5 % pro rok 1998 a 3,5 - 5,5 % pro rok 2000 byl řadou odborníků považován za příliš ambiciózní, jehož plnění bude vyžadovat nadměrně restriktivní měnovou politiku. Srv. např. Tůma, Z. in cit. práce, Kosmata, V.: cit. práce, Mandel, M.: cit. práce, Tomšík, V.: cit. práce

⁷⁹ Nadměrná restrikce by v situaci silného přeúvěrování české ekonomiky spojeného s vysokým podílem klasifikovaných a nedobytných úvěrů mohla vést nejen k poklesu růstu skutečného produktu, ale i potenciálu. V této souvislosti byly uváděny neblahé zkušenosti západních ekonomik v 70. letech, kdy důsledná proticyklická a protiinflační měnová politika spojená se silnou restrikcí vyústila do zastavení investic, do podkapitalizace těchto ekonomik a do snížení tempa růstu potenciálního produktu, což se při následném oživení projevilo růstem cen spojeným nejen s rovními šoky, ale i s růstem agregátní poptávky, který narážel na nedostatek výrobních kapacit. Srv. Tomšík, V.: cit. práce, s. 23

⁸⁰ Hrnčíř, M. - Šmídková, K.: ČNB určila dva časové horizonty cílování inflace. HN 5. 6. 1998

⁸¹ Proto se někteří odborníci vyslovovali spíše pro jádrovou inflaci (tj. spotřebitelský koš očištěný o vliv regulovaných cen a změn cen potravin, na něž má měnová politika jen velmi zprostředkovaný vliv), kterou již předtím vykazoval ČSÚ, ta však představovala příliš úzké vymezení cíle (přibližně 45 % spotřebitelského koše), které by nedávalo jasnou představu o budoucím vývoji inflace. Srv. např. Čihák, M. - Holub, T.: cit. práce s. 227 - 228, Křovák, J. : S cílováním inflace přichází ústřední banka v nevhodný čas a s předčasnými ambicemi. HN, 26. 2. 1998, s. 7

⁸² Tomšík, V.: cit. práce

⁸³ Hrnčíř, M. - Šmídková, K.: Nová kotva měnové politiky. HN 29. 4. 1998; ČNB určila dva časové horizonty cílování inflace. HN 5. 6. 1998

Někteří ekonomové pak dále poukazovali na skutečnost, že ani čistou inflaci nebude možno uspokojivě vysvětlit pouze na základě monetárních faktorů, neboť významnými vysvětlujícími činiteli mohou být také ceny komodit na světových trzích, vztah mezi růstem produktivity práce a mezd a ve střednědobém horizontu i pokračující srovnávání relativních cen v ČR a v zemích EU.⁸⁴

Jako problematické bylo označováno i stanovení inflačního cíle v podobě intervalu s tím, že odvrací pozornost od jeho středu směrem k horní hranici a je spojeno s rizikem, že se inflace může ocitnout mimo něj, což by mohlo mít závažné důsledky pro kredibilitu centrální banky.⁸⁵

Pokud jde o časový horizont plnění inflačního cíle, směřovaly námitky jednak k volbě prosincového, tedy koncového, a nikoli průměrného celoročního indexu, jednak k samotnému přijetí krátkodobého ročního cíle s tím, že nástroje centrální banky většinou nemají tak rychlý účinek, aby byla schopna plně kontrolovat inflaci v horizontu několika měsíců, přičemž dochází k poměrně značným časovým zpožděním.⁸⁶ Jako vhodný se jevil (i vzhledem k šestiletému funkčnímu období guvernéra ČNB) spíše tříletý časový horizont, který by centrální bance poskytl větší prostor.⁸⁷

Tento názor podporovaly i zkušenosti centrálních bank, které přešly na strategii cílování inflace již dříve a které svůj první inflační cíl situovaly zpravidla do období dvou let po jejím zavedení. V souladu s tím považovala i ČNB za hlavní cíl a současně i za svůj závazek vůči veřejnosti svůj střednědobý cíl. Krátkodobý cíl stanovený v ročním horizontu chápala spíše jako podklad pro ukotvení hospodářské a měnové politiky, mzdových vyjednávání a rozhodovacích procesů v ročním výkladu a jako určitou možnost pro ověření vhodnosti zvolené strategie pro dosažení střednědobého cíle a pro její případnou korekci. Jednoznačná orientace na koncovou hodnotu střednědobého cíle pak byla zdůvodňována potřebou flexibility v jednotlivých fázích jeho realizace.⁸⁸

⁸⁴ V této souvislosti označil J. Křovák zavedení cílování inflace za „přicházející v nevhodný čas a s předčasnými ambicemi“ (Křovák, J.: cit. práce). K vyrovnávání cenových hladin viz např. Zamrazilová, E. - Holub, T.: Inflační diferenciál se nesníží. *Ekonom* 1997, č. 10

⁸⁵ Čihák, M. - Holub, T.: cit. práce

⁸⁶ Řada empirických analýz uváděla vliv změny růstu peněžní zásoby v ČR na inflaci se zpožděním minimálně 5 - 8 měsíců.

⁸⁷ Čihák, M. - Holub, T.: cit. práce

⁸⁸ Hrnčíř, M. - Šmídková, K.: ČNB určila dva časové horizonty cílování inflace. *HN* 5. 6. 1998

Zdůrazňována byla i kritická závislost úspěšnosti cílování inflace na zpracování kvalitních prognóz inflačního vývoje a na kvantifikaci vztahu mezi stupněm neplnění inflačního cíle a požadavky na změnu operativního kritéria s tím, že existující objektivní problémy spojené s prognózováním a nejasnosti v politice úrokových sazeb by mohly vyústit až do pouhého sledování vývoje měnového kurzu a jeho ovlivňování pomocí úrokových sazeb.⁸⁹

Představa ČNB, že se jí na základě dosaženého relativně vysokého stupně kredibility za pomoci cílování inflace podaří ovlivnit inflační očekávání hospodářských subjektů natolik, že proces dezinflace nebude spojen s náklady v podobě poklesu reálného produktu, vycházející z koncepce racionálních očekávání narážela na přesvědčení některých odborníků podložených výsledky empirických studií, že v ČR jsou hospodářským subjektům vlastní spíše adaptivní očekávání a že jejich důvěru mohou významnou měrou ovlivnit i významnější odchylky mezi vývojem skutečné a čisté inflace způsobené deregulací cen či daňovými úpravami.⁹⁰

Kritiky nebyl ušetřen ani další z hlavních argumentů ČNB pro zavedení cílování inflace - snaha o vytvoření stabilního makroekonomického prostředí spojeného s postupným snižováním inflace, které považovala za nezbytnou podmínku pro nastoupení cesty stabilního dlouhodobého růstu ekonomiky.⁹¹

M. Čihák a T. Holub v této souvislosti s odvoláním na empirické studie MMF zdůrazňovali, že se dlouhodobé negativní účinky inflace na ekonomiku začínají projevovat až při mírách inflace převyšujících 8 %, takže přínos spojený se zaváděním cílování inflace nemusí být pro českou ekonomiku tak vysoký, jak očekává ČNB.⁹²

J. Frait. sice připustil oprávněnost přechodu k přímému cílování inflace a snahy o nízkou inflaci, zpochybnil však názor, že snížení inflace je důležité pro zvyšování tempa růstu ekonomiky. Na základě studia nejnovějších teoretických poznatků a empirických prací z této oblasti dospěl k závěru, že mezi inflací a dlouhodobým růstem skutečně existuje negativní vztah, který se však projevuje až od míry inflace dosahující 15 - 20 %

⁸⁹ Mandel, M.: cit. práce

⁹⁰ Mandel, M.: cit. práce, Tomšík, V.: cit. práce

⁹¹ Ve Zprávě o měnovém vývoji v ČR za období leden - září 1997 je doslova uvedeno: „Stabilní makroekonomické prostředí, charakterizované především klesajícím tempem růstu inflace, je chápáno jako nezbytná podmínka pro stabilní dlouhodobý ekonomický růst.“ (s. 26)

⁹² Čihák, M. - Holub, T.: cit. práce

ročně, přičemž ale ani v tomto případě není kvantitativně příliš silný. Pro nižší míry inflace není podle jeho názoru prakticky možné identifikovat jakýkoli jednoznačný vliv na dlouhodobý růst. U ekonomiky s relativně stabilní inflací kolem 10 % pak nelze podle něj předpokládat, že snížení inflace významněji přispěje ke zvyšování ekonomického růstu.⁹³

Měnově politická praxe však ukázala, že se ČNB⁹⁴ stejně jako ostatně i řadě dalších centrálních bank příliš nedaří naplňovat své měnové cíle. To se odráží i ve vývoji ekonomické teorie, která přináší nové pohledy na možnost centrální banky ovlivňovat měnový vývoj a v souvislosti s tím i na její postavení jako nezávislé „bankovní moci“ zodpovědné de facto jen sobě samé. S rostoucí intenzitou dochází k prosazování teorie endogenních peněz⁹⁵, v níž má rakouská škola dlouholetou tradici.

I když jsem spíše stoupencem postkeynesovského přístupu k postavení a úloze peněz a centrálních bank⁹⁶ v moderních tržních ekonomikách, domnívám se, že v řadě bodů není jeho učení v rozporu s Hayekovým odmítáním představy, že centrální banky jsou schopny účinně regulovat množství peněz v ekonomice. Proto bych své vystoupení ráda

⁹³ Frait, J.: Je snížení inflace podmínkou pro obnovení ekonomického růstu. *Bankovníctví* 1998, č. 8, s. 7

⁹⁴ ČNB se podařilo splnit stanovený cíl jen ve zcela sporadických případech. Zcela obvyklou záležitostí se stalo jeho „podstřelování“. Inflační cíl byl zpočátku stanoven v podobě čisté inflace (počítané na neúplném spotřebním koši, z nějž byly vyloučeny položky s regulovanými cenami a položky s cenami ovlivňovanými jinými administrativními opatřeními, např. nepřímými daněmi). Pro rok 1998 byl interval 5,5 - 6,5 % byl cíl „minut“ na úrovni 1,7 %; v roce 1999 namísto stanovených 4,0 - 5,0 % dosáhla čistá inflace hodnoty 1,5 %; plánovaný interval pro rok 2000 činil 3,5 - 5,5 % (stanoveno v roce 1997), skutečnost 3,0 %. Rok 2001 byl prvním rokem, kdy se ČNB podařilo splnit inflační cíl (předpoklad pro čistou inflaci 3,0 - 1,0 %, skutečnost 2,4%). Poprvé od zavedení režimu cílování inflace vyjádřila ČNB svůj inflační cíl v tomto roce také v podobě celkové inflace, a to v intervalu 4,3 - 5,8 % (skutečnost 4,7 %), a po projednání jeho výše s vládou. Stanovení inflačního cíle v podobě celkové inflace mělo dále zprůhlednit měnovou politiku ČNB směrem k veřejnosti, neboť se předpokládalo, že takto stanovený cíl bude pro veřejnost srozumitelnější a přispěje ke stabilizaci inflačních očekávání v příštích obdobích. V roce 2002 byl cíl stanoven v podobě cílového pásma celkové inflace v průběhu celého roku, v dubnu byl vyhlášen inflační cíl pro meziroční přírůstek indexu spotřebitelských cen v období leden 2002 až prosinec 2005, a to v podobě průběžného pásma, které rovnoměrně klesá z úrovně 3,0 - 5,0 % v lednu 2002 na úroveň 2,0 - 4,0 v prosinci 2005. Skutečná míra inflace dosáhla v roce 2002 hodnoty 1,8 % a v roce 2003 hodnoty 0,1 %.

⁹⁵ Vedle Rakušanů jsou jejími propagátory - byť v jiné podobě a z jiných teoretických východisek - příslušníci postkeynesovského proudu.

⁹⁶ Z hlediska vývoje českého ekonomického myšlení je zajímavé, že některé úvahy týkající se postavení centrální banky významného českého ekonoma J. Macka jsou podobné pojetí postavení centrální v učení některých stoupenců postkeynesovského proudu. (Macek, J.: *Sociální ekonomika*. Díl III. Nákladem České grafické unie, Praha 1946; resp. některé kritické stati k politice Národní banky československé a k názorům K. Engliše publikované v průběhu 30. let.)

zakončila slovy F. A. v. Hayeka: „But the Central Banks are by no means the only factor capable of bringing about the change in the volume of circulating media.“⁹⁷

Seznam literatury

- [1] Arlt, J.: Inflace v České republice v polovině devadesátých let (srovnání alternativních vysvětlení). *Politická ekonomie*, 1997, č. 5
- [2] Bulíř, A.: Volné bankovníctví: znovunalezené řešení? *Finance a úvěr*, 1992, č. 8
- [3] Bulíř, A. : Československá monetární politika po roce 1989. *FÚ*, 1993, č. 5
- [4] Bulíř, A.: Exogenita nabídky peněz: některé pochybnosti o předchozím výzkumu. *Finance a úvěr*, 1996, č. 1
- [5] Čihák, M.: Ohlédnutí za fluktuálním pásmem koruny. *FÚ* 1997/10
- [6] Čihák, M. – Holub, T.: Cílování inflace v ČR: staré víno v nových lahvích. *FÚ* 1998/4
- [7] Čihák, M. - Holub, T.: Cílování inflace: předběžné poznatky a zkušenosti z prvních měsíců. *HN* 14. 7. 1998, s. 6
- [8] Čihák, M. - Holub, T.: Co říká ekonomická teorie o nezávislosti centrální banky. *Finance a úvěr* 1999, č. 9
- [9] Čihák, M. - Holub, T.: Měnový výbor: cesta z dilemat centrálního bankovníctví v ČR? *Bankovníctví* 2000, č. 6
- [10] Čihák, M. – Janáček, K.: Inflace v České republice v polovině devadesátých let - odpověď. *Politická ekonomie* 1996, č. 1
- [11] Dědek, O.: Currency Convertibility and Exchange Rate Policie in the Czech Republic. *ČNB, IE*, 1995
- [12] Dyba, K.: Navrhovali jsme změnu kurzového režimu. In *Tři roky od měnové krize. Sborník textů*. CEP, Praha 2000

⁹⁷ Hayek, F. A.: *Monetary theory and trade cycle*. Jonathan Cape, London 1933, s. 148

- [13] Frait, J.: Autonomie monetární politiky a monetární přístup k platební bilanci (aplikace na ČR v letech 1992 – 1995). *Finance a úvěr* 1996, č. 5
- [14] Frait, J.: Je snížení inflace podmínkou pro obnovení ekonomického růstu. *Bankovníctví* 1998, č. 8
- [15] Frait, J. – Zedníček, R.: Časová nekonzistence a cíle měnové politiky. *Bankovníctví* 1996, č. 16
- [16] Frait, J. – Komárek, L. – Kulhánek, L.: Analýza dynamiky inflace v ČR pomocí P* modelu. *FÚ* 1998/11
- [17] Guba, M. – Stiller, V. – Arlt, J.: Vztah mezi peněžní zásobou a vývojem inflace v l. 1993 – 1996. *Finance a úvěr* 1997, č. 3 a 4
- [18] Hájek, M. a kol.: Makroekonomická analýza české ekonomiky 1996. *Politická ekonomie* 1997, č. 3
- [19] Havel, J.: *Ekonomika a právo v transformaci*. *Politická ekonomie*, 1998, č. 6
- [20] Hayek, F. A.: *Monetary Tudory and trade cycle*. Jonathan Cape, London 1933
- [21] Hayek, F. A.: *Právo, zákonodárství a svoboda*. Academia, Praha 1991
- [22] Holman, R.: *Transformace české ekonomiky. V komparaci s dalšími zeměmi střední Evropy*. CEP, Praha 2000
- [23] Holman, R.: *Měnová krize 1997 byla trest za nezavedení volného floatingu*. In *Tři roky od měnové krize. Sborník textů*. CEP, Praha 2000
- [24] Holub, T.: Analýza inflace v ČR. *Finance a úvěr* 1997, č. 12
- [25] Hrnčíř, M.: *Devizový kurz a transformace české ekonomiky*. Sborník vědecké konference k 70. výročí centrálního bankovníctví v ČR. ČNB, Praha 1996
- [26] Hrnčíř, M. – Šmídková, K.: ČNB určila dva horizonty cílování inflace. *HN* 5. 6. 1998
- [27] Hrnčíř, M. – Šmídková, K.: *Nová kotva měnové politiky*. *HN* 29. 4. 1998

- [28] Hrnčíř, M. – Šmídková, K.: ČNB určila dva časové horizonty cílování inflace. HN 5. 6. 1998
- [29] Izák, V.: Endogenous or exogenous money supply? – the Czech case. ČNB, IE, 1995
- [30] Izák, V.: Inflation, Wales and Money. Pratur Economic Pápera, 1996, č. 3
- [31] Janáček, K. a kol.: Česká ekonomika začátkem roku 1997. Hospodářské trendy 11. Komerční banka 1997
- [32] Janáčková, S.: Měnová politika - úspěchy i hledání nových cest. ČNB, IE, Praha 1995
- [33] Janáčková, S.: Dilemata současné měnové politiky. Finance a úvěr, 1996, č. 1
- [34] Jonáš, J. – Bulíř, A. – Kerouš, M.(1990): Makroekonomická politika v období privatizace – náčrt základních problémů (I). FÚ, 40, č. 12
- [35] Jonáš, J.: Měnová politika a měnový kurz. Finance a úvěr 1996, č. 1
- [36] Kindlová, E.: Od C. Mengera k F. A. Hayekovi. Habilitační práce. VŠE, Praha 2002
- [37] Klaus, V.: Ekonomická teorie a realita transformačních procesů. Management Press, Praha 1995
- [38] Klaus, V. (1998): Deficit, restrikce, dezinflace a stav ekonomiky. Vystoupení na VŠE v Praze dne 26. 11. 1998. Stručný výtah viz též HN 26. 11. 1998, s. 11
- [39] Klaus, V.: Tři roky po měnové krizi: rekapitulace událostí a jejich souvislostí, aby se na některé věci nezapomnělo. In Tři roky od měnové krize. Sborník textů. CEP, Praha 2000
- [40] Kočárník, I.: Postřehy jednoho z obětí beránek měnové krize z roku 1997. In Tři roky od měnové krize. Sborník textů. CEP, Praha 2000
- [41] Kodera, J. - Mandel, M.: Transmisní mechanismy měnové politiky v podmínkách české ekonomiky. ČNB, IE, Praha 1997

- [42] Kosmata, V.: Přechod na inflační cílení má svá rizika. HN 113. 1. 1998, s. 6
- [43] Kreidl, V.: Česká měnová politika v roce 1996 a 1997. Bankovníctví 1997, č. 3
- [44] Kysilka, P.: Vnější rovnováha jako odraz souladu mezi mírou otevřenosti ekonomiky a parametry makroekonomické politiky. in Otevřenost české ekonomiky a vnější rovnováha: konec 90. let. Sborník textů. CEP, Praha 2001
- [45] Kysilka, P.: Inflace je léčitelná nemoc. Bankovníctví 1998, č. 1
- [46] Mach, P.: Měnová krize z hlediska změn devizových rezerv. In Tři roky od měnové krize. Sborník textů. CEP, Praha 2000
- [47] Mandel, M.: Analýza měnové politiky a měnového vývoje v České republice v letech 1989 - 1995. Bankovníctví, 1997, č. 5
- [48] Mandel, M. – Tomšík, V.: Vývoj platební bilance ČR v období 190 – 1996 pohledem modelu IS-LM-BP. Politická ekonomie 1997, č. 4
- [49] Mandel, M.: Cílování inflace a peněžní zásoby při exogenních šocích (komparace z pohledu modelu AS-AD-NX. FÚ 1998/4
- [50] Martincová, O.: K terminologii. Finance a úvěr 1998, č. 4
- [51] Mataj, J. - Vojtíšek, P. (1992): Právní a ekonomické aspekty nového Zákona o SBČS. FÚ č. 3
- [52] Nešvera, V.: Dilema úrokové politiky. FÚ 1995/4 Postavení centrální banky. 14. seminář České společnosti ekonomické v řadě „Ekonomické teorie a česká ekonomika“. Praha 12. října 1999
- [53] Pospíšil, J.: Inflace a nezávislost centrální banky – máme a můžeme už mít nějaké zkušenosti i v České republice? Sborník vědecké konference k 70. výročí centrálního bankovníctví v České republice. ČNB, Praha 1996
- [54] Revenda, Z. - Mandel, M. - Kodera, J. - Musílek, P. - Dvořák, P. - Brada, J.: Peněžní ekonomie a bankovníctví. 3. vydání. Management Press, Praha 2000

- [55] Šmídková, K.: Účinnost měnové politiky a poptávka po penězích. *Finance a úvěr*, 1996, č. 8
- [56] Šmídková, K.: *Search for the Offset Coefficient*. ČNB, IE, Praha 1995
- [57] Šmídková, K. – Hrnčíř, M.: Přejít na strategii cílování inflace. *Finance a úvěr* 1998, č. 4
- [58] Šmídková, K. – Hrnčíř, M.: Inflační cílení v zahraničí – zdroj informací a inspirace pro českou měnovou politiku. *Bankovníctví* 1998, č. 7
- [59] Teja, F.: *Offset Coefficient for the Czech Republic*. MF ČR, Praha 1995
- [60] Tomšík, V.: Příliv zahraničního kapitálu, režim devizového kurzu, vyrovnávací procesy platební bilance a autonomie monetární politiky ČR v letech 1993 – 1996. *Studie NHÚ J. Hlávky* č. 6, Praha 1997
- [61] Tomšík, V.: Výzkum kompenzačního koeficientu v podmínkách ČR v letech 1993 – 1996. *Finance a úvěr* 1998, č. 6
- [62] Tomšík, V.: Možné problémy současné měnové politiky. *Bankovníctví* 1998, č. 13
- [63] Tošovský, J. (1998): Přednáška na Manažerském fóru v Praze, červen 1998
- [64] Tůma, Z.: Inflace v České republice a nezávislost centrální banky. Sborník vědecké konference k 70. výročí centrálního bankovníctví v České republice. ČNB, Praha 1996
- [65] Tůma, Z. – Kreidl, V.: Stará kotva opuštěna, kde hledat novou? *Ekonom* 1996, č. 7
- [66] Vencovský, F.: *Dějiny českého ekonomického myšlení do r. 1948*. Nadace Universitas Masarykiana. Edice Heureka. Brno 1997
- [67] Vencovský, F.: K nové měnové strategii ČNB. *Bankovníctví* 1998, č. 4
- [68] Zahradník, P.: Volné bankovníctví a slučitelnost s realitou. *Finance a úvěr* 1992, č. 8
- [69] Zeman chce mít ve vedení ČNB své lidi. *Lidové noviny* 26. 9. 1998

Další materiály

- [1] Zprávy ČNB o inflaci, vývoji platební bilance, výroční zprávy a internetové stránky.
- [2] Na mušce politiků. Euro 2. 11. 1998, s. 32
- [3] ČNB chce čistou inflaci kolem šesti procent. Mladá fronta Dnes. 23. 12. 1997, s. 14
- [4] ČNB je příliš nezávislá a nemá zpětnou vazbu. Lidové noviny 1. 10. 1998
- [5] Právní postavení centrální banky v demokratickém státě. Sborník textů č. 3. CEP, Praha 2000

