

économie appliquée

*an international journal of
economic analysis*

Tome XLVII - 1994 - N°1

Revue publiée avec le concours du C.N.R.S.



SOMMAIRE

Irma Adelman et Nobuhiko Fuwa	Income Inequality and Development – The 1970s and 1980s Compared	7
Augusto Graziani	Real Wages and the Loans-Deposits Controversy	31
Philippe Moati et El Mouhoub Mouhoud	Information et organisation de la pro- duction : vers une division cognitive du travail	47
Sylvie Bonny	La standardisation technologique en agri- culture : formes, origine et perspectives à partir du cas français	75
Jérôme Dumoulin	Innovations pharmaceutiques et régle- mentation : le paradigme de l'explication	101
Dimitri Mardas	Spécialisation et échanges intra-branche : reflet de l'absence de politique indus- trielle commune dans le cas des neuf Etats-membres de la CEE	127
Jean-Luc Bailly	Monnaie, « finance » et intérêt	151

Notes et Commentaires

Vladimir Kollontai	Disintegration of the USSR : Economic Implications	177
Amit Bhaduri	Chaotic Targeting on the Market Clearing Price	197

Income Inequality and Development The 1970s and 1980s Compared

Irma Adelman
Nobuhiko Fuwa

University of California at Berkeley

Is there a trade-off between inequality and economic growth? Using the most recent data and simultaneous equation methods we relate the share of income of poor and rich quintiles in LDCs to per capita GNP and some conditioning variables.

We find that in the early stages of development the income share of the poorest declines very steeply and then remains there. The share of the less poor groups reaches its minimum earlier and recovers sooner but slowly. Also the structural adjustment policies of the 1980s exacerbated the growth-equality trade-offs for the poor.

Y-a-t-il un «trade-off» entre inégalité et croissance économique? Utilisant les données les plus récentes et la méthode d'équations simultanées nous établissons une relation entre les quintiles de revenu des pauvres et des riches dans les pays les moins développés et quelques déterminants. Nous avons constaté que dans les premières étapes de développement la part de revenu des groupes les plus pauvres chute de manière très abrupte et demeure à un point bas. La part des groupes moins pauvres atteint son minimum plus tôt et se rétablit plus tôt, mais lentement. Par ailleurs les politiques d'ajustement structurel des années 80 ont exacerbé pour les pauvres les compromis croissance-égalité.

INTRODUCTION

The long-term relationship between income distribution and development has been one of the most closely investigated issues in development economics. In his path breaking article, "Economic Growth and Income Inequality", Kuznets (1955) formulated the hypothesis that early economic growth increases inequality while later economic development narrows it. He based this hypothesis on an analytic model; on data for developed countries since the 1930s, which showed a narrowing of inequality; and on cross-country comparisons between inequality in developed countries and inequality in two developing countries, which showed considerably greater inequality in the latter. This paper formulated the U-hypothesis and posed the research agenda for subsequent studies of the relationship between income distribution and development.

Nevertheless, during the first twenty years immediately following World War II the U-hypothesis was ignored by all but a few radical academic critics, such as Baran (1957) and Myrdal (1968), and by Indian policy makers. It was assumed that the growth of the modern sector, if sustained, would eventually spread the benefits of economic growth to all, including the poorest. In the late 1960s, as a result of the work of the International Labor Organization, mainstream Western economists realized that, despite rapid industrialization and GNP growth, urban unemployment had vastly increased. The combination of slow labor absorption in the modern sector, rapid population growth, education explosion, and exploitation of agriculture had transformed disguised rural unemployment into a combination of low-income, low-productivity disguised urban unemployment, in the informal urban sector, and open educated unemployment of recent, secondary and university, graduates. The realization of these trends revived interest in the Kuznets U-hypothesis. Distributional and poverty issues came to the center of the development agenda.

I. PREVIOUS STATISTICAL STUDIES

The early 1970s led to the initiation of research into the distribution of benefits of growth in developing countries. Data on income distribution in developing countries, however, was (and remains) scant. The first major study of the relationship between income distribution and

economic development was by Adelman and Morris (1973). Completed in 1971, as a report to the Agency for International Development, and based on unpublished income distribution studies in 44 developing countries, their study confirmed the increase in inequality at early stages of development posited in the Kuznets U. It also indicated that the subsequent decrease in inequality with development was not automatic. It depended on which specific policy choices were made in the course of the development process. With policies stressing the reduction of economic dualism and increases in primary and secondary education the later stages of development would reduce inequality; with a continuation of dualistic growth involving neglect of the agricultural sector and a narrow educational pyramid inequality would not decrease, even at the latest stages of development. The Adelman-Morris study did not use regression analysis to establish the relationship between income distribution and development. They argued (in Appendix C to their book) that the heterogeneity of the data and the state of ignorance about the appropriate functional form made the use of regression analysis a dangerous research tool. Instead they relied on a form of analysis of variance (the analysis of hierarchic interactions) which was robust to data quality and did not require prior specification of functional forms.

Following their analysis, a large number of investigators used cross section regressions to study the relationship of inequality to development (Paukert 1973; Bacha 1979; Chenery *et al* 1974; Ahluwalia 1976 a and b; Ahluwalia, Carter and Chenery 1979; Anand and Kanbur 1986; Papanek and Kyn 1986 and 1987; and Bourguignon and Morrisson 1990). These studies generally used a functional form that is quadratic in the log of per capita GNP. They also added some conditioning or policy variables to the regressions such as education, population, or a socialism dummy. The samples of countries varied, sometimes including and sometimes excluding developed and communist countries. Except for Bourguignon and Morrisson (1990), the regressions all confirmed the existence of the Kuznets curve. Anand and Kanbur (1986), however, argue that the location of the minimum point of the U is sensitive to sample composition and to the specific functional form. Such sensitivity is to be expected if, as claimed by Adelman and Morris (1973), the underlying relationship at higher levels of development is either U shaped or J shaped, depending on the policy choices made. Papanek and Kyn (1986) contradicted the Anand and Kanbur contention, and find the relationship to be stable and insensitive to the inclusion or exclusion of specific countries. The conditioning variables they include, however, capture the very policy choices that affect whether the relationship is U

or J shaped. They also find the Kuznets curve to be quite flat (1987). Bourguignon and Morrisson (1990) claim that the Kuznets variables are not statistically significant. But their regressions exclude large countries, such as Brazil, Bangladesh, China, India, Pakistan and Indonesia, where most of the world's poor reside, on debatable *a priori* grounds.

In the present paper we reinvestigate the issue of whether there is a trade-off between inequality and economic growth, using the most recent income distribution data for LDCs compiled by the World Bank for the 1970s and 1980s. We use simultaneous equations estimation techniques to relate the shares of income accruing to the poor and rich quintiles to per capita GNP and a set of conditioning variables.

II. THEORY GUIDING VARIABLE SELECTION

In this section, we focus on the factors that explain the Kuznets curve and the distribution of income in the long and short run and use them to guide our variable selection for the regression studies.

1) Labor absorption into the modern sector

Factor movement among sectors with different wage structures are, potentially, a major source of changes in inequality. Most Kuznets-U generating models rely on intersectoral transfers of population and income to provide the mechanisms which change income-inequality. These transfers were the mechanisms on which Kuznets himself relied. Lewis' 1957 model of development through industrialization also implies a Kuznets curve that is due to intersectoral transfers (Fields 1980). The early stages of industrialization increase the income gap between industry and agriculture and transfer investable funds and labor from the low-productivity, even-distribution, traditional sector to the high-productivity, unequal-distribution modern sector. Up to a point, this process generates an increase in inequality. The per capita income gap between sectors starts closing, in the Lewis model, when a rising wage rate is required to attract more labor into the modern sector. For this to occur, the growth rate of employment in the modern sector must, at least, exceed the growth rate of the supply of labor to the modern sector. Numerical simulations and individual-country studies suggest that, in a given country, the association of decreased inequality with further

growth will start only after more than half the labor force is employed in the modern sector (Fei and Ranis 1964; Robinson 1976; Ranis 1978).

The indices of intersectoral resource allocation used in our study are: the share of agricultural employment in total employment, the share of industrial value added in total GDP and the share of primary exports in total exports. Surprisingly, in view of their theoretical importance, none of these variables proved systematically significant in our regressions for either the 1970s or 1980s. No explanation for their lack of significance suggests itself. Specifically, the insignificant association is not due to a lack of structural change: throughout the period industrialization has continued and the expansion of manufactured exports from developing countries has accelerated.

2) Productivity differentials

The rise and fall in sectoral imbalances in total-factor-productivity change is another prime mover of changes in income inequality with development. Early development is characterized by increasing the imbalance in productivity among sectors, followed by a leveling in productivity differentials and finally by balanced productivity growth, quite late in the development process (Kuznets 1955 and Chenery-Syrquin 1975). Even without population movement from the low-productivity to the high-productivity sector, the U-shaped movement in the productivity gap would suffice to generate a Kuznets Curve.

Agricultural development is the primary method for narrowing productivity gaps among sectors. Many semi-industrial LDCs and several low-income LDCs turned to agricultural development in the 1970s, as a result of the 1973 food crisis and the need to reduce food imports. However, the 1980s were a decade of renewed neglect of agriculture in some countries and emphasis on dualistic agricultural development in others. Dualistic agricultural development is pursued through policies aimed at increasing agricultural output on commercial farms while neglecting the family-farm sector. It has the effect of decreasing the intersectoral productivity differentials while increasing inequality within the rural sector. A low productivity differential may either signify increased productivity on family farms, with a favorable effect on both rural and overall inequality, or it may signify increased productivity on large commercial farms, with unfavorable effects on inequality. This gives rise to a potentially ambiguous effect of diminished productivity increases on inequality. Nevertheless, our

indicator of productivity differentials among sectors – the ratio of productivity in agriculture to productivity in industry – proved uniformly significant in reducing inequality in the distribution of income.

3) Savings generation

More rapid growth requires higher investment. In a closed economy, if more rapid growth is to materialize, income must be shifted from low savers (the poor) to high savers (the rich). This requires decreasing the income share of the poor, especially in the early stages of development. This hypothesis was first advanced by Kalecki (1943) and taken up by Kaldor (1955) and the Latin American structuralist school. In an open economy, other sources of investment-finance are available to supplement domestic savings: foreign aid, foreign investment, and foreign borrowing. Governments can also mobilize domestic savings through inflation.

The variables introduced to represent the methods used to mobilize savings were the rate of domestic inflation, the ratio of outstanding foreign debt to GNP, and the ratio of the stock of foreign capital investment to GNP. In our study, these variables played an important role in explaining inequality during the 1980s, when adjustment policies to the debt crisis dominated development policy. In interpreting their impact on distribution, we must bear in mind that these variables stand not only for the specific methods used to mobilize savings but also for the policy measures to achieve eventual macroeconomic stabilization and adjustment which their use has entailed.

4) The distribution of wealth

One would expect land distribution, the distribution of physical capital and the distribution of human capital to affect the distribution of income. Indeed, many revolutions (1848, in France (land); 1917, in Russia (physical capital and later land); and the 1960s in France and the United States (human capital)) have been based on this theory. As Lindert and Williamson (1985) point out, however, major wealth redistributions such as land reform, slave emancipation, war and losses from economic crises, are too sporadic to offer a systematic explanation of the course of the Kuznets curve. They are likely to lead to important shifts in the Kuznets curve, however.

Both conservative and progressive economists agree that improving the educational attainments of the masses will make the distribution of income more equal. The "human capital" school (Schultz 1971 and Becker 1967) sees education as improving the earnings capacity of individuals. The "redistribution with growth" (Chenery *et al* 1974) and the "redistribution before growth" (Adelman 1978) schools see a broadening of the educational pyramid as a redistribution of wealth. Chenery *et al* argue for increasing the education of the poor as part of redirecting a larger share of investment towards increasing the assets of the poor. Adelman argues for increasing the education of the poor as part of equalizing the distribution of wealth of the major productive asset before improving its productivity, as a means of setting the stage for more egalitarian subsequent growth.

Our primary indicators of wealth distribution were a land concentration Gini coefficient; a communism dummy, signifying no inequality in the distribution of material property and inequality only in the distribution of human capital and access to privileged jobs; primary and secondary school enrollment ratios; and a Latin America dummy, which on the average, indicates a concentrated distribution of land and of real and financial capital.

5) Speed of change

Another explanation of changes in the distribution of income points to the contrast in initial conditions faced by the poor and rich when responding to the new economic opportunities inherent in economic growth. (Adelman and Morris 1973 and Morris and Adelman 1983). Fundamentally, all processes of economic change give rise to both increased absorption of some individuals and displacement and marginalization of others. Those who own (or have access) to factors used disproportionately in expanding sectors or in new technologies, or to factors that are complementary to them, are enriched by change. Those who own factors used in producing substitute commodities or less productive technologies lose from change. How more rapid growth affects different deciles depends on the net balance between the two forces for each decile. Those with assets – including not only financial capital and land but also human capital, information, and networks facilitating migration and access to high-productivity jobs – are better positioned to take immediate advantage of the opening up of any set of new economic opportunities. The poor are slower to respond to

increased opportunities because they have less assets, in both the narrow and the more general sense. Furthermore, since the poor use traditional technologies and combine them with small amounts of low-productivity complementary factors, they are more likely to be marginalized by new technologies, to which they require more time to adapt. Slower growth also allows more opportunity for social adaptation by the poor (through demographic change, migration, and schooling) and hence is likely to affect them less unfavorably ⁽¹⁾.

We did not find the speed of growth, as measured by the rate of growth of per capita GNP, systematically associated with income inequality in either period. Changes in income inequality may have been more affected by social adaptation by the poor than by domestic growth rates. Among methods of social adaptation used by the poor, international migration to the Gulf countries has greatly increased since the formation of OPEC, and the gaps in schooling between the poor and the rich have diminished markedly.

6) Politics and Institutions

The political ideology of the government affects the extent to which equity-oriented policies and strategies are implemented during the development process. It also influences the nature of economic and political institutions. Government investment patterns in infrastructure and education, immigration and trade policies, and agricultural terms of trade are all influenced by the predominant political complexion of the government. In governments dominated by traditional landed and/or foreign elites, roads linking the countryside are few and limited to connecting extractive centers with ports and commercial farming centers to cities; the educational pyramid is narrow; agricultural development patterns are dualistic, favoring large estates and commercial farms over small farms and tenancies; agricultural terms of trade policies reflect the interests of large farmers; and agricultural infrastructure investment patterns in transport and irrigation favor large farming interests. Conservative governments support liberal immigration policies and restrict unionization of the labour force to keep wages low; and

⁽¹⁾ This point is illustrated by the contrast in the extent to which economic growth marginalized the poor in France, Germany and Great Britain during the 19th century. During the early stages of the Industrial Revolution, poverty increased faster with growth in Great Britain and Germany, fast growing countries, than in France, a slow growing country.

they use overvalued exchange rates and tariffs to encourage import-substitution. Elite-oriented government also restrict government social programs and wealth and inheritance taxes.

The economic institutions favored by "liberal" governments are market-oriented and shun safety and equity-oriented regulation of corporations. And, as has been repeatedly demonstrated, exclusive reliance on markets, especially in the absence of safety nets, is good for economic growth and bad for equity. There are also missing institutions for the supply of credit, extension and marketing to small owner-operated farms and for the supply of investable funds to medium-size firms in manufacturing.

The variables we introduced to reflect the political orientation of government are: a communism dummy; a socialism dummy, which describes the extent of socialist influence on government policy in the previous decade, and ranges from none (e.g. Brazil) to significant (India); a variable describing the extent of government involvement in the economy – the share of government expenditure in GNP; and a variable describing the extent of redistributive government policies – the share of government social expenditures in total government expenditures. Of these, the communism variable was significant in the 1980s and the socialism variable in the 1970s.

7) Initial conditions

The initial conditions of a country affect its options for development strategies, and have an impact on its geo-political position. These factors, in turn, influence a country's distribution of income. The variables we introduced to reflect the country's initial conditions are: population, a broad measure of resource abundance, and population density.

A nation's size, as reflected by its population, affects its development strategy choices. Scale economies are more pronounced in large countries, making import-substitution a more appealing development strategy. International trade cannot play as large a role in the economies of large countries as it does in the economy of small countries. Large countries therefore tend to be more closed and less subject to external influences. Large countries have more varied resource bases, making for less need for trade and enabling greater diversification of production. Finally, large countries tend to have greater regional differences, and therefore more unequal distributions of income.

Natural resource abundance makes primary specialization more likely. It tends to bring in expatriate settlers, and encourage colonization.

(Natural-resource poor countries are not worth colonizing). Natural resource abundance is therefore likely to generate foreign dependence and to have governments that are dominated by primary-export interests. In a previous study, (Adelman and Morris 1973) greater natural resource abundance was associated with more unequal distribution of income. Our present analysis for the 1970s confirms that finding.

Greater population density implies less arable land per agricultural worker. It generates a predominance of small, uneconomic holdings, and fosters low-productivity, subsistence agriculture. (The East Asian economies are important exceptions to this generalization). The low productivity subsistence sector may coexist with a high-productivity commercial farming sector that provides staple exports and food for cities. Income distribution within the rural sector is therefore likely to be more unequal in land-scarce countries. Despite the existence of some commercial farming, the productivity gap between agriculture and industry is likely to be larger. We would therefore expect less arable land per agricultural worker to be associated with more unequal distribution of income. In our study for the 1970s, we find this expectation confirmed.

III. THE STATISTICAL ANALYSIS

All less developed countries for which income distribution data could be found were included in the study. The 1970 sample consists of 45 developing countries of which one is East European. The 1980 sample includes 38 developing countries of which 3 are East European. The country-data and sources are reproduced in Tables A1 and A2 of the Appendix. The dependent variables are the quintile shares. In addition to the Kuznets-curve variables, the log of per capita GNP and its square, the independent variables include: two variables, unit1 and unit2, which represent the types of measures to which the decile shares refer; the factor allocation variables; the productivity differential variable; the wealth distribution variables; the savings generation variables; the speed of change; the politics and institution variables; and the initial condition variables.

We then used stepwise regressions to select, from this extensive list of variables, a subset of statistically significant variables for inclusion in the simultaneous-equations estimation. The significant variables in the stepwise regressions were then incorporated into a set of seemingly unrelated regressions, that were estimated under the restriction that the

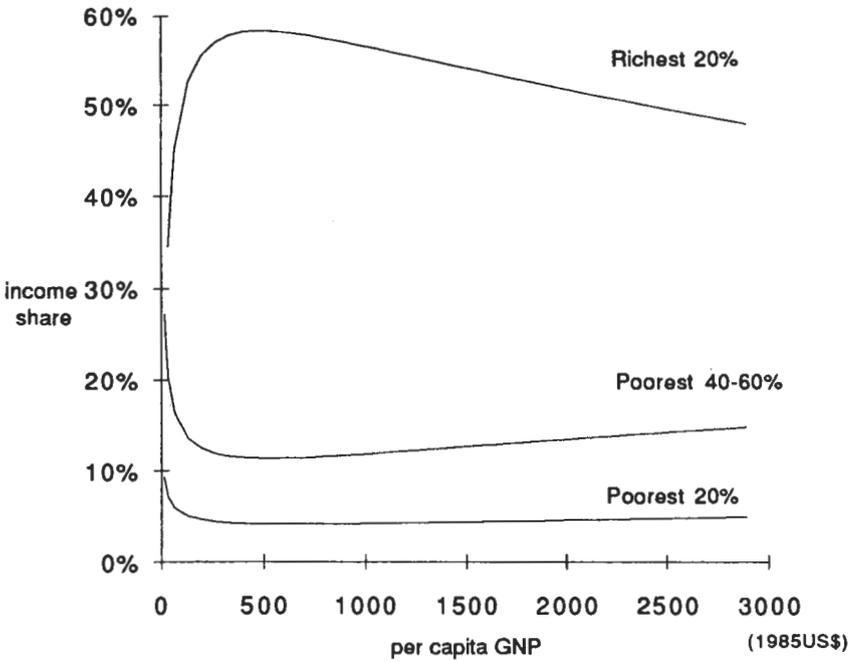


Figure 1
Estimated Kuznets Curve 1970s

estimated income share for all quintiles add up to unity. Tables 1 and 2 present the results of our simultaneous equation estimates for the 1970s and 1980s respectively. The R-squares for the individual regressions are high for both periods.

IV. THE KUZNETS U

The U-hypothesis is confirmed by our regression systems for both periods. The income share of the poorest is negatively related to the log of per capita GDP and positively related to its square for the four poorest quintiles; and it is positively related to the log of per capital GNP and negatively related to its square for the two richest deciles. For both the 1970s and 1980s F-tests for the combined hypothesis that the coefficients of both the log of per capita GNP and its square are equal to zero reject this hypothesis at a 10% significance level for all

Table 1
Seemingly Unrelated Regression Estimates For the 1970s

	Poorest 20%	Next 20%	Next 20%	Next 20%	Richest 20%
Log GNP/cap	-4.268 (-1.82)	-9.824 (-3.52)	-14.133 (-4.32)	-10.516 (-2.32)	38.743 (3.89)
Log GNP/cap squared	0.345 (1.94)	0.836 (3.93)	1.182 (4.719)	0.917 (2.62)	-3.275 (-4.29)
Unit1	-0.819 (-2.02)	-1.837 (-3.82)	-2.076 (-3.66)	-0.501 (-2.73)	4.295 (2.49)
Productivity Ratio	0.018 (1.86)	0.021 (1.91)	0.032 (2.46)	0.041 (2.36)	-0.112 (-2.84)
Land Gini	-0.035 (-3.42)	-0.046 (-3.78)	-0.040 (-2.84)	-0.034 (-4.82)	-0.156 (-3.68)
Socialism		-0.015 (-3.22)	-0.024 (-3.41)	-0.030 (-2.74)	0.068 (4.73)
Foreign Investment	-0.015 (-2.50)	-0.025 (-3.67)	-0.028 (-3.35)	-0.023 (-2.06)	0.092 (3.67)
Resource Abundance	-0.030 (-4.05)	-0.047 (-5.05)	-0.041 (-4.37)		0.118 (5.17)
Pop Density Agriculture		-0.005 (-2.21)	-0.001 (-2.12)		0.001 (2.32)
Constant	21.340 (2.85)	42.99 (4.80)	60.21 (5.71)	52.74 (4.38)	-77.28 (-2.41)
R-square	0.662	0.762	0.747	0.447	0.735

but the poorest LDC-quintile in the 1980s regressions. This contradicts the study of Bourguignon and Morrisson (1990), who, in an LDC sample, which refers to the 1970s and excludes large LDCs such as India, Bangladesh, Brazil, China, and Pakistan, find GNP and its square statistically insignificant.

Our results trace out a Kuznets Curve for both periods, showing that income inequality first increased with development and then declines. But, as we can see from Table 3 and Figure 2, especially for the 1980s, the U-curve looks more like a J-curve than like a U-curve in the income

Table 2
Seemingly Unrelated Regression Estimates For the 1980s

	Poorest 20%	Next 20%	Next 20%	Next 20%	Richest 20%
Log GNP/cap	-3.575 (-1.67)	-5.204 (-2.45)	-4.805 (-2.29)	-2.901 (-1.41)	16.486 (2.29)
Log GNP/cap squared	0.240 (1.48)	0.391 (2.44)	0.378 (2.38)	0.257 (1.67)	-1.266 (-2.26)
Unit1	-2.771 (-5.39)	-3.048 (-5.87)	-2.861 (-5.53)	-1.852 (-3.73)	10.472 (5.81)
Productivity Ratio	0.032 (3.13)	0.034 (3.86)	0.025 (4.20)		-0.091 (-3.71)
Communism	2.523 (3.28)	3.149 (4.08)	3.567 (4.67)	3.417 (4.69)	-12.626 (-4.73)
Foreign Investment	-0.031 (-4.96)	-0.072 (-4.34)	-0.062 (-3.76)	-0.029 (-1.86)	0.222 (3.88)
Inflation	-0.013 (-2.52)	-0.017 (-3.27)	-0.017 (-3.32)	-0.012 (-2.37)	0.060 (3.30)
Foreign Debt	-0.031 (-4.97)	-0.029 (-5.49)	0.020 (-5.39)		0.080 (5.31)
Constant	21.173 (2.99)	29.253 (4.18)	31.168 (4.50)	29.577 (4.38)	-11.139 (-0.46)
R-square	0.729	0.762	0.729	0.518	0.745

range relevant for developing countries confirming Adelman and Morris (1973), Papanek and Kyn (1987), and Anand and Kanbur (1986). First, the declines in the shares of the poorest deciles are steep and occur very early in the development process. The major declines occur at income levels corresponding to those of South Asia, where the bulk of the world's poor are. Second, the curves are flat around the turning points, especially for the 1980s. The income share of the poorest quintile is within 3% of the minimum at income levels ranging from \$485 to \$1000 for the 1970s and \$1689 to \$5000 for the 1980s. Third, due to the flatness of the U-curve, the share of the poorest does not recover

Table 3
The Turning Points of the Kuznets Curve
 (per capita GNP in 1985 US\$)

Population Group	1970*		1980	
	min/max GNP	Recovery GNP level	min/max GNP	Recovery GNP level
Poorest 20%	591	2 500	1 689	15 000
Next 20%	432	1 000	767	1 300
Next 20%	479	1 500	574	2 000
Next 20%	389	1 000	281	1 000
Richest 20%	451	900	671	1 300

* The data in the 1970 regressions was in 1980 US\$. For this table, the GNP figures estimated from the 1970 regressions were converted into 1985 US\$ using the US inflation rate.

to that corresponding to an income of \$100 until developing countries reach income levels corresponding to those of developed countries.

V. RESULTS FOR THE 1970s

Aside from per capita GNP, the significant variables in the regression systems for the 1970s represent a combination of structural and institutional variables and initial conditions. Macroeconomic variables play essentially no role in the regressions for the 1970s. Among the variables significant in our regressions, the greater the productivity differentials among sectors the larger the degree of inequality in the distribution of income, as expected on theoretical grounds. Again as expected, the larger the extent of inequality in the distribution of tangible wealth, the more unequal the distribution of income. The more concentrated the ownership of land, as indicated by the Gini coefficient for land, the lower the share of the poorest four quintiles and the higher that of the richest. But, surprisingly, the distribution of human wealth, as indicated by school enrollment ratios, does not appear to have exerted a systematic effect on income inequality, perhaps because the enrollment ratios represent flows rather than stocks and, during the seventies, countries with low stocks have increased school enrollment

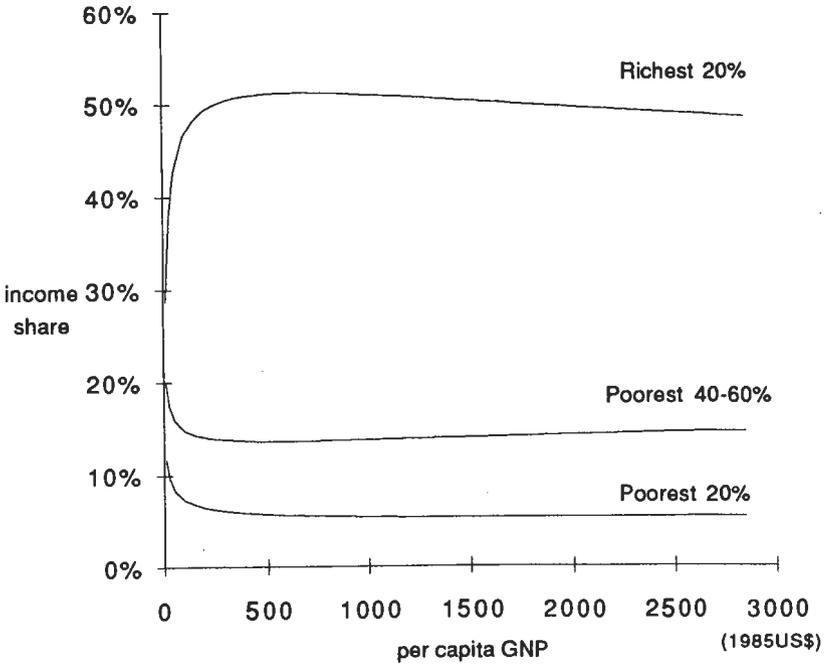


Figure 2
Estimated Kuznets Curve 1980s

ratios very significantly. Among initial conditions for development, the greater the degree of natural resource abundance and the higher the population density the lower the income share of the poorest and the higher the share of the richest quintile. As indicated earlier, these initial conditions are important not only in themselves but also because they lead to different development strategies and different political orientations of governments. Finally, the politico-institutional environment, as represented by our capitalism variable, indicates that the more market-oriented the domestic institutions in commodity and factor markets, the greater the extent of inequality in the distribution of income. Also, the greater the role of foreign elites in the economy and policy, as represented by the stock of foreign direct investment, the lower the share of the poorest four quintiles and the higher that of the richest.

VI. RESULTS FOR THE 1980s ⁽²⁾

The results of the stepwise regressions indicate that in the 1980s distributional consequences were largely dominated by macroeconomic adjustment policies and by the initial asset and political power distribution represented by communism, and foreign investment. Of the full list of variables, aside from the log of GNP and its square and sectoral productivity differentials, only inflation, foreign debt, net direct foreign investment, and the communism dummy were consistently significant across all equations. Most of the structural variables found significant in the 1970 appear, in the 1980s, to have been replaced by variables portraying adjustment policies. This contrast is perhaps the most interesting aspect of our results.

As expected, more pronounced intersectoral productivity differentials, which indicate greater intersectoral dualism, increase inequality. The elasticity of the share of the poorest 20% with respect to the productivity differential variable is 0.2 at the mean share. These results thus suggest that policies aimed at closing the productivity gap between sectors, by increasing the productivity of agriculture, can, as suggested in Adelman (1984), play a significant role in poverty alleviation.

On the average, communism increases the share of the poorest income groups by about 3 percentage points. This estimate implies that, in the LDC sample, the share of the poorest 20% is, on the average 30% higher under communism; that of the next quintile is 20% higher; that of the two middle quintiles is 15% higher. By contrast, in an average LDC, communism decreases the share of the richest 10% by 36%. Communism clearly made the distribution of income significantly more equal probably because it reflects a very egalitarian distribution of assets, no capital incomes, and a fairly narrow distribution of wages. But the greater apparent equality of the distribution of income in communist countries is probably substantially overstated by the data: the income distribution data for Eastern European countries does not include the distribution of public consumption (wages in kind by enterprises – cars, educational privileges, housing, and health coverage – which are more unequally distributed than wages); the measured income inequality

⁽²⁾ An earlier study by the authors, of income distribution and development in the 1980s, was based on data in the 1990 World Development Report and included 21 developed and 20 developing countries. The statistical results were very similar in nature to those reported below and will be published in a special issue of the *Indian Economic Review* in memory of S. Chakravarty.

also does not reflect the non-market rents accruing from status-related differential access to goods and services; and the distribution data refers only to the economically active population, excluding such groups as pensioners. It should also be noted, that in the 1980s the milder forms of socialism (represented by our capitalism variable), which tinker with pricing and investment policies at the margin without touching the basic distribution of wealth that were significant in the 1970s, do not appear to have significantly affected the shares of income of the poor during the 1980s. This may be because, during the 1980s, even socialist governments were forced to reduce or eliminate their subsidies to the poor and curtail their social expenditures.

Greater net foreign investment, our third significant conditioning variable, is consistently associated in our results with larger inequality, as in the 1970s. There are many possible interpretations for the negative association between net foreign investment and the share of income accruing to the poor. First, greater net foreign investment may be associated with a smaller share of income accruing to the poor because it represents more substantial foreign dependence, a tighter political alliance between domestic and foreign interests, and a more export-interest dominated economy and society; such a constellation of socio-political forces tends to lead to policies and investment patterns that discriminate against the poor (Baran 1957 and Morris and Adelman 1988). Second, increased foreign investment may increase inequality by augmenting the technological dualism and capital intensity of the economy through greater transfer of technology appropriate to the developed countries in which the investment originates. Third, greater foreign investment may increase inequality by flowing to natural-resource intensive activities, such as raw material extraction, and to commercial export crops produced on large estates and plantations; the returns from both of these types of activities are very concentrated. Finally, more direct foreign investment may be associated with greater inequality since it is likely to flow to countries following IMF adjustment policies more closely, and to countries that are more free-market oriented in their policies and institutions. The first three reasons for the association of greater inequality with greater foreign investment operated during the 1970s as well; only the last reason is specific to the 1980s.

The fourth systematically significant conditioning variable for the Kuznets curve in our seemingly unrelated regression systems is inflation. Our analysis indicates that, during the 1980s, inflation has systematically reduced the shares of nominal incomes of the poor while increasing the

shares of the rich. On the average, for LDCs, one percentage point increase in inflation reduces the share of income of the poorest by between .01 and .02 percentage points while increasing the share of the richest 10% by .06. At the average rate of inflation for the LDC sample, inflation reduces the share of the poorest by about one sixth. These results are probably due to the nature of IMF-inspired adjustment policies to inflation in the 1980s: wage repression, – to achieve international competitiveness, reduce domestic absorption, generate expenditure switching among domestic and foreign goods and decrease inflationary pressures – and cuts in governments subsidies to wage goods, to reduce government deficits. These policies have tended to exacerbate the usual differences in how assets owned by the poor and rich fare during inflationary times. Even without conscious wage repression, there is usually a lag between increases in wage incomes and inflation while the returns to the assets of the rich (land, foreign exchange, and physical capital) keep pace with inflation. Thus the rich are more successful in defending their incomes and assets from erosion by inflation than the poor and are not as hard hit relatively by adjustment policies.

A large literature on the incidence of adjustment to the debt crisis points to the fact that with IMF adjustment policies the poor have borne the brunt of the cost of adjustment (for a summary see Pinstrup-Anderson 1986, Taylor 1988 and The Human Development Report 1990). Our regression results confirm these findings and suggest the need to protect the vulnerable quintiles from the wage repression policies and sever cuts in government expenditures on social services used to maintain debt service.

CONCLUSIONS

Comparing our results for the 1970s with the 1980s, we find that there has been a structural difference in the nature of systematic influences on poverty between the debt-led growth period of the 1970s and the adjustment-oriented period of the 1980s. The Kuznets factors, productivity differentials among sectors, and the political complexion of the government were important in both periods. During the 1970s, however, long term factors, such as the wealth distribution and the extent of natural resource abundance exerted a systematic impact on inequality. By contrast, during the 1980s, short-term factors portraying

the nature of adjustment policies to internal and external imbalances were the significant additional influences on income distribution.

Our results indicate that, for the poorest of the poor, the Kuznets U is a Kuznets J: in the early stages of development, the income share of the poorest declines very steeply, reaches its minimum, and, for all practical purposes remains there. Recovery toward pre-development shares is very slow, and is not achieved until well after countries become fully developed. The share of less poor groups reaches its minimums earlier and recovers relatively sooner. For all groups, the Kuznets curve is very flat around the minimum. Our results further suggest that, during the 1980s, the structural adjustment policies and the investment financing methods used to support growth have exacerbated the growth-equality trade-offs for the poor.

Thus, the prospects for the bulk of the poor in developing countries can hardly be considered satisfactory. Our results indicate that, in the income range where the overwhelming majority of the poor are, the deterioration in the share of income of the poor with growth was quite substantial. The income range at which the major deterioration in income share of the poor occurs comprises South Asia and is characteristic of much of sub-Saharan Africa. Our findings thus support the view that, in the lowest income countries, the primary hope of the poor lies not in raising their countries' growth rates, but rather in changing the structure of assets of the poor and the nature of macroeconomic adjustment policies.

The flatness of the Kuznets curve gives rise to both optimistic and pessimistic interpretations. On the optimistic side, it suggests that the trade-off between growth and inequality is confined to the early and middle stages of development. Once countries reach the semi-industrial state, the decline or rise in shares is quantitatively so small that there is no effective trade-off between growth and inequality. Therefore, when countries reach a semi-industrial state, growth can unambiguously be used to reduce poverty. On the pessimistic side, the flatness of the curve suggests that relative deprivation will continue for a very long time, and that it will be a very long time indeed before a typical developing country can attain the degree of equality characteristic of very primitive developing countries.

REFERENCES

- Adelman I., *Redistribution before growth – a strategy for developing countries*, Martinus Nijhof, The Hague, 1978.
- Adelman I. and Morris C.T., *Society politics and economic development: a quantitative approach*, Johns Hopkins Press, Baltimore, 1967.
- Adelman I. and Morris C.T., *Economic growth and social equity in developing countries*, Stanford University Press, Stanford, 1973.
- Ahluwalia M.S., "Inequality, poverty and development", *Journal of Development Economics*, vol. 6, 1976a, p. 307-342.
- Ahluwalia M.S., "Income distribution and development: some stylized facts", *American Economic Review*, vol. 66, 1976b, p. 128-135.
- Ahluwalia M.S., Carter N.G. and Chenery H.B., "Growth and poverty in developing countries", *Journal of Development Economics*, vol. 6, 1979, p. 299-341.
- Anand S. and Kanbur S.M.R., *Inequality and development: a critique*, 25th Anniversary Symposium, Yale Growth Center, 1986.
- Bacha E., "The Kuznets curve and beyond: growth and change in inequalities", in: E. Malinvaud (ed.): *Economic growth and resources*, St. Martin's Press, New York, 1979, p. 52-73.
- Baran P., *The political economy of growth*, Monthly Review Press, New York, 1957.
- Becker G.S., *Human capital and the personal distribution of income: an analytical approach*, Institute of Public Administration, Ann Arbor, Michigan, 1967.
- Bourguignon F. and Morrisson C., "Income distribution, development and foreign trade: a cross-sectional analysis", *European Economic Review*, vol. 34, 1990, p. 1113-1132.
- Chenery H.S., Ahluwalia M.S., Bell C.G. et al., *Redistribution with growth*, Oxford University Press, Oxford, 1974.
- Chenery H.S. and Syrquin M., *Patterns of Development 1950-1970*, Oxford University Press, London, 1975.
- Fei J.C. and Ranis G., *Development of the surplus labor economy*. Irwin, Homewood, Illinois, 1964.
- Fields G.S., *Poverty, inequality and development*, Cambridge University Press, New York, 1980.
- Food and Agricultural Organization of the United Nations. *FAO Yearbook: production*. FAO, Rome, (various issues).
- Food and Agricultural Organization of the United Nations. *1970 world census of agriculture: analysis and international comparison of the results*. FAO, Rome, 1981.
- Gastil R.D., *Freedom in the world*, Freedom House, New York, (various issues).
- International Monetary Fund, *International financial statistics yearbook*, IMF, Washington, D.C., (various issues).
- International Monetary Found, *Government finance statistics yearbook*, IMF, Washington, D.C. (various issues).
- Kalecki M., *Studies in economic dynamics*, Allen and Unwin, London, 1943.
- Kaldor N., "Alternative theories of distribution", *Review of Economic Studies*, vol. 23, 1955, p. 83-100.

- Kuznets S., "Economic growth and income inequality", *American Economic Review*, vol. 45, 1955, p. 1-28.
- Lecaillon J., Paukert F., Morrisson C. and Germidis D., *Income distribution and economic development: an analytical survey*. International Labour Office, Geneva, 1984.
- Lewis W.A., "Economic development with unlimited supplies of labor", *Manchester School of Economic and Social Studies*, vol. 22, 1954, p. 139-91.
- Lindert P.H. and Williamson J.G., "Growth, equality and history", *Explorations in Economic History*, vol. 22, 1985, p. 341-377.
- Morris C.T. and Adelman I., "Institutional influences on poverty", *Journal of Economic History*, vol. 43, 1983, p. 43-55.
- Morris C.T. and Adelman I., *Comparative patterns of economic development, 1850-1914*. Johns Hopkins University Press, Baltimore, Maryland, 1988.
- Myrdal G., *The Asian Drama*, Pantheon Press, New York, 1968.
- Organization for Economic Cooperation and Development-Development Assistance Directorate. *Stock of private direct investments by D.A.C. countries in developing countries, end 1967, 1971, and 1973*. OECD, Paris, 1973.
- Paukert F., "Income distribution at different levels of development: a survey of evidence", *International Labor Review*, vol. 108, 1973, p. 97-125.
- Papanek G.S. and Kyn O., "The effect of income distribution on development, the growth rate, and economic strategy", *Journal of Development Economics*, vol. 23, 1986, p. 55-66.
- Papanek G.S. and Kyn O., "Flattening the Kuznets curve: the consequences for income distribution of development strategy, government intervention, income and the rate of growth". *The Pakistan Development Review*, vol. 26, 1987, p. 1-54.
- Pinstrup-Anderson P., *Macroeconomic adjustment policies and human nutrition: available evidence and research needs*, International Food Research Institute, Washington, D.C., 1986.
- Ranis G., "Equity and growth in Taiwan: how 'special' is the 'special case'?", *World Development*, vol. 6, 1978, p. 397-409.
- Robinson S., "A note on the U-hypothesis relating income inequality and economic development", *American Economic Review*, vol. 66, 1976, p. 437-440.
- Schultz T.W., *Investment in human capital: the role of education and research*, Free Press, New York, 1971.
- Taylor L., *Varieties of stabilization experience*, WIDER studies in development, Clarendon, Oxford, 1988.
- United Nations Development Programme, *Human development report*, Oxford Press, New York, 1990.
- Van Ginneken W. and Park J. (eds), *Generating internationally comparable income distribution estimates*. International Labour Office, Geneva, 1984.
- The World almanac and book of facts*. World Almanac, New York, 1991.
- World Bank, *World Tables*, Johns Hopkins Press, Baltimore, Maryland, 1989-90.
- World Bank, *World development report*, Oxford Press, New York, 1993.

Real Wages and the Loans-Deposits Controversy

Augusto Graziani

Faculty of Economics
University of Rome

The classical model assumes that real wages are paid in advance, namely before the production process in started. In the neoclassical model, according to which the real wage rate equals the marginal product of labour, wages have to be paid after the production process is over. In the classical model, savings are a prior stock of wealth available at the beginning of each period. The monetary counterpart of the construction is that bank deposits equivalent to real savings exist at the beginning of each period. In the classical model, therefore, a sort of priority of deposits over loans can be detected. Neoclassical authors insist on the priority of deposits over loans; but in their model, since real wages are not paid in advance and no prior savings fund exists, a similar statement cannot be accepted.

Le modèle classique suppose que les salaires réels sont payés à l'avance, précisément avant que le processus de production ait démarré. Dans le modèle néo-classique suivant lequel le niveau du salaire réel est égal au produit marginal du travail, les salaires doivent être payés une fois le processus de production achevé. Dans le modèle classique, les épargnes sont un stock préalable de richesse disponible au début de chaque période. La contrepartie monétaire de la construction classique est que les dépôts bancaires équivalant à l'épargne réelle existent au début de chaque période. Dans le modèle classique, par conséquent, une sorte d'antériorité des dépôts sur les prêts peut être décelée. Les auteurs néo-classiques insistent sur l'antériorité des dépôts par rapport

aux prêts; mais, dans leur modèle, comme les salaires réels ne sont pas payés à l'avance et qu'il n'existe pas d'épargne préalable, une telle affirmation ne peut être acceptée.

FOREWORD

The problem we shall examine in the present paper concerns the old and never ending controversy concerning the relationships between saving and investment on one side, bank loans and deposits on the other. Paradoxical as it may seem, in order to fully understand the nature of the problem and to evaluate the solutions given to it over time, it will prove useful to have recourse to a debate belonging to an apparently widely distant field of theoretical economics, namely the *theory of wages*. The specific point we shall examine is whether, under a macroeconomic viewpoint, real wages are to be understood as being paid in advance, that is before the production process is started, or only after it has been completed. As we shall show, the answer given to this question will throw light on the apparently totally different problem indicated above.

It goes without saying that if we consider the timing of wage payment in relation to the working period, no question arises, since it is commonly agreed that, as a rule, wages are paid *at the end of the working period* (end of the week, end of the month). The problem we shall investigate is instead whether wages are paid at the beginning or *at the end of the production period*; or, to put it differently, the question is whether workers, when consuming real wages, are consuming commodities which they have themselves produced in the same period, or commodities produced beforehand, possibly by others, and subsequently saved.

In the history of economic doctrines, opposite views have been held on this point. Classical economists agreed on the fact that real wages are paid in advance. In contrast to the position of the classics, neo-classical economists were thinking in terms of posticipated wages, the main reason being that if the real wage-rate has to be equal to the marginal product of labour, wages can only be paid after the production process is over and the product is available. If real wages were to be paid before production were started, the available real product (except in long period equilibrium positions) might simply be insufficient to cover the wage bill.

Let us now turn to the macroeconomic problems we intend to analyse. They can be summarized in the two following questions, both concerning two points which seem still open to doubt:

a) Is it legitimate to identify bank deposits with money savings? A superficial consideration of the relevant literature would suggest that the general answer is in the negative; still, frequent hints, to be found also in the writings of most authoritative scholars, seem to reveal a preference for a positive answer;

b) Is it legitimate to say that, for the banking system as a whole, loans make deposits? Again, while a general agreement on this view seems to prevail in the literature, it often happens to read totally opposite statements.

The main conclusions of the present paper are the following ones:

1) If we think in terms of the classical model, in which wages are paid in advance, the following propositions hold: a) it is correct to conceive of savings as a prior stock of wealth existing at the beginning of each period; b) it is also correct to consider investment of any period as determined by the savings inherited from the preceding period; c) and finally a correspondence exists between bank deposits and savings on the one hand, and between investment and bank loans on the other;

2) If we think in terms of the neoclassical model, in which real wages are paid at the end of the production process, none of the above statements is correct.

I. THE CLASSICAL MODEL

The Wage-Fund. Classical economists agreed on the fact that real wages are paid in advance. As Adam Smith writes:

“It seldom happens that the person who tills the ground has wherewithal to maintain himself till he reaps the harvest. His maintenance is generally advanced to him from the stock of a master, the farmer who employs him...” (Smith, 1776, Book I, chap. 8, p. 83).

“What is annually saved is as regularly consumed as what is annually spent, and nearly in the same time too; but it is consumed by a different set of people” (Smith, 1776, Book II, chap. 3, pp. 327-8).

Similar statements are to be found in Ricardo's works. Wages, in Ricardo's writings, are paid out of capital, that is by means of:

“that part of the wealth of a country which is employed in production, and consists of food, clothing, tools, raw materials, machinery, etc. necessary to give effect to labour” (Ricardo, 1817 (1951), chap. V, p. 95).

The assumption of advanced wages went hand in hand with the idea of a production process tied to a yearly harvest and to the related idea that the workers would have to wait for a whole year if they were to be paid by means of the wheat they have themselves produced. Consequently, in the classical model, both wage earners and capitalists are supposed to consume in each period commodities produced in the preceding one. As Morishima writes:

“At the time of Ricardo, agriculture was the dominant sector, and the production lag significant. Corn produced in one year was retained and used as wage funds for the following year” (Morishima, 1989, p. 12).

Beside such technical considerations, it also seemed proper to explain that wages are paid out of the capitalists' own savings. If this is true, what the poor eat is the result of the abstinence of the rich. It need not be reminded that Marx placed himself on a strongly opposite position, by his general statement that wage-earners are in any case granting credit to capitalists:

“Überall kreditiert der Arbeiter dem Kapitalisten” (Marx 1872 [1975], vol. I, chap. IV, Sec. 3, “Kauf und Verkauf der Arbeitskraft”).

The first of the two arguments for considering real wages as paid in advance, the one related to the technical aspects of agricultural production, has often been considered not only as the main one but also as vital to the whole theory. Consequently some authors think that the classical view would lose weight if applied in a different technical context, such as the one prevailing in modern industrial production where production takes place in a chain of synchronised cycles each one starting before the preceding one has come to an end. Morishima, on emphasizing this aspect, points out that the classical view is obsolete nowadays, just because of the declining importance of agriculture in the modern world (Morishima, 1989 p. 12, and 1990). The same position is also held by Blaug, who surprisingly seems to think that synchronised production, typical of modern industry, makes the advance payment of wages irrelevant (Blaug, 1968, Chap. 6; we shall suggest later on that the root of a similar idea can perhaps be found in Marshall).

However, as Wicksell had already made totally clear (Wicksell, 1926 [1934], Part II, chap. 2, Sec. E), the synchronization of industrial production is irrelevant. It may be true that a stationary economy, in which production takes place on the basis of a chain of synchronized cycles, apparently does not require any previous saving. In a similar economy, in any single moment there will be one production cycle which is completed and supplies the amount of commodities needed to pay the wage bill required for starting the following cycle; it may seem

therefore that, if production is run by means of synchronized cycles, wages can be paid in advance even in the absence of accumulated saving. Even if this is true, it still remains to be explained how the prevailing structure of production was initially set up or how it can expand. If production has to be increased and new lines of production, in addition to the already existing ones, have to be set up, and if the wage bill has to be advanced in real terms, the fact that the pre-existing processes are run in synchronized cycles is of no help. The conclusion cannot be avoided that the fact of moving from a yearly agricultural production to a synchronized industrial production does not reduce the requirements consequent to the advance payment of real wages. The assumption of an agricultural yearly output, typical of classical authors, simplifies the model without altering its features. (Those who think that synchronized production makes the advance payment of wages easier resemble those who think that the existence of a revolving fund of liquidity makes bank credit no longer necessary).

Let us go back to the classical model in which wages are paid in advance. At the beginning of each period, the output of the previous period is in the hands of capitalists, who must allocate it between direct consumption and advanced wages (if the additional assumption were made that capitalists do not consume, commodities produced in the previous period would go totally into the wage bill). At the end of the following production cycle, the harvest is used to restore the initial the stocks of consumption goods. If a profit has been earned, the amount of commodities passed on to the following period is higher than the one inherited from the past. The amount of commodities allocated to wages is therefore equal to the amount produced in the previous period less the consumption of capitalists, if any. This amount is the celebrated *wages fund* which constitutes the wage bill of the current period.

It is thus clear that the wage bill of each year *comes out of the savings of capitalists* (Senior, 1836 [1951], p. 143, 153, 198; Mill, 1848 Book II, chap. XI). It is also clear that, in line with the teachings of the Wage-Fund Doctrine, the wage rate is determined by the ratio between the part of total output that capitalists destinate to the wage bill and the number of employed workers. This does not mean that the wage rate is rigidly determined. Capitalists can allocate available output at their discretion either to personal consumption or to increasing the wage bill. This point, which according to some (Forget 1992) was accepted by Mill, is explicitly made by Taussig, who emphasizes the fact that, while the wage bill of any period is always "the product of past labour" (Taussig 1986a, p. 90), still:

“[wages] depend on the sums which the employers judge it expedient to turn to the hire of labor” (Taussig 1896a, p. 89; also Taussig 1896b, Part I, chap. IV).

Therefore, even in the short period, nothing prevents wage-earners from getting higher wages, even in the presence of a constant level of employment. If, in the long run, real wages are set at the subsistence level, this is not due to the wages fund doctrine, but to the theory of population adopted by the classics.

Contrasting definitions of Saving and Investment. The definition of investment adopted by the Classics follows from the above description of the process of production. In the Classical model, in contrast to what will be typical of the Neoclassical school, investment is not defined as equal to the production of new capital goods, but as the amount of resources engaged in current production. Therefore for capitalists as a whole, *investment is equal to the wage bill*. Since, as already said, investment is equal to the fraction of total output that capitalists decide not to consume, investment is also equal to saving.

A first set of conclusions can now be formulated:

- a) in the classical model, the wage-fund doctrine seems to be fully correct;
- b) in the classical model, the conception of saving as a stock of goods, which is available in real terms and can be allocated to consumption or to investment, is also correct;
- c) in the classical model, it is correct to maintain that savings sets a limit to investment.

It should be emphasized that each of the above conclusions is only valid within the classical model, in which real wages are paid in advance.

When we leave the classical model and move to the neoclassical model in which real wages are paid only after the production process has been completed, savings, as the fund of consumption goods needed for paying wages in advance, disappear. Many continue to consider saving as being in the nature of free or disposable capital existing prior to investment; but, as Hayek had to admit, these are only “vague and undefined terms” (Hayek 1941, pp. 9-10; it should be noticed however that Hayek himself had been guilty of the same fault ten years before; Hayek 1935, pp. 9 and 10).

It may seem that the above conclusions are totally irrelevant for the problems in the theory of money and banking mentioned at the outset. We shall see instead that such conclusions pertaining to the real sector are the source of other statements belonging to the theory of banking, which, while being perfectly valid within the classical model, have been

unduly transferred to the neoclassical framework, where they have been the source of totally wrong deductions.

II. THE NEOCLASSICAL MODEL

The neoclassical authors reject the idea of wages being paid in advance. Mainly on analytical grounds, they move to a definition of wages as being paid at the end of the production process.

This deviation from the classical model is mainly connected to the marginal theory of distribution. Let us assume that all the necessary conditions which are needed for the marginal rule of distribution to hold, including Euler's theorem, are satisfied. In equilibrium, firms pay each factor according to its own value marginal productivity. For this to be feasible, factors must be paid from the output that they have themselves produced. This means that labour, along with all other resources, can only be paid after the production process has been completed. If the factors of production are paid in advance, the output from which they get their remuneration would be the result of production performed in a prior period. In this case, no sure correspondence would exist between total available output and the total of factor remunerations computed according to the marginal principle.

This would be true also in a strictly theoretical general equilibrium model, in which *tâtonnement* rigorously precedes exchanges. Once *tâtonnement* has been performed and equilibrium prices have been determined, the marginal productivities of each resource are also determined. But, while this makes it possible to negotiate the remuneration of each factor in advance (this is in fact the meaning of *tâtonnement*), it does not allow actually to pay it. For this to be done, production must have already been completed, which means that real wages, just as any other kind of remuneration, cannot be paid in advance.

The neoclassical critique of the Wage-Fund doctrine. The Neo-Classical school refuses the Wage-Fund doctrine on the ground that such doctrine establishes no definite relationship between the real wage rate and the productivity of labour. It is a fact that, in general, the real wage rate cannot satisfy at the same time both the wage-fund principle and the marginal productivity rule. A similar correspondence would only be possible in a stationary economy with a constant population, in which a wage-fund, exactly equal in amount to the wage-bill computed

on the basis of the marginal principle, had been previously accumulated; or in a growing economy, with a constant rate of growth of population and a capital stock growing at the same rate such as to produce a constant labour productivity.

Wicksell's model, to be discussed below, is the only model in which the advance payment of real wages is consistent with the marginal productivity principle. In general however, the conclusion is that the marginal theory of distribution, being inconsistent with the wages fund theory, is also inconsistent with the advance payment of real wages.

Jevons and Marshall. Both Jevons and Marshall saw the contradiction between the marginal principle and the advance payment of wages. Both tried to overcome it by a somewhat loose treatment of the problem.

Jevons starts by defining the wage rate as being strictly connected to current product:

“... the wages of a working man are ultimately coincident with what he produces, after deduction of rent, taxes, and the interest of capital” (Jevons 1871 [1957], p. 270).

At the same time, Jevons supports the idea of an advance payment of wages:

“It is the proper function of capitalists to sustain labour before the result is accomplished,...” (Jevons 1871 [1957], p. 271).

Jevons tries to keep together the two contrasting views by saying that, when wages are paid in advance, the wage rate is set on the basis of a reasonable forecast as to the results of production. In stationary conditions, forecasts are by definition correct.

Marshall's treatment of the problem is equally instructive. For him, as for Jevons, the crucial point lies in the new theory of distribution, according to which:

“Wages tend to equal the net product of labour; its marginal productivity rules the demand price for it” (Marshall 1922, Book VI, chap. II, paragraph 3).

Adherence to this principle induced Marshall to reject the old wage-fund doctrine, a doctrine which:

“suggested a correlation between the *stock* of capital and the *flow* of wages, instead of the true correlation between the *flow* of the products of labour aided by capital and the flow of wages” (Marshall 1922, Book VI, chap. II, paragraph 10).

In adherence to the marginal principle, Marshall admits that, if wages have to be set equal to the marginal productivity of labour, they must also be paid out the *current* product, and therefore after the production process is over. No difference is made between labour and capital: each

of them anticipates its own contribution to production and draws its income at the end of the process. Marshall writes:

“[the] national dividend is at once the aggregate net product of, and the sole source of payment for, all the agents of production” (Marshall 1922, Book VI, chap. II, paragraph 6).

“Capital in general and labour in general co-operate in the production of the national dividend, and draw from it their earnings in the measure of their respective (marginal) efficiencies” (Marshall 1922, Book VI, chap. II, paragraph 10).

Still Marshall does not want wholly to reject the principle of the advance payment of wages. In fact he writes:

“... when anyone works for hire, his wages are, as a rule, advanced to him out of his employer’s capital - advanced, that is, without waiting till the things which he is engaged in making are ready for use” (Marshall 1922, app. J, paragraph 2).

In an attempt to reconcile the marginal principle with the advance payment of wages, he adds that even if workers, being paid in advance, do not get a share of current product, still they get “the equivalent” of it:

“the share of the national dividend, which any particular industrial class receives during the year, consists either of things that were made during the year, or of the equivalents of those things. For many of the things made... during the year are likely to remain in the possession of capitalists or undertakers... while in return they... hand over to the working classes some things that had been made in previous years” (Marshall 1922, Book VI, chap. II, paragraph 10).

Such reconciliation has clearly no analytical basis. In an attempt to make his analysis more precise, Marshall seems to make recourse to the principle of synchronized production (Marshall Book VI, chap. II, paragraph 10). Marshall’s treatment might therefore be the origin of Blaug’s and Morishima’s belief that synchronized production might reconcile the marginal principle with the advance payment of wages. In Marshall’s text, the point is however left somewhat on the air.

The consequences of a posticipated wage. Dropping the principle of advanced wages produces far-reaching consequences; all of them contribute to widen the gap between the classical and neo-classical schools.

If wages are not paid in advance, production no longer requires, as Marshall would say, “a good store of wealth set by” (Marshall 1922, App. J, paragraph 1). The clear-cut distinction between the capitalist class, which saves and invests, and wage earners who live on the verge of subsistence, disappears. The same category of investment, which for the classical authors was identified with the wage bill advanced by

capitalists, no longer exists, and the term investment is used instead to indicate the share of total product that the community as a whole decides to produce in the form of capital goods.

Once the concept of investment, in the classical meaning of the term, is dropped, also saving, as a previously available fund, disappears. In the new analytical setting, the function of saving is no longer of making possible the advance payment of wages, but of granting the equilibrium between supply and demand in the goods market in a world in which capital goods are produced along with consumption goods.

In a setting in which savings and investment are no longer identical by definition, a mechanism has to be found which explains how saving and investment are brought to equality. Such a mechanism necessarily involves the working of the monetary and banking system. In the investigation of this question, the analytical problems mentioned at the outset (relationships between bank loans and investment, or between bank deposits and savings) come to the foreground.

III. MONETARY ANALYSIS

In the extension of the analysis from the real to the monetary sector, the dominant contribution is due to Wicksell.

Wicksell's aim is of defining a monetary economy which, in the equilibrium position, can reconcile the apparently unattainable requirement of a marginal productivity wage which is also paid in advance. Wicksell starts by building a model for many aspects similar to the classical model. He imagines a society made up of three main classes, landowners, capitalists, and wage-earners. In his model, just as in the classical one, the wage is paid in advance. In his own words:

“In most phases of production, however, there is a longer or shorter interval between the employment of labour and the final production of an article for sale. Since the labourer does not usually wait for his wages for the whole of this period..., it must be evident that he does not obtain them *from the product of his labour...*” (Wicksell 1926 [1934], Part II, chap. II, sec. E, p. 189-90);
 “... [the wage fund theory] is correct only on the assumption that the labourers take their wages in kind at the same time as they render their services. If,...., they wish... to wait for their wages until their own product is ready for the market, then wages may... rise to any height whatsoever and be independent of the size of the wage-fund or capital;... by such a procedure, the workers would themselves become capitalists” (*ibidem*, p. 195).

Wicksell, along with the classical authors, in order allow for the advance payment of real wages, assumes that the necessary amount of

wage-goods is available at the beginning of each period; such amount comes from the output of the previous period and has been set aside by savers who, in his model, are at the same time capitalists and dealers. Under this viewpoint Wicksell fully accepts the wage-fund doctrine.

So far as the monetary mechanism is concerned, Wicksell starts by rejecting the assumption of a given money stock. Wicksell imagines that initially no money, neither fallen from heaven nor inherited from the past, is available the economy. Money is created inside the market mechanism by means of bank credit; the money stock appears therefore as a strictly endogeneous magnitude. Bank credit makes possible the so-called "prior availability of money" which is necessary to start the mechanism of exchanges; in particular, money is created by the banks in order to finance the expenditure of the firms and the payment of the wage-bill.

When production is started, firms are granted a bank loan and use it for purchasing the labour force and paying the wage bill. Wage-earners spend their money wages on consumption goods sold by dealers. Equilibrium between the monetary demand and the real supply on the goods market determines the money price of goods. At this point, money initially created by the banks is in the hands of the dealers, who can hand over their liquid balances to the banks by opening bank deposits on which a yearly interest is reckoned.

Wicksell's equilibrium can only be reached if the banks set the money rate of interest at a level which is equal to the real yield of investment. Some typical features of Wicksell's model deserve being pointed out.

As already explained, the money price of goods is determined at the beginning of each period. It depends on the amount of available real saving and on the amount of bank credit granted by the banks and has no direct relation with the price level of the preceding periods. The price level therefore need not be stable over time. On the other hand, even if, in any period, the level of money prices is higher than the one prevailing in the past, this can by no means be identified with the beginning of an inflationary process: a new and higher price level may prevail, without being in itself the source of any further increase in prices. Inflation in fact does not originate from an increase in the quantity of money or in the amount of bank loans granted, but from a divergence between the natural and the money rates of interest. More precisely, the origins of inflation lie in the money rate of interest being too low as compared to the real yield of investment.

In equilibrium, capital is paid according the natural rate of interest, which is a measure of the real yield, and therefore of the

marginal productivity, of investment. In a position of general economic equilibrium, if investment is paid according to its marginal productivity and Euler's theorem holds, labour will also be paid according to its own marginal productivity. As a consequence, real wages, in spite of being paid in advance, will be set according to the marginal rule.

Let us now go back to the questions initially mentioned, namely the correspondence between savings and bank deposits on one side, and the parallel correspondence between investment and bank loans on the other. Some features typical of the classical model are also present in Wicksell's model:

a) wage-earners can convert their money wages into real commodities as soon as they get them, which means that *wages are paid in advance in real terms*;

b) *saving precedes investment*; if, with the classics, by real investment we mean the wage bill, real investment is determined by real saving;

c) bank loans are used by firms for paying the wage bill, that in the terminology of the classical model corresponds to the money value of investment. We can say therefore that *bank loans correspond to the investments of the firms*;

d) real savings of the capitalists, namely the wage goods set aside by savers, are sold to wage-earners and subsequently converted into bank deposits: we can therefore say that bank deposits are the result of savings.

The conclusion emerges that the double correspondence between bank loans and investment on one side, and bank deposits and savings on the other, is only valid when the theory of a monetary economy is constructed starting on the basis of the classical model with real wages paid in advance. At the same time, Wicksell's model shows beyond any possible doubt that the recurring idea that deposits make loans is totally wrong. In Wicksell's model, at the beginning of each period, savings exist as an amount of real goods and not in the form of monetary savings or bank deposits. Bank deposits come into existence as a consequence of loans initially granted by the banks. It is clear therefore that loans make deposits and deposits cannot exist if not as a consequence of prior loans.

IV. HAYEK'S MACROECONOMICS

Hayek's macroeconomics was the source of a celebrated debate between Hayek and Keynes and has been since then the object of

severe criticism (Graziani 1992). In contrast to the dominant view, we shall argue that Hayek's analysis can be considered as internally consistent, provided it is interpreted in the framework of the classical model with advanced real wages. In what follows, we shall refer to the opinions expressed by Hayek in his famous essays of 1931, written immediately after the publication of Keynes' *Treatise*, and to his *Prices and Production* (Hayek, 1931, 1935).

The main problems raised by Hayek's analysis lie in his monetary analysis, in that Hayek firmly maintains that a basic requirement of macroeconomic equilibrium is *that loans granted by the banks should not exceed savings collected through bank deposits*:

"... it is perfectly clear that, in order that the supply and demand for real capital should be equalised, the banks must not lend more or less than has been deposited with them as savings (and such additional amounts as may have been saved and hoarded). And that means naturally that... they must never allow the effective amount of money in circulation to change" (Hayek 1935, p. 27).

A similar statement is obviously unacceptable in its literal formulation, since loans make deposits and not the other way round. Still, as we shall try to show, it is possible to find the logical origins of Hayek's mistake and partly to justify his position.

Let us go back to Wicksell's model. As we know, Wicksell unifies in the same social group savers, capitalists, and dealers. Wicksell himself declares that this simplification was adopted for the sake of being clear ("der größeren Übersichtlichkeit wegen", Wicksell 1898, chap. IX, Sec. B, p. 127). If we separate the two groups of capitalists-savers and dealers, we can imagine the process to work as follows.

At the end of each period, capitalists, instead of keeping their saving in kind (for instance, as a stock of corn), sell the corn to dealers who buy it by means of loans granted them by the banks. Savers, once converted their real saving into money, pay the money into a bank deposit. At the end of the period, the banks' accounts will show loans granted to dealers equal in amount to deposits collected from savers. At the beginning of the following period, firm will get from the banks the usual loans which are needed in order to pay wages. Wage-earners, as usual, will buy consumption goods from the dealers, who will now be able to extinguish their debt towards the banks. At this point, we are back to Wicksell's simplified model: the banks' accounts will show loans granted to the firms (equal in amount to the wage bill) and an equal amount of deposits owned by wage-earners.

Clearly, the fact of dealing with dealers and capitalists as separate social groups does not alter the working of the real part of the model. But

the equilibrium conditions *can now be restated in monetary terms*. In the simplified model, initial savings are represented by a physical amount of consumption goods owned by capitalists-dealers; in the disaggregated model, while obviously dealers own the same amount of goods, savers, having sold the same goods to dealers, have converted their real saving into bank deposits and are now in possession of monetary wealth. Available saving can therefore be defined either as a stock of physical goods in the hands of dealers, or as a stock of liquid resources in the hands of savers. As a consequence the savings-investment equality can now be stated in two different forms:

a) we can say that loans granted to firms (representing investment of the current period) should be equal to the value of the stock of goods in the hands of dealers;

b) or we can say that loans granted to the firms should be equal to deposits that the banks have collected from savers.

Let us now imagine that the banks follow Hayek's prescription and grant new loans equal in amount to deposits collected from savers. Firms, as we know, pass on the liquidity to wage-earners who in turn pass it on to dealers who pay back their debt to the banks. The net result of the chain of operations is that, any time a unit of credit is granted to the firms, a unit of debt is extinguished by the dealers. Just as Hayek recommends, if banks follow his prescription, the money stock (that in Wickseil's model is pure credit in the form of bank deposits) is constant.

Hayek's analysis contains however a clear mistake. In Hayek's presentation, it seems that the banks can get deposits from savers *without any previous loan having been granted*. Our reconstruction shows clearly that the savings which give rise to bank deposits originate from loans that the banks themselves have granted to dealers; and that when the banks grant loans to the firms, they obviously are not making use of deposits collected from savers (something that they clearly would be unable to do), but they are granting new loans. Such loans, the amount of which is totally independent from the previously existing deposits, generate in turn an exactly equal amount of deposits.

V. CONCLUDING REMARKS

The object of the present article has been of investigating three main ideas still dominating in the common opinion, namely: a) that at the beginning of each period, an initial stock of free saving is available; b)

that bank deposits originate from the flow of savings; c) that the role of bank loans is to finance investment. We have shown that the above propositions hold only in the framework of the classical model with advanced real wages. We have also shown that, as soon as we move to the neo-classical model, none of the above propositions holds. The neoclassical model, in which the assumption of advanced real wages is dropped, does not tolerate real savings pre-existing to investment or bank deposits pre-existing to loans. The surviving idea of a presumed correspondence between savings and bank deposits, even if analytically wrong, can be justified as a residual of the classical model with wages paid in advance. The equally surviving, and equally wrong, idea of a presumed priority of deposits over loans lacks any historical or analytical justification.

REFERENCES

- Blaug M., *Economic Theory in Retrospect*, Homewood, Ill, R. Irwin, 1968.
- Forget E., "J.S. Mill and the Tory Movement. The Theoretical Value of the Recantation", *History of Political Economy*, vol. XXIV, 1992, p. 31-59.
- Graziani A., "A Note on Hayek's Macroeconomic Equilibrium", unpublished manuscript, July 1992.
- Hayek F.A., "Reflections on the Pure Theory of Money of Mr. J.M. Keynes", *Economica*, XI, 1931, p. 270-295 et 398-403; XII, 1932, p. 22-44.
- Hayek F.A., *Prices and Production*, London, Routledge & Kegan Paul (1st ed. 1931), 1935.
- Hayek F.A., *The Pure Theory of Capital*, London, Routledge & Kegan Paul, 1941.
- Jevons H.S., *The Theory of Political Economy*, New York, Kelley & Millman Inc, 1871 [1957].
- Marshall A., *Principles of Political Economy*, Londra, MacMillan, 1922.
- Marx K., *Das Kapital*, Frankfurt, Ullstein Verlag, 1872 [1975].
- Mill J. St., *Principles of Political Economy*, ed. by W.J. Ashley, New York, A.M. Kelley, 1848 [1964].
- Morishima M., *Ricardo's Economics*, Cambridge University Press, 1989.
- Morishima M., "Economic Theory and Industrial Evolution", in M. Baranzini and R. Scazzieri eds., *The Economic Theory of Structure and Change*, Cambridge University Press, 1990.
- Ricardo D., *Principles of Political Economy and Taxation*, edited by P. Sraffa, Cambridge University Press, 1817 [1951].
- Senior J.N., *An Outline of the Science of Political Economy*, London, Allen & Unwin, 1836 [1951].
- Smith A., *An Inquiry into the Nature and Causes of the Wealth of Nations*, The Glasgow Edition of the Works and Correspondence of Adam Smith, Oxford University Press, 1770 [1976].

Taussig F., "The Employer's Place in Distribution", *Quarterly Journal of Economics*, vol. X, 1896, pp. 67-94.

Taussig F., *Wages and Capital*, Londra, 1896.

Wicksell K., *Geldzins und Güterpreise*, Leipzig, Jena. G. Fischer, 1898 [1989].

Wicksell K., *Lectures on Political Economy*, London, Routledge & Sons, 1926 [1934].

Information et organisation de la production : vers une division cognitive du travail *

Philippe Moati

Crédoc, Cessefi-Paris I

El Mouhoub Mouhoud

Université d'Evry
Cessefi-Paris I

La diffusion dans les années 80 de nouveaux modes d'organisation de la production a constitué une réponse des entreprises à l'évolution de leur environnement économique : changement dans les normes de consommation, modification de la division internationale du travail, accélération et complexification du progrès technique, le tout dans un contexte de turbulences monétaires et financières. L'idée avancée est que ces nouveaux modes d'organisation sont la manifestation d'une transformation plus fondamentale du principe de division du travail qui glisserait d'une logique technique à une logique de compétences et d'apprentissage. Nous examinons comment le nouveau principe de division « cognitive » du travail se distingue du principe de division technique du point de vue du critère de segmentation des processus

* Une version antérieure de cet article a fait l'objet d'une présentation au cours du séminaire Metic du CEPREMAP du 22 septembre 1992 et de la réunion du Cessefi de l'université Paris I, le 12 octobre 1992. Nous remercions les participants à ces séminaires, ainsi qu'un rapporteur anonyme, pour leurs remarques qui ont permis une amélioration du texte initial. Nous remercions également le Prédit et France Télécom pour le financement apporté à la réalisation de l'étude dont est issu cet article.

de production, des modes de coordination et de réintégration, et des déterminants de la localisation des activités productives.

The implementation of new production organization patterns in the 1980's results from the firms' response to the evolution of their economic environment : change in consumption standards, change in international division of labour, speeding-up and increasing complexity of technological change, in a context of financial and monetary turbulence. Our idea is those new production organization forms are the reflection of a more fundamental change in the principle of division of labour which is supposed to pass from a technical nature to a logic of skills and learning. The new principle of «cognitive division of labour» can be distinguished from the technical principle at the levels of the criterion of the production process segmentation, the mechanisms of coordination and reintegration, and the factors of location of economic activity.

INTRODUCTION

La grande majorité des économistes étudiant l'entreprise et l'organisation industrielle s'accordent à reconnaître que d'importantes modifications des modes d'organisation de la production ont commencé à apparaître dans le courant des années 70 et se manifestent aujourd'hui de plus en plus clairement⁽¹⁾. Cette évolution comporte de multiples dimensions et a donné lieu à l'observation de «faits stylisés» : remplacement dans les ateliers des chaînes de production par des groupes semi-autonomes s'accompagnant d'une certaine recomposition des processus productifs; développement des formes «partenariales» de relations industrielles (sous-traitance, alliances, accords de coopération); recentrage des entreprises sur leurs «métiers»; développement des dépenses de R&D et, plus généralement, de l'investissement incorporel; accroissement du poids des PME mais renforcement de l'emprise des groupes; diffusion des pratiques de «juste à temps» (JAT) et de production en flux tendus; changements dans les déterminants et les modalités de la délocalisation industrielle et recomposition régionale des processus productifs,...

(1) Durand (1993), Dockès (1993), Jacot (1990), Commissariat général du plan (1990), Insee-L'entreprise (1992),...

Un important effort de conceptualisation a été accompli au cours des dernières années pour tenter de rendre compte, plus ou moins localement, de ces différents faits stylisés. La théorie de la régulation (Boyer (1992), Coriat (1990), Du Tertre (1989),...) et l'approche évolutionniste (Freeman et Soete (1990), Freeman et Perez (1988), Dosi (1992),...) s'efforcent de fournir un cadre global rendant compte des changements dans les modes d'organisation de la production. Ces changements sont souvent ramenés à une opposition entre rigidité et flexibilité des conditions de l'offre. Cet article avance une hypothèse d'interprétation de ces changements comme une manifestation d'un glissement, plus fondamental, du principe de la division du travail.

Les dernières décennies ont été le témoin de changements structurels dans l'environnement des entreprises, dont on peut brièvement rappeler les principales manifestations.

– *L'évolution des normes de consommation.* La saturation des besoins primaires dans les pays industrialisés a entraîné un ralentissement de la croissance des marchés de biens de consommation s'accompagnant de nouveaux comportements des consommateurs, caractérisés par une plus grande versatilité. Les modalités de la concurrence reposent davantage sur la différenciation des produits et sur la capacité des entreprises à s'adapter à la volatilité des marchés.

– *L'évolution de la division internationale du travail.* La mondialisation des marchés et l'évolution de la division internationale du travail (montée du Japon et des nouveaux pays industriels) ont provoqué l'exacerbation de la concurrence sur des marchés déjà moins dynamiques et ont renforcé les incitations à mener des stratégies de différenciation pour les firmes des pays industrialisés. La concurrence devient hétérogène; les avantages comparatifs en compétition sont fondés sur une palette élargie de déterminants (minimisation des coûts, qualité et renouvellement des produits par l'innovation,...). Les marchés deviennent à la fois plus concurrentiels et le caractère polymorphe des régimes de concurrence s'accroît.

– *Le changement technologique.* Les dernières décennies ont été marquées par l'accélération du progrès scientifique et technique, contribuant à une complexification et à un éclatement des savoirs. Le développement et la diffusion des technologies de l'information et de la communication (TIC) modifient en profondeur la manière de produire. Par ailleurs, l'exigence de différenciation et de flexibilité liée au fonctionnement contemporain des marchés incite les entreprises à incorporer les dernières découvertes technologiques dans les produits et les procédés. En outre, la science et les technologies se complexifient

en se développant. Ce mouvement conduit à la formation de nouvelles disciplines et à la constitution de corps de spécialistes devant maîtriser des compétences de plus en plus approfondies et de plus en plus pointues. Comme le mentionne Witkowski, le progrès des connaissances mène « à une recherche hyperspécialisée, dont le moindre domaine est si vaste qu'aucun spécialiste ne saurait en avoir une vue d'ensemble » (Witkowski, 1991, p. 8). Cette hyperspécialisation provoque un certain cloisonnement. « C'est un peu comme une galaxie en expansion, dont le volume gonfle mais dont les parties s'écartent les unes des autres » (Calame, 1991). Pourtant, en même temps que les savoirs s'écartent et se cloisonnent, on relève un accroissement de l'importance des aspects combinatoires dans le développement technologique et l'innovation-produit (Chesnaï, 1990). Les entreprises sont donc soumises à un paradoxe : l'activité économique réclame de plus en plus la combinaison de technologies et, au-delà, de corps de savoirs, alors que ces technologies sont de plus en plus sophistiquées et qu'il devient de plus en plus difficile d'en maîtriser plusieurs simultanément.

– *Les turbulences financières.* Les turbulences financières, qui ont marqué les années 80⁽²⁾, ont contribué au climat d'instabilité et d'incertitude qui a caractérisé l'univers de décision des entreprises. Par exemple, l'instabilité des taux de change a créé une forte incertitude sur le coût des approvisionnements en matières premières importées; simultanément, en modifiant très brutalement les niveaux de compétitivité-prix, elle a altéré les bases mêmes de la concurrence internationale⁽³⁾. Ces facteurs monétaires ont sans doute favorisé la recherche par les entreprises d'une compétitivité hors prix. Les turbulences financières se sont aussi accompagnées d'une montée de la logique financière dans le fonctionnement de l'économie. Les entreprises ont dû faire preuve d'une plus grande sélectivité dans leurs projets d'investissement productif, et ce d'autant plus que l'incertitude généralisée (au niveau de la demande, de l'évolution technologique,...) rendait cet investissement plus risqué. Les stratégies de diversification se sont révélées plus mesurées et les entreprises ont eu tendance à privilégier des formes d'engagement peu coûteuses en capital (développement de la sous-traitance, des alliances stratégiques, ...). Enfin, le coût des actifs circulants est apparu d'autant plus important

(2) Instabilité des parités monétaires, crise de l'endettement international, tension sur l'épargne mondiale, gonflement de la bulle spéculative entretenu par le décroissement des marchés nationaux et l'accélération des transactions, prolifération des nouveaux instruments financiers, vagues d'OPA,...

(3) Lafay et Herzog (1989, p. 20).

que les taux d'intérêt étaient élevés. Ainsi, l'optimisation de la gestion des stocks a pu apparaître comme un moyen efficace d'allègement du compte financier et a constitué une incitation forte à la mise en place d'organisation de la production en flux tendus.

L'ensemble de ces évolutions a contribué à augmenter le degré de complexité et d'instabilité de l'économie qui, aujourd'hui, nécessite et génère des flux croissants d'informations. Ce renforcement du rôle de l'information dans la vie économique en a accru sa valeur potentielle. Il convient cependant d'éviter une confusion : l'information ne possède pas une valeur intrinsèque. Au contraire, il est même clair que les TIC, le développement des médias,..., ont réduit considérablement le coût d'accès à l'information. Le caractère stratégique de l'information relève de la connaissance et du savoir. Si, comme l'affirme Toffler (1990), «le savoir devient la ressource décisive de l'économie avancée» (p. 119), le véritable pouvoir réside dans la maîtrise de la transformation de l'information en connaissance. Cette transformation cruciale réclame la compétence ⁽⁴⁾.

Si l'information peut souvent être considérée comme un bien libre, il est loin d'en aller de même des compétences. Les compétences intègrent une composante tacite qui leur donne la propriété d'être imparfaitement communicables (Nelson et Winter (1982), Dosi (1988),...). Par ailleurs, le développement des compétences est le plus souvent local et obéit à une logique cumulative. Ceci signifie que lorsque des compétences progressent, elles tendent à se spécialiser. A mesure qu'elles se spécialisent, les compétences deviennent de plus en plus difficiles à communiquer et à acquérir. D'une manière qui peut paraître paradoxale, en démultipliant les capacités de transmission et de traitement de l'information, les TIC ont accéléré ce processus de spécialisation des compétences et de morcellement du savoir. La mise en avant des compétences, ainsi que les caractéristiques de leur processus de développement, ne seraient-elles pas en train de modifier la nature de l'activité de l'entreprise et les conditions de sa prospérité?

Traditionnellement, l'entreprise est un lieu de combinaison de *facteurs* de production plus ou moins génériques. Le critère de sélection sur le marché est la capacité des entreprises à optimiser le rendement des facteurs en sélectionnant la meilleure combinaison (la technique la plus efficiente). C'est en gros le principe à la base de l'entreprise «fordienne» qui émerge au début du siècle et qui est fondée sur la production en série (Piore et Sabel, 1984).

⁽⁴⁾ Voir Pelikan (1989).

Avec la complexification de l'environnement économique et l'accroissement de son instabilité, la viabilité de l'entreprise passe de plus en plus clairement par la «*capacité à découvrir la bonne adaptation entre des technologies particulières et des marchés particuliers*» (Reich, 1991, p. 84). Le cœur de l'activité de l'entreprise consiste alors à combiner des compétences spécialisées. L'obsolescence rapide, à laquelle l'accélération du changement soumet les compétences, impose à l'entreprise de mettre en œuvre les conditions favorables à l'adaptation rapide des compétences par un processus de transformation de l'information en nouvelles connaissances pertinentes en regard des sollicitations du marché.

L'écart entre les deux modèles est important. L'«*entreprise-facteurs*» privilégie l'efficacité physique (maximisation du rendement) et statique (logique d'affectation (Perrin, 1990)); l'«*entreprise-compétences*» recherche l'efficacité qualitative (innovation et qualité) et dynamique (créativité et flexibilité). L'entreprise-facteurs utilise principalement des facteurs de production génériques, c'est-à-dire qu'elle procède à l'affectation de ressources «*données*» (Gaffard, 1990). L'entreprise-compétences, par l'apprentissage collectif qu'elle cherche à favoriser, développe la spécificité de ses facteurs de production. La compétitivité de l'entreprise-compétences réside dans la qualité de ses facteurs de production de laquelle elle cherche à tirer «*l'intégration permanente de l'innovation et de la production*» (Gordon, 1990, p. 322), alors que la compétitivité de l'entreprise-facteurs découle de l'efficacité de la combinaison de facteurs génériques retenue.

A travers cette modification de la nature de l'activité de la firme, c'est le principe même à la base de la division du travail qui est renouvelé. **Le principe de la division du travail serait ainsi en train de glisser d'une logique technique et de rendement à une logique de compétence et d'apprentissage. On passerait ainsi d'une division technique à une division cognitive du travail.**

La manière dont est divisé le travail dans un système économique comporte trois aspects : le critère de segmentation de la production; la manière dont est coordonnée l'activité des différentes parties prenantes et dont sont réintégrés les fragments de processus de production; la localisation de la production de chaque produit ou fragment. Examinons ce qui différencie la division cognitive de la division technique du travail sur chacun de ces trois aspects.

I. CRITÈRE DE SEGMENTATION DE LA PRODUCTION

Dans la division technique du travail, le processus de production est segmenté dans l'optique de l'optimisation du rendement. A. Smith identifiait trois raisons aux gains de productivité associés à la division du travail de détail : l'ouvrier améliore sa dextérité lorsqu'il ne l'exerce que sur un nombre limité d'opérations; il perd moins de temps à passer d'une opération à une autre; sa spécialisation le rend plus à même de penser à des innovations sur la manière d'effectuer sa tâche de façon plus efficace. Pour bénéficier des avantages de la division technique du travail, il convient donc d'identifier les morceaux de processus de production homogène du point de vue de la nature des opérations à effectuer. La segmentation est bien de nature technique et dépend des propriétés physiques singulières à chaque processus de production. Cette logique pousse à parvenir à des tâches suffisamment élémentaires et indépendantes pour qu'elles puissent être accomplies de manière mécanique, d'abord par un ouvrier spécialisé, puis par une machine qui aura été spécialement conçue pour optimiser l'accomplissement de cette tâche⁽⁵⁾. Cette logique de spécialisation technique tend à séparer les tâches de conception et d'exécution et trouve sa forme la plus significative dans la chaîne de production fordienne. La division technique du travail présente le caractère remarquable d'amener une spécialisation toujours plus poussée des facteurs de production, qui les rend de plus en plus attachés à une tâche donnée. Du côté du capital, l'hyperspécialisation des équipements est source de rigidité et est créatrice de coûts irrécupérables. Le travail, en dépit de son hyperspécialisation, reste un facteur générique car la spécialisation se traduit ici par un appauvrissement conceptuel du contenu du travail, rendant très court le temps d'assimilation du savoir-faire nécessaire pour passer d'un emploi particulier à un autre⁽⁶⁾. En d'autres termes, la main-d'œuvre est spécialisée, non polyvalente mais interchangeable.

La logique de cette forme de division du travail revient à morceler toujours davantage le processus de production. Cependant, comme l'avait montré A. Smith, la capacité à mettre en œuvre une division très fine du processus de production dépend de l'étendue du marché. Cette relation entre l'étendue du marché et le degré de division du travail

⁽⁵⁾ Dockès et Rosier (1988).

⁽⁶⁾ Voir, par exemple, Coriat (1991), Freyssenet et Thénard (1988).

doit être perçue de manière dynamique. En effet, la division du travail favorise la réalisation de gains de productivité et le développement de nouvelles machines. Le prix relatif des marchandises ainsi produites a tendance à baisser, ce qui accroît leur marché. Cette extension du marché autorise une division plus fine du processus de production et permet ainsi la réalisation de nouveaux gains de productivité... (Boyer et Schmeder, 1990). Cette logique s'épuise avec la difficulté croissante qu'il y a à trouver de nouveaux débouchés lorsque les consommateurs sont de plus en plus éloignés géographiquement, faisant perdre en coût de transport une part croissante des gains de productivité, et lorsque l'inclination des consommateurs à consommer le produit est satisfaite et ne peut être relancée par une nouvelle baisse de prix. C'est précisément en cela que la saturation des besoins primaires dans les pays industrialisés a contribué à remettre en cause le principe de la division technique du travail comme moteur de la croissance dans certaines branches.

L'importance croissante de l'information et du savoir dans l'activité économique modifie le critère de la division du travail en favorisant un découpage de l'activité de production fondé sur les savoirs. L'efficacité dans l'acquisition des connaissances réclame la spécialisation (Demsetz, 1991). Dans cette perspective, le processus de production est décomposé en blocs de savoirs homogènes⁽⁷⁾. En première approximation, on peut définir un bloc de savoirs comme un ensemble de connaissances rattachées à un même corps de principes scientifiques et techniques, soumises à une dynamique d'évolution commune impulsée par une activité de recherche et de transformation des informations en nouvelles connaissances obéissant à certaines heuristiques partagées par une communauté de spécialistes. On le voit, cette notion de bloc de savoirs est très proche de celle de paradigme technologique développée par Dosi⁽⁸⁾.

Le progrès des sciences et des techniques a eu un effet d'éclatement sur les blocs de savoirs (Ayres, 1991), faisant évoluer chacun le long d'une trajectoire propre, conduisant à un éclatement rendant plus difficile la communication entre spécialistes de blocs différents. Les progrès réalisés au sein de chaque bloc s'accompagnent généralement d'un accroissement simultané de sa complexité et de sa spécificité, plaidant toujours davantage pour une spécialisation des individus utilisant ces savoirs ou engagés dans le processus de création de

(7) Piore (1991) utilise l'expression de «catégories conceptuelles» dans un sens très proche de celui que nous donnons aux blocs de savoirs.

(8) Voir Dosi (1984, 1988).

nouvelles connaissances. Dans une activité économique où l'information et la connaissance ont acquis une valeur stratégique, où la capacité à tirer profit du changement permanent de l'environnement est un avantage concurrentiel majeur, où le coût de développement de nouvelles connaissances s'élève..., la fragmentation de l'activité économique selon la nature des blocs de savoirs apparaît comme efficace à la fois pour la pleine maîtrise des connaissances constituées, pour réaliser des économies d'informations (Demsetz, 1991) et pour favoriser le développement des compétences et la création de nouvelles connaissances.

Le glissement du critère de la division du travail d'une logique technique à une logique cognitive s'accompagne souvent d'une certaine recomposition du processus de production⁽⁹⁾, la logique technique de décomposition menant généralement à un plus grand fractionnement du processus que la logique cognitive. La logique cognitive ne saurait s'épanouir dans un processus de production trop morcelé (à l'intérieur des blocs de savoirs) qui jouerait certainement comme un frein à la créativité et un obstacle à la réalisation d'économies de champ. La séparation stricte entre la conception et l'exécution s'estompe; la qualité, la créativité, la capacité à apprendre... de l'ensemble dépend notamment de la facilité de communication entre les parties, du degré d'engagement de chacune dans l'activité de l'entreprise,... La division cognitive du travail s'accompagne donc d'une certaine recomposition fonctionnelle. Le facteur travail tend à perdre de son hyperspécialisation caractéristique de la division technique du travail; s'il reste spécialisé, c'est sur un champ conceptuel homogène où la compétence s'approfondit avec la pratique. En revanche, il étend sa polyvalence sur un ensemble de tâches plus large (à l'intérieur d'un bloc de savoirs). En ce sens, le degré de spécificité du facteur travail s'accroît dans la mesure où le passage d'un bloc de savoirs à un autre pour un travailleur est moins aisé que le passage d'un poste de travail à un autre dans la division technique du travail. Grâce à l'incorporation des TIC, le capital perd en revanche

⁽⁹⁾ Ceci a amené des auteurs tels que Kern et Schumann (1989) à parler de la fin de la division du travail. Cette expression nous paraît abusive. D'une part, elle réduit le principe de division du travail à l'organisation dans les ateliers. D'autre part, même au sein des ateliers, l'application du nouveau principe de division du travail peut conduire à désintégrer certaines activités plus finement que sous la logique technique. Par exemple, dans le domaine des composants électroniques, Carroué (1990) montre comment la complexification de la technologie et les impératifs commerciaux ont conduit à une désintégration de la fonction de conception (s'accompagnant d'ailleurs de l'éclatement international de cette activité).

beaucoup de sa spécificité (mais aussi de son caractère prioritairement déterminant dans la compétitivité de l'entreprise).

Cette division du travail fondée sur les savoirs n'est pas une nouveauté. C'est elle qui a toujours régi l'organisation de la production artisanale qui était dominante avant l'ère de la production en série. Piore et Sabel (1984) voient dans la reconfiguration des processus de production en blocs de savoirs homogènes (ainsi que dans le développement des outils de production flexibles) une bifurcation historique pouvant mener au retour des formes artisanales comme mode dominant d'organisation de la production. Ceci ne nous semble pas constituer une conséquence nécessaire de la généralisation du principe de division cognitive du travail.

II. LES MODES DE COORDINATION ET DE RÉINTÉGRATION

Si la division du travail conduit à une structure de l'activité de production ne correspondant plus à la structure de la consommation, une phase de «réintégration» des fragments de production est nécessaire pour aboutir à des biens consommables. Autrement dit, les fragments de produits ou les morceaux de processus de production doivent être assemblés pour aboutir à la création d'une valeur d'usage finale. Cette réintégration suppose la coordination, sous une forme ou sous une autre, de l'activité des différentes parties prenantes de la production du bien. La manière dont s'opère cette réintégration du travail, dont sont coordonnées les activités, ainsi que la nature des contraintes que fait peser cet impératif de réintégration sur l'organisation de la production, dépendent du type de division du travail qui a été mis en œuvre.

Dans la division technique du travail, les différents fragments de processus de production doivent se révéler techniquement compatibles de façon à ce que la réintégration ne soit pas de nature à faire perdre le bénéfice de la fragmentation. Si l'on compare la phase de réintégration à la réalisation d'un puzzle, l'impératif de rendement commande que chaque pièce ait été précisément façonnée de manière à trouver sa place dans la construction d'ensemble sans avoir besoin d'une adaptation préalable. L'efficacité dans cette phase de réintégration est d'autant plus fondamentale que le puzzle a été composé d'un très grand nombre de très petits morceaux, chacun présentant un contour très

spécifique. Les fragments du processus de production se doivent d'être techniquement compatibles et d'obéir à des standards de qualité. Cet impératif pousse à une standardisation des fragments (à l'intérieur d'un processus de production donné) qui contribue à la rigidité d'ensemble du mode de production. Assurer une parfaite compatibilité des divers fragments du processus de production exige une coordination étroite des activités individuelles et plaide pour un large degré d'intégration de l'activité au sein de l'entreprise. La question de la réintégration interfère donc avec celle des contours de l'entreprise. Par la coordination hiérarchique à l'intérieur de la firme, il est possible de veiller *ex ante* à la compatibilité des fragments de processus. Ceux-ci étant conçus par l'entreprise, elle est en mesure de leur donner le design le plus adapté. Le recours à une fourniture extérieure risque de se révéler très onéreux en termes de coût de transaction, étant donné la spécificité des « actifs » considérés. La coordination par le marché se rencontre ainsi surtout au niveau des fragments les plus standardisés. L'incitation à l'internalisation est d'autant plus forte que la division du travail est profonde, aboutissant à la multiplication de fragments très fins, organiquement interdépendants, et exigeant des équipements productifs très spécialisés⁽¹⁰⁾. C'est probablement une des raisons qui ont poussé à la constitution de très grandes entreprises au cours des cinquante dernières années à mesure que s'approfondissait la logique de la production en série. La sous-traitance est une forme de coordination intermédiaire entre l'organisation interne et le marché. Elle permet de bénéficier des avantages de l'externalisation (rendements d'échelle, économie en capital...), mais, grâce au cahier des charges et aux divers contrôles exercés par le donneur d'ordre, elle offre davantage de garanties de compatibilité que le marché. Dans cette optique, la sous-traitance est pensée (et vécue) comme une quasi-intégration (Houssiaux, 1957).

Relevons ici une des limites de la division technique du travail. A mesure que le processus de production se fragmente, la hiérarchie

(10) La spécificité des actifs est, selon Williamson, un facteur essentiel de génération de coûts de transaction rendant plus incertain le recours au marché comme mode de coordination. En effet, lorsque la transaction fait intervenir des actifs spécifiques, du fait de son extrême spécialisation, le capital du fournisseur (ou le type de qualification de sa main-d'œuvre) n'a qu'une très faible valeur pour une autre utilisation, et le fournisseur se trouve prisonnier de son client. Ce dernier est lui aussi prisonnier de son fournisseur car il ne peut trouver facilement d'autres fournisseurs disposant d'un capital présentant une telle spécialisation, et le recours à un fournisseur non pareillement équipé se révélerait moins efficace, donc plus coûteux. Les deux parties se trouvent ainsi en situation de monopole bilatéral engendrant de coûteuses négociations pour le partage de la rente commune (Williamson, 1979).

s'impose avec plus de force pour assurer une coordination de plus en plus complexe et qui s'accommode mal des incertitudes de la relation marchande. Ce faisant, la dimension de l'entreprise augmente, son organisation s'alourdit et les coûts de coordination s'élèvent. Cette inflation des coûts d'organisation plaide pour une externalisation, alors même que la coordination hiérarchique s'imposait pour contourner les difficultés de la coordination marchande face à l'échange de fragments de plus en plus spécifiques. Cette contradiction pousse à la recherche de formes alternatives de coordination. Les TIC ont permis dans une certaine mesure de réduire cette contradiction. En améliorant les possibilités de contrôle et en facilitant les procédures de coordination, elles autorisent la poursuite de la croissance de l'entreprise, notamment au travers de la mise en place de structures organisationnelles plus décentralisées⁽¹¹⁾, et elles réduisent simultanément les possibilités de comportements opportunistes à l'occasion de relations de marché (Malone *et alii*, 1987, Brousseaux, 1992). Ceci est une illustration de ce que, en soi, le développement des TIC ne s'accompagne pas nécessairement de la mise en place de nouvelles pratiques organisationnelles mais, au contraire, peut assurer une certaine pérennité à des pratiques qui paraissaient condamnées⁽¹²⁾.

La division cognitive du travail impose une réintégration des fragments produits à l'intérieur des différents blocs de savoirs. Là aussi, les fragments doivent présenter un niveau élevé de compatibilité. Cette nécessité plaide pour l'internalisation dans l'entreprise de la production des différents fragments concourant au bien final. Cependant, si la coordination hiérarchique offre une garantie de compatibilité, à mesure que la firme grandit, elle est aussi de nature à étouffer la créativité de ses membres et à nuire à l'efficacité de l'apprentissage (Gordon, 1990). Or, au sein de la division cognitive du travail, sans doute plus que la compatibilité technique, ce qui importe c'est la compatibilité

(11) «Au début, les micro-ordinateurs autonomes entraînent un déplacement de pouvoir vers le bas. Avec ces nouvelles armes, les cadres moyens et même les simples employés goûtèrent à une liberté et à un pouvoir inaccoutumés. Mais quand leurs appareils furent effectivement reliés à l'unité centrale, la haute direction fut en mesure de surveiller de près certains paramètres clés de l'activité de très nombreuses petites unités : il devenait possible de laisser une autonomie d'action considérable, tout en gardant une emprise sur leur comptabilité financière» (Toffler, 1990, p. 268).

(12) Cette idée s'oppose aux théories technologistes de Perez, Freeman, Soete..., qui voient dans la diffusion des TIC le vecteur essentiel de transition vers un nouveau mode de fonctionnement de l'économie et de l'apparition de nouveaux modes d'organisation de la production.

dynamique des fragments issus des différents blocs de savoirs, c'est-à-dire **leur capacité à progresser de concert**, sans que l'inertie au niveau de l'un des blocs de savoirs n'entrave la dynamique de l'amélioration globale des performances du produit fini. Ceci pousse l'entreprise à se spécialiser sur un nombre limité de blocs de savoirs, et ce d'autant plus que l'environnement évolue rapidement, exigeant une forte capacité d'apprentissage et de flexibilité (Dosi, Teece et Winter, 1990). En effet, l'instabilité de l'environnement impose de pouvoir reconfigurer très rapidement la combinaison des blocs de savoirs : dépérissement de certains blocs, apparition de nouveaux, modification en chaîne des différents fragments en raison de la relation organique qui les lie... Dans ce contexte, une intégration de l'entreprise poussée trop loin se traduit par une rigidité préjudiciable à sa compétitivité (Storper, 1992). A cela s'ajoute le coût croissant de la maîtrise des blocs de savoirs qui incite les entreprises à concentrer leur effort d'apprentissage. On comprend aisément, par exemple, que la dispersion d'un budget de R&D sur des domaines variés risque de se révéler inefficace face à des concurrents qui auraient concentré les mêmes ressources sur le développement de blocs de savoirs plus étroits. C'est précisément à une concentration des budgets de recherche sur des champs étroits et homogènes que l'on est en train d'assister, par exemple dans l'industrie pharmaceutique, en réponse au coût croissant de la découverte et du développement des médicaments et à l'accroissement de la pression concurrentielle (Loustau, 1991).

A l'autre extrémité, une spécialisation trop étroite pose un problème de communication avec les partenaires assurant d'autres fragments du processus de production. La spécialisation par blocs de savoirs se justifie notamment par les gains d'information qu'elle permet de réaliser. La condition de ce gain est cependant que l'utilisation d'un bien (ou d'un fragment) exige moins de savoirs et de compétences que sa production (Demsetz, 1991). S'il en était autrement, l'utilisateur devrait détenir les mêmes compétences que le producteur et serait donc *a priori* en mesure de produire le bien lui-même. L'entreprise engagée dans une division cognitive du travail doit témoigner d'un degré d'intégration suffisant pour que son produit (ou service) épargne à l'utilisateur la maîtrise des connaissances nécessaires à sa production.

Le choix des blocs de savoirs qui seront maîtrisés directement par l'entreprise dépendra de plusieurs facteurs :

– Etant donné le caractère cumulatif et idiosyncratique des compétences, l'histoire de la firme, de ses succès et de ses échecs aboutit généralement à faire apparaître des zones de supériorité (ou

de moindre infériorité) sur certains blocs de savoirs nécessaires à la production, par rapport à ses concurrents. C'est bien sûr en priorité sur les blocs de savoirs où l'entreprise bénéficie d'éléments de supériorité et d'une capacité à apprendre construite au cours du temps que va porter la spécialisation.

– Les blocs de savoirs nécessaires à une production donnée n'ont pas tous la même importance stratégique. Si pour vendre des chaussures de sport, il faut mobiliser des blocs de savoirs relatifs au design, au travail du cuir, au marketing..., la configuration du régime de concurrence et la nature du processus de fabrication rendent la maîtrise du marketing bien plus stratégique que celle de la fabrication. C'est bien ce qu'ont compris les firmes américaines Nike et Reebok qui, en sous-traitant quasi intégralement leur production et en concentrant leurs ressources sur la communication, ont complètement déstabilisé les leaders traditionnels intégrés comme Adidas et Puma (Moati, 1989).

– Contrairement à la logique des coûts de transaction où l'on raisonne en termes de coûts comparés, l'entreprise n'a souvent pas à se poser la question de savoir si elle doit recourir à l'extérieur pour la mobilisation d'un bloc de savoirs nécessaire à la production de son produit. La compétence n'est pas un bien libre et est d'autant plus difficilement imitable qu'elle se développe selon une trajectoire qui lui est spécifique. Ainsi, bien souvent, le problème n'est pas de savoir s'il est plus économique de faire ou de faire faire, mais simplement de faire faire parce qu'on ne sait pas faire. La décision d'internaliser, dans un contexte de division du travail fondée sur les savoirs, est donc souvent moins une affaire d'avantage comparatif qu'une affaire d'avantage absolu ou d'indisponibilité pure.

Au total, c'est bien la nature des compétences qu'a réussi à créer et à développer l'entreprise qui lui indique les blocs de savoirs qu'elle est en mesure de prendre en charge efficacement. Il lui faudra organiser le travail en son sein de manière à favoriser l'apprentissage collectif et développer sa «réactivité» aux menaces et opportunités d'un environnement instable (Cohendet et Llerena, 1990). Cet impératif engendre des modifications des modalités de la coordination hiérarchique telle qu'elle est pratiquée traditionnellement dans la division technique du travail. L'hyperspécialisation de la main-d'œuvre est révisée au profit d'une certaine recomposition des tâches effectuées par chacun. Les équipements productifs programmables jouent un rôle très important dans cette recomposition. L'objectif est d'obtenir un plus grand engagement des individus devant se

traduire par une plus forte responsabilisation, plus d'initiative face à l'occurrence d'aléas, et une meilleure capacité à développer ses compétences par l'apprentissage. Favoriser l'apprentissage collectif signifie revoir la structure organisationnelle en effaçant quelque peu la rigidité des relations hiérarchiques qui cloisonnent les fonctions, figent les flux d'informations et limitent la capacité d'initiative des individus (Aoki, 1990). On cherche plutôt à favoriser des modes de coordination horizontaux, afin de faciliter les échanges entre des spécialistes complémentaires (pas nécessairement situés sur la même ligne hiérarchique) et d'aboutir à l'intégration efficace des savoirs de chaque membre de l'entreprise. A cet effet, la bonne circulation de l'information est un objectif prioritaire. L'utilisation des TIC en réseau apparaît alors comme un moyen efficace d'assurer une très large diffusion d'informations de format normalisé.

Si la contrainte d'efficacité dynamique pousse les entreprises à se spécialiser sur un nombre limité de blocs de savoirs, menant en général à un degré d'intégration moins poussé que dans la logique technique de division du travail, le marché se révèle souvent imparfait comme mode de coordination et d'intégration de ces blocs de savoirs en vue de la production de valeurs d'usage. L'échange d'outputs de blocs de savoirs est empreint d'une dose importante d'incertitude : en raison du principe de spécialisation, l'acheteur est difficilement en mesure d'apprécier pleinement les caractéristiques et la qualité du produit vendu, ce type d'information étant très mal transmis par le système de prix. Cette asymétrie d'information entre le fournisseur et le client encourage les comportements opportunistes de la part du premier. Par ailleurs, les fragments issus des différents blocs de savoirs concourent organiquement aux performances globales du produit final et leur combinaison est ainsi génératrice d'une quasi-rente irréductible à la somme des apports de chacun (Amendola et Gaffard, 1988). De même, l'amélioration constante des produits exige une étroite collaboration entre les parties prenantes (Lundvall, 1985). *«L'intensité des interactions parmi les spécialistes est trop intense pour permettre un marché»* (Piore, 1991, p. 149). L'extension de la division cognitive s'est donc accompagnée du développement de nouveaux modes de coordination, dépassant la dichotomie marché-hiérarchie. Ces nouveaux modes de coordination ont pour objet de faciliter l'intégration d'informations et de savoirs bien plus que le simple transfert de marchandises. Ils doivent permettre une communication efficace entre spécialistes possédant chacun une culture propre à leur bloc de savoirs et difficilement communicable, et favoriser l'apprentissage collectif

nécessaire à l'amélioration des performances du produit et à la découverte de nouvelles associations technologie-marché⁽¹³⁾. Ainsi les modes de coordination au sein d'une division cognitive du travail doivent garantir, outre la compatibilité matérielle des fragments issus de la division, la compatibilité des hommes impliqués dans la production du bien final (capacité de se comprendre et d'apprendre ensemble). Seul, le système des prix se révèle peu adapté à cette fonction. Il a besoin d'être complété par un certain nombre de «conventions» régissant les relations entre spécialistes de différents blocs de savoirs. Les «connaissances communes» (common knowledge), comme le fait de partager un langage, des références culturelles, la connaissance de principes scientifiques et techniques de base... sont des formes de conventions réduisant le coût de la communication entre spécialistes (Demsetz, 1991). De telles connaissances communes peuvent aussi se construire au cours des relations, par un mécanisme d'apprentissage réciproque (Williamson, 1979). L'efficacité de la coordination entre les spécialistes est donc favorisée lorsque les relations s'inscrivent dans la durée. L'établissement de «codes de conduite», plus ou moins explicites, peut également se révéler comme un moyen efficace de coordonner l'activité des spécialistes : les codes de conduite réduisent les coûts de transaction en épargnant des négociations coûteuses et limitent les tentations de comportements opportunistes.

Les exemples de ces nouveaux modes de coordination sont nombreux. Ils sont souvent résumés par l'expression de «réseau» : des entités spécialisées dans des blocs de savoirs spécifiques sont reliées dans un ensemble de relations coopératives, dans le cadre d'un «marché organisé» (Lundvall, 1985). Le réseau est construit sur les compétences technologiques des participants et non pas sur des liaisons financières (Michalet, 1990) : le réseau peut ainsi rassembler des partenaires juridiquement indépendants, des joint-ventures, des filiales d'un même groupe, des établissements d'une même firme... Le réseau peut être articulé autour d'une firme pivot : c'est fréquemment le cas au Japon, et c'est le modèle que l'on rencontre dans l'industrie automobile. La firme pivot acquiert généralement ce statut en raison de sa maîtrise des fragments stratégiques du processus de production (ceux dont dépendent le plus étroitement les performances du produit final) et de ses conditions d'accès au client final (image de marque, réseau de distribution constitué...). La firme pivot joue un rôle important dans la

(13) En s'inspirant de Guilhon (1992), on peut avancer que la division cognitive du travail implique à la fois une désintégration cognitive et une intégration informationnelle.

définition des codes de conduite au sein du réseau (normes techniques, pratiques commerciales, orientation des recherches...). Ces codes de conduite sont souvent formalisés à travers des contrats de long terme liant les partenaires à la firme pivot. Les avantages de ces contrats de long terme sont nombreux :

– Ils permettent d'économiser en coûts de transaction en évitant des négociations permanentes et coûteuses. Le contrat est suffisamment général pour autoriser une adaptation de la prestation à l'évolution des besoins (modification des volumes commandés, changements dans la spécification des pièces...).

– Les partenaires, rassurés par la stabilisation de l'horizon temporel de la collaboration, sont incités à engager des investissements spécifiques visant l'amélioration de la productivité et de la performance du produit intermédiaire (Guilhon et Gianfaldoni, 1990). La compatibilité statique et dynamique des fragments s'en trouve renforcée.

– Ils assurent des formes de contrôle proches de celles à l'œuvre à l'intérieur de l'entreprise. Par exemple, au travers du concept d'assurance qualité, le client est en mesure de s'assurer de la qualité du produit de son fournisseur en procédant à des audits auprès de ce dernier afin d'avaliser ses processus de fabrication et de contrôle (Garreau, 1989).

Il peut également exister des réseaux n'impliquant pas de firme pivot. C'est le cas notamment lorsque des grandes firmes concluent une alliance stratégique pour mener à bien des programmes de R&D dans des domaines où le coût de développement du savoir dépasse les capacités de firmes isolées, ou bien lorsque des avancées significatives réclament le croisement des apports de plusieurs blocs de savoirs (Delapierre et Michalet, 1989)⁽¹⁴⁾. C'est également le cas au sein des « districts » où se forme un maillage étroit de coopération entre des PME⁽¹⁵⁾. En l'absence d'une firme pivot jouant le rôle d'attracteur et de diffuseur d'un code de conduite assurant la compatibilité nécessaire à une réintégration efficace des fragments, ce type de réseaux est davantage fondé sur la coïncidence des intérêts des parties prenantes et sur une certaine homogénéité culturelle garantissant une compatibilité de langage, d'habitudes, de valeurs, de vision du monde..., autant d'éléments favorisant une coordination conventionnelle des comportements des membres du réseau. L'existence de telles conventions est essentielle

(14) C'est la conjugaison de ces deux facteurs qui a poussé à la constitution de consortiums internationaux dans la construction aéronautique.

(15) Voir, par exemple, Pyke, Becattini et Sengenberger (1990), Peyrache (1992),...

afin de réduire les tensions qui ne sauraient manquer de naître de la difficulté de partager la rente créée collectivement par la coopération de facteurs spécifiques. Comme le rappelle A. Leijonhufvud (1989), «*les communautés stables et relativement soudées sont considérées comme jouissant en fait d'un avantage comparatif pour la stabilisation des modèles de coopération entre de nombreuses entreprises qui perçoivent ensemble des rentes communes. Elles peuvent faire peser des valeurs culturelles communes sur les parties contractantes et au moins rétrécir ainsi l'éventail des indéterminations inhérentes au monopole bilatéral*» (p. 187). Des éléments culturels intériorisés par les agents se substituent alors à la pression de la firme pivot pour définir les codes de conduite nécessaires.

Il convient d'éviter la conclusion hâtive que la division du travail ramène à un capitalisme de petites entreprises indépendantes. La division cognitive du travail favorise très certainement la décentralisation de l'activité économique. Toutefois, que ce soit par l'exigence d'une forte compatibilité statique et dynamique des fragments de processus de production qui réclame souvent l'action d'un «*attracteur*» capable d'imposer des normes, ou que ce soit par l'élévation des coûts fixes associés au développement des savoirs et à la réussite de stratégies commerciales fondées sur la différenciation, les rendements croissants sont encore suffisamment importants dans la vie économique pour justifier l'existence et le rôle capital des grandes firmes (Boyer, 1991). Leur rôle évolue toutefois vers celui d'animateur et de coordonnateur d'un réseau d'unités (contrôlées financièrement ou non) disposant d'une large autonomie de gestion.

La maîtrise d'un bloc de savoirs autorise souvent de multiples applications dans des secteurs variés. Ainsi, les entreprises engagées dans la division cognitive du travail peuvent présenter des portefeuilles de produits et d'activités en apparence hétérogènes. Leur cohérence vient du rattachement à un même bloc de savoirs. Cette particularité explique que ces entreprises puissent participer simultanément à plusieurs réseaux.

III. LES DÉTERMINANTS DE LA LOCALISATION DES ACTIVITÉS PRODUCTIVES

Conformément à la logique de minimisation des coûts à la base de la division technique du travail, le principe de localisation de

l'activité productive sera de produire chaque fragment là où les coûts de production pourront être les plus faibles. La décomposition internationale des processus productifs (DIPP) (Lassudrie-Duchêne, 1982, note de référence) est l'application de cette logique à l'échelle internationale. C'est à partir d'une décomposition technique du processus de production que s'établit la DIPP⁽¹⁶⁾. Le principe de localisation selon les avantages comparatifs de pays correspond bien, pour l'essentiel, à une logique de minimisation des coûts (Lassudrie-Duchêne, 1982, Lassudrie-Duchêne *et alii*, 1986).

L'impératif de réintégration pèse sur la possibilité de bénéficier pleinement des opportunités de DIPP. Là encore, il ne s'agit pas de perdre en coûts de transfert les gains obtenus par les avantages de localisation. On comprend que les coûts de transport puissent avoir une influence déterminante sur la faisabilité et l'étendue du champ géographique de la DIPP. Le degré de l'interdépendance technique entre fragments influence aussi l'opportunité de localisation. C'est ainsi par exemple que les «produits systèmes», grâce au caractère synchronique de leur processus de production qui concentre les contraintes d'interdépendance technique sur la phase d'assemblage, sont de meilleurs supports à la DIPP que les «produits filières». Une défaillance à un stade quelconque d'un processus de production diachronique interrompt l'ensemble de l'activité. Or, l'éloignement géographique accroît les risques de défaillance (aléas de transport, circulation de l'information ralentie,...).

La division cognitive du travail implique d'autres critères de localisation de la production. De manière relativement triviale, on peut dire que chaque fragment du processus de production ne pourra être localisé que dans les régions (au sens large) où est maîtrisé le bloc de savoirs requis. A première vue, on semble retrouver ici une logique d'avantages comparatifs. Toutefois, rappelons-nous que la maîtrise d'un bloc de savoirs, ainsi que la capacité à en repousser les limites, sont rarement le résultat de l'usage de facteurs de production génériques. C'est la raison pour laquelle l'approche de la théorie du commerce international, fondée sur les différences de dotation des pays en facteurs de production, est d'un intérêt limité pour comprendre les critères de localisation de l'activité productive dans une logique de division cognitive du travail. Retenons toutefois que maîtriser des blocs de savoirs requiert le plus souvent une main-d'œuvre bénéficiant d'un haut

(16) «La DIPP est avant tout un phénomène technique. Aussi convient-il de rechercher les déterminants techniques du fractionnement des processus avant de poser le problème de la localisation des unités de production» (Fontagné, 1991, p. 99).

niveau de formation et de qualification. Ainsi, le degré de qualification du facteur travail d'un pays définit de manière très globale la position que ce pays pourra prétendre occuper dans une division cognitive internationale du travail, c'est-à-dire le degré de sophistication des blocs de savoirs pour lesquels il sera susceptible de bénéficier d'une compétitivité suffisante. Mais pour dépasser ce principe général, qui est peu prédictif des spécialisations fines, il faut passer à une approche fondée sur les dotations en facteurs spécifiques.

Ces facteurs spécifiques se créent, se développent ou disparaissent, fondamentalement à travers la dynamique individuelle des entreprises et les relations qui se nouent entre elles au cours de l'activité de production. Or, étant donné les mécanismes de formation et de développement des compétences, ces facteurs spécifiques sont justement tellement spécifiques qu'ils sont souvent idiosyncratiques à l'entreprise ou à un groupe d'entreprises en étroite collaboration au sein d'un réseau : ils naissent de la capacité d'apprentissage et de la créativité des entreprises; ils se développent et se diffusent à travers les stratégies d'innovation et d'imitation menées au niveau micro-économique. Le stock de chaque facteur spécifique dont bénéficie un pays est ainsi indissociable de l'histoire de ses entreprises et de la dynamique que celles-ci insufflent aux structures sectorielles. On glisse d'une logique d'avantage comparatif de pays (dotation en facteurs de production génériques) à une logique d'avantage spécifique de firme (formation de facteurs spécifiques). Les conséquences sont importantes. On peut comprendre de la sorte que des pays, pourtant caractérisés par le même degré de qualification de leur main-d'œuvre, affichent des positions très différentes en termes de compétitivité sur les différents blocs de savoirs. Les écarts naissent des différences de trajectoires suivies par les industries dans les différents pays. Ces différences de trajectoires sont elles-mêmes le résultat de la dynamique engendrée par la lutte concurrentielle à laquelle se livrent des firmes hétérogènes dans chaque pays. A partir d'un modèle d'échange international avec des firmes hétérogènes dans leur capacité à transformer le progrès des connaissances scientifiques en accroissement de la productivité, Moati (1992) montre que cette hétérogénéité des compétences micro-économiques fait suivre à ces industries des trajectoires pouvant être très différentes entre des pays partageant rigoureusement les mêmes caractéristiques macro-économiques (coût des facteurs, niveau technologique,...). De ces dynamiques contrastées naissent des flux d'échange international (inimaginables, étant donné les hypothèses, pour

les théories traditionnelles) pouvant donner lieu à des spécialisations inter-sectorielles marquées.

Comme dans la division technique du travail, l'impératif de réintégration de fragments pèse sur la localisation des activités de production. La division cognitive du travail réclame souvent la proximité géographique des parties prenantes au processus de production. Cette proximité géographique est de nature à favoriser l'apprentissage collectif, en facilitant la communication des idées et en accélérant les transactions. Elle implique souvent une proximité culturelle des partenaires qui facilite l'établissement d'une coordination conventionnelle améliorant la qualité de la communication (Pyke, Becattini et Sengenberger, 1990; Giordano, 1991). C'est notamment pour cette raison que, par exemple, la division avion de l'Aérospatiale a procédé au regroupement sur un même site spécialiste de tous les moyens rattachés à une même technologie (de Bortoli, 1992). Ces incitations à la proximité géographique sont bien sûr renforcées par la mise en œuvre des pratiques de gestion de type JAT. Toutefois, comme l'a écrit J.-C. Perrin (1990) au sujet de la territorialisation des entreprises associées à Toyota au Japon, *«cette forme de localisation répond tout autant à la contrainte des flux tendus qu'à l'objectif de mieux intégrer les développements technologiques initiés par Toyota et les améliorations techniques réalisées par les associés»* (p. 299). Néanmoins, la mise en œuvre de la division cognitive du travail n'entraîne cependant pas nécessairement de l'absence de DIPP, même si elle s'accompagne généralement d'une certaine recomposition et d'un certain rapprochement géographique des fragments du processus de production (Mouhoud, 1989). Les nouvelles formes de «DIPP» associées à la division cognitive du travail ont pour critère, non pas la minimisation des coûts, mais l'accès à des facteurs spécifiques maîtrisant les blocs de savoirs nécessaires. De façon générale, la diffusion de la division cognitive du travail tend à renforcer le poids de la technologie (en tant que créatrice de monopoles et d'indisponibilités) et des facteurs tenant aux avantages spécifiques de firmes dans la détermination des flux d'échanges internationaux.

L'ensemble de ces considérations plaide pour une concentration des activités de production obéissant à une logique cognitive de division du travail, au sein des pays industrialisés. C'est un des facteurs qui incitent à penser que les opérations de délocalisation vers les pays à bas salaires ne devraient pas connaître un important dynamisme au cours des prochaines années (Michalet, 1990), accélérant le processus en cours

de marginalisation des pays en développement au sein de l'économie mondiale (Mouhoud, 1992).

CONCLUSION

Les deux dernières décennies ont été marquées par des changements structurels de l'environnement des entreprises. Ces changements se sont traduits par une redéfinition de la nature de l'activité de l'entreprise qui passe ainsi d'un lieu de combinaison optimal de facteurs de production génériques à un lieu de combinaison de compétences et d'apprentissage de nouvelles connaissances à partir de facteurs spécifiques. Cette redéfinition s'accompagne de la mise en œuvre d'une division cognitive du travail (décomposition des processus de production par blocs de savoirs), qui tend à se substituer à la division technique du travail caractéristique des modes de production fordien. Le JAT, le «toyotisme»... sont analysés comme étant des formes organisationnelles associées à ce nouveau principe de division du travail. Mais elles ne résument pas à elles seules la nouvelle nature de la firme et la division cognitive du travail, qu'il est concevable de voir associée à d'autres formes organisationnelles ou d'autres techniques de gestion. L'instauration d'une division cognitive du travail implique une redéfinition complète de la manière de segmenter le processus de production, de délimiter les contours des entreprises, de définir les critères de localisation des unités de production, et d'élaborer de nouveaux modes de coordination de l'activité des différentes parties prenantes à chaque processus de production.

Le champ d'application de la division cognitive semble être très général puisque, comme nous l'avons vu, elle guide l'organisation interne de l'entreprise, elle définit le partage de l'activité entre entreprises, et elle est un vecteur de la définition de la spécialisation internationale des nations.

L'idée d'un glissement complet du principe de division technique à celui de division cognitive du travail doit être nuancée. La division cognitive du travail est un mode d'organisation de la production émergent. Le processus en cours reste encore loin d'être généralisé. Comme le rappellent Boyer et Schmeder (1990), «*on continue à enregistrer la prégnance des principes fordien dans la plupart des branches industrielles*» (p. 174). Au niveau macro-économique, on observe bien la coexistence des deux logiques d'organisation de

la production, l'une dominant l'autre selon les compartiments de l'économie. L'hypothèse est que la division cognitive du travail a vocation à s'étendre et à devenir le mode dominant d'organisation de la production. Dans cette hypothèse, ce glissement du principe de la division du travail revêtirait un caractère irréversible. Certes, le passage à une division cognitive du travail se trouve favorisé par le contexte d'incertitude, de concurrence exacerbée, de contrainte financière... qui caractérise l'environnement des entreprises dans la période de récession. Toutefois, la division cognitive du travail ne nous semble pas devoir être analysée comme le simple fruit de stratégies d'adaptation à une conjoncture défavorable et qui pourrait être remise en cause avec la reprise. La division cognitive du travail apparaît plutôt comme une réponse des entreprises à la fois à la crise des modes d'organisation issus du fordisme et à la contradiction entre la complexification des techniques et leur nécessaire hybridation. Si les formes organisationnelles de sa mise en œuvre dépendent sans doute de la conjoncture macro-économique, sa logique semble se dégager du processus de transformation structurelle associé à la crise.

Même dominante, la division cognitive du travail ne saurait cependant s'imposer comme un modèle universel. Outre que certains secteurs (probablement ceux mobilisant les savoirs les plus banalisés) continueront de fonctionner sous la logique technique, on pourra observer la coexistence des deux logiques à l'intérieur même des processus de production. Par exemple, la domination de la logique cognitive au sein d'un processus de production n'implique pas qu'une fragmentation selon les blocs de savoirs se révèle être pertinente d'un bout à l'autre. Pour les parties du processus de production les plus banalisées, l'application d'une logique de division technique du travail peut se révéler être un meilleur moyen de contribuer à la compétitivité globale du produit final. Symétriquement, un processus de production obéissant de manière dominante à une logique de division technique du travail peut connaître certains fractionnements conformes à la logique cognitive (en particulier sur les aspects du processus de production faisant appel aux technologies les plus sophistiquées).

L'hypothèse de glissement du principe de la division du travail demanderait à être validée par des travaux empiriques systématiques, qui devraient en particulier s'attacher à identifier les caractéristiques sectorielles favorables à la diffusion de la logique cognitive, les modalités nationales et sectorielles de sa mise en œuvre, les caractéristiques structurelles et stratégiques d'entreprise rendant compte de la dispersion intra-sectorielle de l'engagement des firmes dans l'une ou l'autre

logique de division du travail... L'étude par les auteurs ⁽¹⁷⁾ des secteurs de l'automobile et de l'habillement apporte des éléments partiels de réponse à ces questions. L'inégal engagement des secteurs dans la division cognitive du travail repose notamment sur les caractéristiques des produits (contenu en technologies, caractère diachronique ou synchronique des processus...), sur l'existence d'opérateurs capables de jouer le rôle d'«attracteurs» dans le secteur, en diffusant des codes de conduite et en assurant une forte cohésion d'ensemble, sur les modalités de la concurrence sur le marché... Au niveau des entreprises, la spécificité des produits de l'entreprise, son positionnement stratégique, le rapport de force à l'égard des fournisseurs et des clients, sa capacité à affirmer un savoir-faire reconnu,...., sont quelques-unes des variables rendant compte des écarts d'engagement dans les nouvelles formes d'organisation de la production.

D'autres travaux empiriques sont nécessaires, en particulier afin de fournir une meilleure définition empirique de la notion de blocs de savoirs. De même, un effort de réflexion doit être mené afin de tirer les conséquences de la diffusion de la division cognitive sur le fonctionnement global de l'économie. Une piste de recherche porte sur le bouclage avec la demande. Nous avons vu la relation dynamique vertueuse qui liait la division technique du travail et l'extension du marché. Un tel bouclage avec la demande semble moins évident dans le cas de la division cognitive du travail. Est-ce à dire qu'une économie régit par ce principe de division du travail n'est pas entraînée dans une logique de croissance soutenue?

RÉFÉRENCES BIBLIOGRAPHIQUES

- Amendola M., Gaffard J.-L., *La dynamique économique de l'innovation*, Economica, Paris, 1988.
- Aoki M., «Toward an Economic Model of the Japanese Firm», *Journal of Economic Literature*, vol. XXVIII, March 1990, p. 1-27.
- Ayres R., «Evolutionary Economics and Environmental Imperatives», *Structural Change and Economic Dynamic*, vol. 2, n° 2, December 1991, p. 255-273.
- Beniger J.R., *The Control Revolution : Technology and the Economic Origins of the Informations*, Harvard University Press, Cambridge, Mass, 1986.
- Bortoli (de), «L'Aérospatiale», *Communication au séminaire «Compétitivité des entreprises et gestion des réseaux»*, Ecole nationale des Ponts et Chaussées, Paris, 11 et 12 mars 1992.

⁽¹⁷⁾ Moati et Mouhoud (1992).

- Boyer R., *Déterminants et évolutions probables de la productivité et de l'emploi*, Cepremap, n° 7922, Paris, 1979.
- Boyer R., «Les alternatives au fordisme. Des années 1980 au XXI^e siècle», in Benko et Lipietz éditeurs, *Les régions qui gagnent*, PUF, Economie en liberté, Paris, 1992.
- Boyer R., Schmeder G., «Division du travail, changement technique et croissance. Un retour à Adam Smith», *Revue française d'économie*, vol. V, n° 1, hiver 1990, p. 125-194.
- Brousseau E., «Coûts de transaction et impact organisationnel des technologies d'information et de communication», *Réseaux*, n° 54, 1992.
- Calame P., «Plaidoyer pour une redistribution des savoirs», in *L'état des sciences et des techniques*, sous la direction de Witkowski N., La Découverte, Paris, 1991.
- Carroué L., «Réseaux de communication et organisation de la production dans les composants électroniques : les ASIC», in Bakis H., *Communication et territoires*, La Documentation française, Paris, 1990, p. 93-108.
- Commissariat général du Plan, *L'Usine du Futur. L'entreprise communicante et intégrée*, La Documentation française, Paris, 1990.
- Coriat B., *Penser à l'envers*, Bourgeois, Paris, 1990.
- Chesnaï F., «Accords de coopération interfirmes, dynamique de l'économie mondiale et théorie de l'entreprise», in Humbert M., *Investissement international et dynamique de l'économie mondiale*, Economica, Paris, 1990, p. 457-494.
- Cohendet P., Llerena P., «Nature de l'information, évaluation et organisation de l'entreprise», *Revue d'économie industrielle*, n° 51, 1er trimestre 1990, p. 141-165.
- Demsetz H., «The Theory of the Firm Revisited», in Williamson O.E. and Winter S.G. (ed), *The Nature of the Firm. Origins, Evolution, and Development*, Oxford University Press, New York et Oxford, 1991, p. 159-178.
- Dockès P., «Les recettes fordistes et les marmites de l'histoire (1907-1993)», *Revue économique*, vol. 46, n° 3, mai 1993, p. 485-527.
- Dockès P., Rosier B., *L'histoire ambiguë. Croissance et développement en question*, Paris, PUF, 1988.
- Dosi G., *Technical Change and Industrial Transformation. The Theory and an Application to the Semi-Conductor Industry*, Mc Millan Press, New York, 1984.
- Dosi G., «Sources, Procedures, and Microeconomic Effects of Innovation», *Journal of Economic Literature*, vol. XXVI, n° 3, September 1988, p. 1120-1171.
- Dosi G., «Industrial Organisation, Competitiveness and Growth», *Revue d'économie industrielle*, n° 59, 1er trimestre 1992.
- Dosi G., Teece D., Winter S., «Les frontières des entreprises : vers une théorie de la cohérence de la grande entreprise», *Revue d'économie industrielle*, n° 51, 1er trimestre 1990, p. 238-254.
- Durand J.-P. (éd.), *Vers un nouveau modèle productif?*, Paris, Syros, 1993.
- Farhi S., «Le juste-à-temps : comment les entreprises s'adaptent-elles aux à-coups de la demande?», *Problèmes économiques*, n° 2235, 24 juillet 1991, p. 17-20.
- Fontagne L., *Biens intermédiaires et division internationale du travail*, Economica, Paris, 1991.
- Freeman C., Perez C., «Structural Crises of Adjustment, Business Cycles and Investment Behavior», in Dosi G., Freeman C., Nelson R., Silverberg G. and Soete L. (ed.), *Technical Change and Economic Theory*, Pinter Publishers, London & New York, 1988, p. 38-66.

- Freeman C., Soete L., «Fast Structural Change and Slow Productivity Change : Some Paradoxes in the Economics of Information Technology», *Structural Change and Economics Dynamics*, vol. 1, n°2, December 1990.
- Freyssinet M., Thénard J.-C., «Choix d'automatisation, efficacité productive et contenu du travail», *Cahiers du GIP-Mutations industrielles*, Paris, 1988.
- Gaffard J.-L., «Innovations et changements structurels : Revue critique de l'analyse économique moderne de l'innovation et des changements structurels», *Revue d'économie politique*, vol. 100, n° 3, mai-juin 1990, p. 323-382.
- Garreau A., «Transport et juste-à-temps», *Les Cahiers de recherche*, Centre HEC-ISA, CR345, 1989.
- Giordano Y., «Décision et organisations : quelles rationalités?», *Economies et Sociétés*, Série «Sciences de Gestion», S.G., n° 17, avril 1991, p. 161-194.
- Gordon R., «Systèmes de production, réseaux industriels et régions : les transformations dans l'organisation sociale et spatiale de l'innovation», *Revue d'économie industrielle*, n° 51, 1er trimestre 1990, p. 304-339.
- Guilhon B., «Technologie, organisation et performances. Le cas de la firme-réseau», *Revue d'économie politique*, vol. 102, n° 4, juillet-août 1992, p. 563-592.
- Guilhon B., Gianfaldoni P., «Chaînes de compétences et réseaux», *Revue d'économie industrielle*, n° 51, 1er trimestre 1990, p. 97-112.
- Houssiaux J., «Quasi-intégration, croissance des firmes et structures industrielles», *Revue économique*, vol. n° 3, mai 1957.
- INSEE-L'Entreprise, *La France des entreprises*, n° spécial de l'entreprise, 1992.
- Jacot J.-H. (éd.), *Du fodisme au toyotisme? Les voies de la modernisation du système automobile en France et au Japon*, Paris, La Documentation française, 1990.
- Kern H., Schumann M., *La fin de la division du travail?*, Paris, Editions de la MSH, 1989.
- Lafay G., Herzog C., *Commerce international : la fin des avantages acquis*, Economica, Paris, 1989.
- Lassudrie-Duchêne B., «Décomposition internationale des processus productifs et autonomie nationale» in H. Bourguinat éd., *Internationalisation et autonomie de décision*, Economica, Paris, 1982, p. 45-56.
- Lassudrie-Duchêne B., Berthélémy J.-C., Bonnefoy F., *Importation et production nationale*, Economica, Paris, 1986.
- Leijonhufvud A., «Les coûts d'information et la division du travail», *Revue internationale des sciences sociales*, n° 120, mai 1989, p. 177-189.
- Loustau S., *L'industrie de la pharmacie et de la parapharmacie*, Crédoc, Paris, 1991.
- Lundvall B., *Product Innovation and User-Producer Interaction*, Aalborg University Press, Aalborg, 1985.
- Malone T.W., Yates J., Benjamin R.I., «Electronic markets and electronic hierarchies», *Communication of the ACM*, vol. 30, n° 6, June 1987, p. 484-497.
- Mathis J., Mazier J., Rivaud-Danset D., *La compétitivité industrielle*, IRES, Dunod, Paris, 1988.
- Michalet C.A., «Où est la notion d'économie mondiale?», in Humbert M., *Investissement international et dynamique de l'économie mondiale*, Economica, Paris, 1990, p. 16-34.
- Moati P., *La filière sport*, rapport pour le ministère de l'Industrie et le secrétariat d'Etat à la Jeunesse et aux Sports, Crédoc, Paris, 1989.
- Moati P., *Hétérogénéité des entreprises et commerce international*, Economica, 1992.

- Moati P., Mouhoud E.M., *Les transports et les télécommunications dans l'arbitrage entre modes d'organisation de la production*, Rapport Crédoc pour le Prédit, Paris, 1992.
- Mouhoud E.M., «Les stratégies de relocalisation des firmes multinationales», *Revue d'économie politique*, vol. 99, n° 1, janvier-février 1989.
- Mouhoud E.M., *Changement technique et division internationale du travail*, Economica, Paris, 1992.
- Nelson R.R., Winter S.G., *An Evolutionary Theory of Economic Change*, The Belknap Press of Harvard University Press, Cambridge Mass., 1982.
- Pelikan Pavel, «Evolution, Economic Competence, and the Market for Corporate Control», *Journal of Economic Behavior and Organization*, vol. 12, 1989, p. 279-303.
- Perrin J.-C., «Organisation industrielle : la composante territoriale», *Revue d'économie industrielle*, n° 51, 1er trimestre 1990, p. 276-303.
- Pyke F., Becattini G., Sengenberger W. (eds), *Industrial Districts and Inter-firm Co-operation in Italy*, ILO, Genève, 1990.
- Peyrache V., «L'innovation dans les districts industriels», *Problèmes économiques*, n° 2262, 12 février 1992, p. 25-30.
- Piore M.J., «Nouvelles remarques sur les trajectoires technologiques», in Boyer R., Chavance B., Godard O. (ed.), «*Les figures de l'irréversibilité en économie*», éditions de l'Ecole des Hautes Etudes en sciences sociales, Paris, 1991, p. 137-150.
- Piore M.J., Sabel C., *The Second Industrial Divide*, Basic Books Inc., New York, 1984; traduction française : *Les chemins de la prospérité : de la production de masse à la spécialisation souple*, Hachette, Paris, 1989.
- Reich R.B., *The Work of Nations. Preparing Ourselves for 21st Century Capitalism*, Simon & Schuster, Londres, 1991.
- Storper M., «The Limits to Globalization : Technology Districts and International Trade», mimeo, March 1992.
- Tertre C. (du), *Technologie, flexibilité, emploi : une approche sectorielle du post-taylorisme*, L'Harmattan, Paris, 1989.
- Toffler A., *Powershift : Knowledge, Wealth and Violence at the Edge of the 21st Century*, Bantam Books ed., New York, 1990; traduction française : *Les nouveaux pouvoirs : Savoir, richesse et violence à la veille du XXIe siècle*, Fayard, Paris, 1991.
- Williamson O.E., «Transaction Cost Economics : The Governance of Contractual Relations», *Journal of Law and Economics*, vol. 22, n° 2, October 1979, p. 233-261.
- Witkowski N., «Présentation», in *L'état des sciences et des techniques*, La Découverte, Paris, 1991.

La standardisation technologique en agriculture : formes, origine et perspectives à partir du cas français

Sylvie Bonny

INRA, Economie et Sociologie rurales
Grignon

Ce texte analyse le processus, les formes, l'origine et les perspectives de la standardisation technologique en agriculture, particulièrement en ce qui concerne les produits et les procédés. On examine les voies par lesquelles celle-ci s'est imposée en agriculture (raisons techniques, économiques, rôle de la formation, exigences de l'aval, mesures réglementaires, etc.) et le cas particulier des standards de qualité qui correspondent à une diversification. Actuellement le modèle technique agricole est en crise et on pourrait observer demain un modèle plus diversifié et flexible et une agriculture plurielle et multifonctionnelle. La dynamique des relations standardisation-diversification-innovation qui présente des traits spécifiques en agriculture est également abordée et on souligne la dualité de la standardisation.

This paper analyses the forms and origin of technological standardization in agriculture, the processes involved and its future prospects, with particular emphasis on products and processes. The paper examines how and why technological standardization has arisen in agriculture (technical reasons, economics, role of training, downstream requirements, regulations, etc.) and the particular case of quality standards which correspond to a diversification. At the present time, the technical

model of agriculture is in crisis; in future, there may be a more diversified and flexible model and a plural and multifunctional agriculture. The dynamics of relations between standardization, diversification and innovation, which have some distinctive features in agriculture are also discussed and the duality of standardization is stressed.

INTRODUCTION

Les processus de standardisation jouent un rôle important en économie industrielle comme l'ont montré plusieurs auteurs (cf. par exemple David, 1987; Ferné, 1992; Foray, 1990 et 1993; OTA, 1992). Qu'en est-il en économie agricole? Il y a en agriculture, même si l'on considère seulement l'exemple français, une grande diversité dans les façons de produire en raison du caractère artisanal de la production, de la variabilité inhérente aux phénomènes biologiques, de la diversité des conditions pédoclimatiques et des dotations en facteurs de production, mais aussi du fait qu'une partie des modes de conduite des cultures et des élevages provient des modes de savoir-faire paysans, même si une part importante est désormais « imposée » par l'appareil de formation et d'encadrement et par les industries agro-alimentaires d'aval qui achètent les produits agricoles. Il y a également une grande diversité dans les produits obtenus en raison de leur origine biologique ⁽¹⁾, de l'existence de nombreux segments de marché et de différences de goût chez les consommateurs conduisant à la multiplication des produits proposés et à des stratégies de différenciation des aliments. Toutefois, malgré cette diversité, il existe également une forte standardisation dans les façons de produire et dans les produits obtenus, c'est-à-dire **une standardisation technologique**. Nous allons chercher à étudier les formes qu'elle prend, son origine et ses perspectives, notamment dans ses relations avec la différenciation, la diversification et l'innovation à partir du cas de l'agriculture française actuelle. Notons que nous traiterons ici

⁽¹⁾ Sauf dans le cas de la multiplication végétative et du clonage qui existent pour certains végétaux, les produits biologiques diffèrent entre eux en raison notamment du phénomène de recombinaison génétique qui s'opère à chaque fécondation et qui redistribue les différents gènes. A ces variations génotypiques s'ajoutent des variations phénotypiques liées à l'interaction avec l'environnement : le climat, la nature des sols, etc., différent d'un lieu à l'autre. Malgré cette diversité des produits biologiques, il existe bien sûr des caractéristiques moyennes liées à l'espèce, la variété, etc., comme le montre une courbe de Gauss.

essentiellement de la production agricole, et non de l'industrie agro-alimentaire où les processus mis en œuvre sont différents (car de nature industrielle et non plus agricole).

La standardisation se caractérise par sa dualité : pour beaucoup de produits industriels, comme elle permet une compatibilité des équipements entre eux, l'interchangeabilité des pièces, une réduction d'incertitude, etc., elle est recherchée; pour d'autres produits, notamment les aliments, les vêtements et divers biens de consommation courante, elle est au contraire souvent rejetée par le consommateur car considérée comme synonyme d'uniformité appauvrissante. Toutefois les standards de qualité sont une source de diversification des denrées et les normes de salubrité ou de définition des produits évitent au consommateur d'être berné dans un domaine où l'asymétrie d'information est forte; ils sont un élément important de coordination industrielle et marchande. Ainsi donc la standardisation est un processus complexe qui nécessite d'être étudié plus en détail en agriculture où il revêt des formes et des caractéristiques particulières.

I. STANDARDISATION ET MODÈLE TECHNIQUE DE PRODUCTION

Un **standard** est un ensemble de spécifications techniques auquel adhère un producteur, soit tacitement, soit au terme d'un processus formel de standardisation volontaire, soit dans le respect d'une décision de l'autorité publique (David, Greenstein, 1990). P.A. David (1987) a distingué les standards de référence ou de définition, les standards établissant des attributs minimums admissibles et les standards de compatibilité et d'interface. Ce sont essentiellement les deux premiers types qui concernent l'agriculture. Mais le terme standard renvoie souvent trop étroitement à une norme technique de fabrication, dans ce texte nous l'emploierons aussi dans le sens plus large de **modèle technique** ou de **modèle de production**. Il s'agit d'une représentation simplifiée, théorique – idéal-typique au sens webérien – des principales caractéristiques techniques de la production à une période donnée, la notion de modèle technique restant davantage au niveau des techniques mises en œuvre que celle de modèle de production qui prend aussi en considération les aspects socio-économiques, voire politico-économiques de la production. Malgré la diversité de l'agriculture française qu'ont soulignée de nombreux auteurs, on peut être autorisé

à parler d'un modèle technique ou d'un modèle de production. Il s'agit en fait plus exactement du **modèle dominant** à une période donnée, dominant non pas parce qu'il est le plus largement répandu mais parce qu'il est la **référence-type** qui définit les objectifs à viser et les moyens pour les atteindre, c'est-à-dire les standards de conception technique.

Cette notion de modèle technique est de ce fait assez proche de celle de **paradigme technologique** introduite par Dosi (1982, 1988) et celle de modèle de production du concept de **paradigme technico-économique** développé par C. Perez et C. Freeman (Perez, 1983; Freeman in Salomon, Schméder, 1986; Freeman in Heertje, 1988; Freeman, Perez in Dosi *et al.*, 1988). «Un paradigme technologique définit, dans leur contexte, les besoins que l'on estime devoir satisfaire, les principes scientifiques utilisés pour la tâche, la technologie qu'il importe d'utiliser. En d'autres termes, un paradigme technologique peut être défini comme un «modèle» [pattern] de solution à des problèmes technico-économiques sélectionnés, basé sur des principes fortement sélectionnés dérivant des sciences de la nature, associés à des règles spécifiques visant à acquérir une connaissance nouvelle et à la préserver, quand cela est possible, d'une diffusion rapide aux concurrents» (Dosi, 1988, p. 1127). Ainsi donc les notions de standard de production (dans une acception large), de paradigme technologique et de modèle technique sont proches.

II. LES FORMES DE LA STANDARDISATION TECHNOLOGIQUE EN AGRICULTURE

La standardisation peut concerner soit les **processus de fabrication**, soit les **produits** eux-mêmes. Dans l'industrie la standardisation proprement dite est apparue en tant que technique de production au XIX^e siècle aux USA dans la fabrication des armements. C'est le «système américain» de manufactures. «Dans un certain nombre d'industries, l'enjeu fondamental du passage du travail manuel à la mécanisation a été la recherche d'une parfaite uniformité des différentes pièces produites. L'identité des pièces devait permettre de remplacer les longues séances d'ajustage caractéristiques du système antérieur par le montage beaucoup plus rapide de pièces interchangeables» (Boyer, Schméder, 1990). La standardisation consiste à fixer des normes de dimensions et de qualité rigoureusement définies pour les produits ou leurs composantes. Elle permet le remplacement des pièces et facilite

aussi beaucoup la division du travail et la production en grandes séries à bas prix. Elle conduit ainsi *in fine* à la production de masse et à la consommation de masse qui sont l'un des éléments caractéristiques du fordisme.

En agriculture elle est apparue plus tardivement. Dans l'agriculture française, elle s'est surtout diffusée après la deuxième guerre mondiale, avec ce que l'on a appelé la deuxième révolution agricole, qui s'est traduite sur le plan technique par :

- l'utilisation d'un nouveau matériel génétique amélioré et de ce fait plus standardisé : par exemple, les variétés locales de pommes et les races animales régionales ont laissé place à la prédominance de quelques variétés (golden, ...) et de quelques races;

- la motorisation, la mécanisation : le tracteur, la moissonneuse-batteuse ont remplacé le dur labeur des animaux et des hommes et induit une certaine standardisation dans les façons culturales. La récolte mécanique par exemple exige des produits mûrs au même moment, ce qui exclut les mélanges variétaux et plus encore les associations de cultures. Elle a entraîné également une standardisation spatiale : champs rectangulaires, suppression des haies, talus, chemins creux. Ce processus a abouti aux opérations foncières de remembrement : cela consiste à mettre fin au morcellement excessif de la propriété foncière par un système d'échanges obligatoires de parcelles dans le but de regrouper celles d'une même exploitation; le remembrement s'est accompagné d'opérations connexes d'arrachage des haies et d'arasement de talus. Cela aboutit *in fine* à une plus grande standardisation des paysages agricoles, avec par exemple la régression du bocage;

- la chimisation : les engrais chimiques ont supplanté le fumier pour la fertilisation, et la lutte contre les maladies et les parasites s'est faite au moyen de pesticides chimiques. On peut parler d'une certaine standardisation car les agriculteurs se réfèrent à des recommandations et règles pour leur emploi; toutefois l'épandage d'un engrais ou d'un pesticide dépend de la culture, du sol, du climat, du précédent cultural, de l'état physiologique de la plante, de l'expérience de l'agriculteur, de son information, etc., donc il subsiste une grande diversité. Dans les élevages de porcs et de volailles et dans une mesure bien moindre en élevage bovin, les producteurs ont utilisé de plus en plus d'aliments fabriqués par l'industrie et non plus les céréales ou sous-produits de la ferme. Des normes d'origine scientifique fondées sur les besoins des animaux à leurs différentes périodes de croissance se sont imposées, et cela d'autant plus que certaines de ces productions se sont faites sous contrat entre une firme d'aliments et l'éleveur : c'est le processus

d'intégration observé dans les années 60 où l'agriculteur doit appliquer les prescriptions de l'industriel pour la conduite de l'élevage.

Cette **standardisation dans les processus de production** a entraîné une **standardisation dans les produits obtenus**, mais cette dernière provient aussi des **exigences et des normes imposées** par la réglementation sanitaire et par l'aval. En effet, diverses réglementations visent à protéger les consommateurs de la fraude sur la nature du produit et de la mauvaise qualité hygiénique des denrées comestibles. Par ailleurs les industries agro-alimentaires qui achètent les produits agricoles pour les transformer demandent une certaine composition en glucides, lipides et protéines, une bonne conformation pour les animaux, une apparence pour les végétaux (les fruits par exemple sont calibrés), une qualité bactériologique, une teneur en matière sèche,... Compte tenu de la diversité des produits biologiques, il s'agit surtout de seuils. Pour orienter dans le sens souhaité, le paiement des produits agricoles se fait désormais le plus souvent à la qualité, selon la conformation ou la composition.

A cette standardisation des processus techniques de fabrication et des produits obtenus, s'ajoute une **standardisation d'ordre organisationnel et économique** ⁽²⁾. Le modèle technique de production agricole mis en place dans les décennies d'après-guerre est souvent qualifié de **productiviste** : l'objectif essentiel était de produire et aussi d'exporter (le solde des produits agro-alimentaires a été déficitaire jusqu'au milieu des années 70). Cela nécessitait d'augmenter la productivité de la terre et du travail, généralement fort basse jusqu'en 1945, et de diminuer les coûts de production. Cela a conduit à la spécialisation des exploitations (auparavant le système polyculture-élevage prédominait, les cultures étaient associées à l'élevage tandis qu'aujourd'hui les exploitations sont spécialisées dans un nombre très restreint de productions) et par l'intensification : on a cherché à produire davantage par hectare et par animal pour accroître les volumes et réduire les coûts de production unitaires. On a également développé les échanges entre l'agriculture et les autres secteurs, ainsi qu'avec les autres pays : l'agriculteur a utilisé de plus en plus d'intrants industriels, a transformé de moins en moins lui-même ses produits à la ferme et a produit de plus en plus pour le marché et non plus pour l'autoconsommation (en 1960, un agriculteur nourrissait environ neuf personnes; en 1992, il en nourrit près de cinquante). Comme dans les autres secteurs, il y a eu un fort mouvement de substitution

⁽²⁾ De même que Schumpeter a distingué les innovations de produit, de procédé et d'organisation, on peut distinguer une standardisation en ces trois domaines.

du capital au travail et de concentration des producteurs. Toutes ces transformations ont entraîné une standardisation des procédés et des produits, en raison notamment de la réduction du caractère artisanal et local de la production agricole.

Les pratiques agricoles se sont «rationalisées» : en effet d'autres formes de rationalité que la rationalité micro-économique standard du producteur ont longtemps existé dans les campagnes; mais dans les années 60 le paysan a fait place à l'agriculteur et au producteur agricole comme l'a analysé H. Mendras dans son ouvrage *La fin des paysans* (Mendras, 1967 et 1991). Les modes de savoir-faire et tours de main traditionnels assez diversifiés ont cédé le pas à des modes de conduite technicisés plus standardisés tant sur le plan des façons de produire que de la gestion des exploitations. Cela a entraîné une transformation profonde des rapports aux ressources comme le souligne L. Thévenot (1989) en reprenant les analyses de P. Boisard et M.T. Letablier sur la production de camembert. La citation suivante de ces auteurs reprise par L. Thévenot permet par exemple de comparer le rapport domestique au rapport industriel dans le cas de la production laitière :

rapport domestique : «Louis connaît personnellement les vaches de son troupeau de normandes, qu'il nourrit traditionnellement du fourrage issu du domaine qu'il a hérité de son père, ce qui est la garantie d'un bon lait. Il obéit à la coutume ancestrale qui veut que les vaches donnent comme on leur donne. Il lui suffit d'un coup d'œil pour connaître que quelque chose ne va pas avec la plus âgée d'entre elles, la Noiraude, qui a son caractère. Il sait bien alors, pour l'avoir appris de son père lorsqu'il était jeune, ce qu'il a à faire.»

rapport industriel : «Le bétail, sélectionné pour son matériel génétique, est identifié au moyen de son collier électronique. Les principaux paramètres (poids, température, etc.) sont saisis par des capteurs électroniques. La production laitière est enregistrée automatiquement. Ces informations sur la productivité et l'état sanitaire du bétail sont exploitées par l'unité centrale d'un ordinateur à partir d'un programme de calcul préalablement introduit. En fonction des résultats, sont commandés des actionneurs électromécaniques assurant la distribution ajustée d'aliments concentrés dont la composition en protéines est contrôlée.»

Dans les années 70, cette évolution a fait naître chez les économistes un débat sur l'industrialisation de l'agriculture (cf. par exemple CNEEJA, 1971) et sur les formes particulières que prenaient les rapports capitalistes de production dans ce secteur. En effet le fordisme et le taylorisme qui se traduisent par une standardisation des procédés et des produits ont pris en agriculture des formes spécifiques (le salariat ne s'y est pas développé, au contraire). Certains auteurs ont parlé d'une soumission formelle au capitalisme : «la quasi-totalité des producteurs directs de l'agriculture est soumise formellement au capital. En effet, la

soumission formelle du travail au capital touche, bien qu'à des degrés divers, 90% des producteurs de la sphère de production agricole» (Faure, 1974). Pour d'autres auteurs la petite production marchande se maintient en agriculture car elle est plus efficace compte tenu des spécificités de la production agricole (Servolin, 1972). Quoi qu'il en soit, les processus de standardisation technologique sont bien à l'œuvre en agriculture, mais avec des modalités particulières et une variété notable persiste, notamment en raison des phénomènes biologiques et agro-climatiques en jeu : comme disent les agriculteurs, il n'y a pas deux terres, ni deux vaches pareilles et le climat est toujours aléatoire; toutefois certaines productions se sont beaucoup artificialisées et sont davantage standardisées : les cultures sous serre, l'élevage de porcs et de volailles notamment.

III. LES ORIGINES DE CE PROCESSUS DE STANDARDISATION ET DE SA DIFFUSION

Le processus de standardisation que nous venons de décrire s'est imposé par plusieurs voies :

Pour des raisons techniques

Par exemple l'agriculteur n'a plus trouvé sur le marché que de nouvelles semences; avec l'insémination artificielle largement développée dans l'élevage il n'a eu le choix que pour des races dominantes; avec la motorisation il a dû agrandir et équarrir ses parcelles.

Pour des raisons économiques

Le modèle productiviste permet généralement une diminution des coûts de production et une meilleure rentabilité (du moins jusqu'à une époque assez récente). A un niveau plus macro-économique il correspondait aux fonctions attribuées à l'agriculture dans les trois décennies d'après-guerre : produire une nourriture abondante à bas prix, libérer la main-d'œuvre agricole. Cela aboutit à la production de masse fordiste de «commodités» comme on les appelle parfois en transposant le terme anglais (Chalmin, El Alaoui, 1990). La standardisation

technologique permet aussi, en agriculture comme dans les autres secteurs, une économie des coûts de transaction en particulier entre producteurs et firmes d'aval ou consommateurs du fait de l'homogénéité et de la plus grande régularité du produit⁽³⁾. On retrouve ainsi en agriculture les effets des standards – économies d'échelle dans la production, réduction des coûts de transaction notamment – signalés dans les autres secteurs (Kindleberger, 1983; David, 1987; Foray, 1990).

Par l'appareil d'encadrement et de formation

Les agriculteurs sont conseillés par des techniciens qui peuvent appartenir à des chambres d'agriculture, à des groupements de vulgarisation, ou aux organismes (négociants, coopératives) qui commercialisent les produits industriels utilisés en agriculture. L'école joue également un rôle important puisqu'un nombre croissant d'agriculteurs a suivi une formation technique agricole, celle-ci étant même nécessaire pour recevoir certaines aides, alors que naguère l'agriculteur apprenait son métier de ses parents. Or l'école et l'appareil d'encadrement diffusent un modèle relativement normalisé. Dès 1955, l'historien M. Augé-Laribé écrivait dans son ouvrage *La révolution agricole* : « on est passé de l'homme de la houe à l'agent des laboratoires, l'exécutant de règles scientifiques ». L'agro-industrie d'amont vulgarise également certaines recommandations pour l'emploi des engrais, des pesticides et des aliments du bétail.

Par l'aval

Les industries agro-alimentaires qui achètent les produits agricoles exigent qu'ils répondent à certaines normes. Il existe même comme on l'a déjà noté des contrats d'intégration où la firme donne à l'agriculteur un cahier des charges qui indique les règles de production qu'il doit suivre. Cela concerne notamment les légumes destinés à la conserverie, la production de veaux en batterie ou de volailles hors sol⁽⁴⁾. On se rapproche ici quelque peu des standards de compatibilité fréquents dans l'industrie car ces cahiers des charges visent notamment à obtenir

(3) A l'opposé pour l'achat de produits non standards comme un cheval de course, les coûts de transaction sont élevés car l'acheteur doit s'informer des qualités du produit, le tester, etc., l'asymétrie d'information et l'incertitude étant fortes.

(4) Dans l'élevage hors sol on utilise pour nourrir les animaux presque uniquement des aliments achetés à l'extérieur de l'exploitation et non les céréales et fourrages produits sur l'exploitation.

des produits adaptés à une certaine transformation et homogènes. La standardisation permet également de réduire les coûts de transaction et d'avoir un approvisionnement plus constant quant à sa composition.

Par des mesures réglementaires

Ces dernières peuvent viser par exemple la qualité bactériologique, l'absence de produits toxiques (qui peuvent être d'origine naturelle), la limitation des résidus de pesticides, mais aussi la composition du produit. Ces réglementations sont importantes compte tenu des risques en jeu : les intoxications alimentaires peuvent être extrêmement graves, voire mortelles. Il s'agit de veiller à la santé publique et de protéger le consommateur de la fraude sur la nature du produit vendu. On déplorait en effet autrefois le mouillage du lait, l'adjonction de plâtre et autres substances au pain, etc. Dès le Moyen Âge des ordonnances condamnaient de tels abus. Des lois sont passées en ce domaine à partir du XIX^e siècle. Le code napoléonien de 1810 réprime les « tromperies, falsifications et altérations volontaires de marchandises ». La loi-cadre de 1905 institue en France le service de la répression des fraudes. De même aux Etats-Unis, « one of the first product areas to benefit from standards was that of food. Responding to scandals in the meat packing industry, Congress passed the Pure Food and Drug Act of 1906. This legislation not only protected against misbranding and food adulteration; it also standardized containers of marketing fruits and vegetables, thereby eliminating false measurements and deceptive shapes. Later the Department of Agriculture, continuing the standards program initiated during the First World War, developed standards for fruits, vegetables, peanuts, honey, butter, cheese, eggs, and meat, and established inspection stations at a number of key distribution centers » (OTA, 1992). De telles mesures devraient augmenter avec le souci croissant de la qualité chez les consommateurs et la sensibilisation accrue à la présence de pesticides ou d'hormones ⁽⁵⁾.

En raison du développement des exportations

Pour qu'un produit puisse être exporté, il doit être défini et répondre à certaines normes. Le standard devient alors parfois une barrière

⁽⁵⁾ Il faut souligner que l'établissement de normes en matière de teneur en résidus admissibles dans les aliments est un processus complexe résultant de négociations difficiles où interviennent scientifiques, juristes, politiques et industriels : il s'agit souvent de domaines controversés caractérisés par l'incertitude comme l'ont analysé certains sociologues (cf. par exemple Rip, Groenewegen, 1989).

protectionniste non tarifaire comme on a pu le voir pour les fromages, la bière ou la viande : des pays refusent l'importation de certains produits en arguant qu'ils ne correspondent pas à leurs normes sanitaires. Lors de l'intégration européenne, l'harmonisation des normes et des critères nationaux d'homologation se révéla très difficile. Dans un premier temps la Communauté s'efforça de bâtir des directives ou des règlements très complets sur la composition de certains produits comme la confiture ou le chocolat; mais cela se révéla très long et très complexe. De ce fait, après avoir eu de grandes ambitions en la matière, la Commission a adopté une approche plus minimaliste : tout produit fabriqué et commercialisé légalement dans un Etat-membre peut circuler librement à l'intérieur de la Communauté sauf s'il ne répond pas à certaines nécessités impérieuses, notamment dans la santé publique d'un Etat-membre. Dès lors l'harmonisation passe par la reconnaissance mutuelle des réglementations respectives de chacun des Etats-membres.

Pour des raisons idéologiques

Le modèle de production dominant s'est parfois imposé comme «le progrès inéluctable»... C'était le «bon» modèle qu'il fallait suivre, sinon on était considéré comme arriéré, retardataire; une norme sociale s'est ainsi imposée : «ces modèles d'intensification sont ainsi devenus des symboles de la modernité tels que les agriculteurs qui ne les ont pas adoptés ont souvent intériorisé eux-mêmes leur «infériorité» ou leur exclusion de l'agriculture rentable, de l'agriculture d'avenir» (Alphandéry *et al.*, 1980).

Malgré certains de ses avantages que nous avons notés, la standardisation des aliments est souvent rejetée par les consommateurs qui déplorent l'uniformisation de la nourriture. Les étals des supermarchés donnent cependant plutôt une impression de diversité et de multiplicité des denrées proposées (*cf.* paragraphe suivant). L'insatisfaction des consommateurs en ce domaine provient notamment de la distance croissante à la fois physique et symbolique qui s'est établie entre eux et l'origine des aliments : ces derniers deviennent des produits non identifiés, sans passé ni origine connus (Fischler, 1990; Valceschini, 1993). D'où le succès de la vente directe du producteur au consommateur, notamment de la vente à la ferme quand elle est possible, car elle rend le produit identifiable et même unique. De leur côté les généticiens et les écologistes déplorent souvent la diminution de la diversité génétique, avec la réduction du nombre de variétés végétales cultivées et de races

animales utilisées (*cf.* par exemple Chauvet, Olivier, 1993); tous les biologistes s'accordent en effet sur la nécessité de conserver la biodiversité. D'autres personnes regrettent la standardisation des paysages et des pratiques agricoles. Des protestations se sont aussi élevées quand la définition de normes hygiéniques rigoureuses dans la Communauté a failli faire disparaître les trente-deux fromages français au lait cru. Ces divers exemples montrent bien la dualité de la standardisation, tantôt recherchée, tantôt honnie.

IV. LES STANDARDS DE QUALITÉ, FORME DE DIFFÉRENCIATION DES PRODUITS

Jusqu'ici nous avons surtout considéré la standardisation en tant que «production de modèles standards fabriqués en série» (définition du Petit Robert) qui correspond à une homogénéisation des pratiques et produits agricoles. Mais il existe une **production de standards** au sens de «spécifications techniques» **qui participe au contraire à un mouvement de diversification et différenciation**. Ce sont les réglementations concernant notamment **les appellations contrôlées** (vins, fromages) et **les labels** (poulet fermier, agriculture biologique) : le **standard de qualité** est alors demandé par des groupes de producteurs pour protéger une image de marque et assurer une meilleure valorisation de leur marchandise; il concerne notamment le mode d'obtention des produits. Il correspond aux aspirations des consommateurs qui recherchent de plus en plus la diversité et l'authenticité : produits du terroir, produits typés, produits fermiers, etc. Ce n'est plus seulement aujourd'hui pour se différencier comme dans le cas de la consommation de type ostentatoire analysée au début du siècle par T. Veblen aux Etats-Unis, mais aussi par besoin de réassurance : on recherche désormais souvent les produits traditionnels, les produits verts fleurant bon la campagne d'autrefois, au moins pour les repas festifs (pour les autres repas les consommateurs visent assez souvent des aliments vite prêts et peu chers). Il existe de plus du fait de la diversité des consommateurs et de la stratégie des firmes différents segments de marché : produits bon marché, aliments diététiques ou allégés, denrées biologiques, produits du terroir, aliments haut de gamme ou gastronomiques, etc. B. Sylvander (1992) indique ainsi qu'on estime à environ 400 000 références le nombre actuel de variantes de produits alimentaires (y compris les variantes de marque).

La production de standards de qualité, en particulier par la spécification de l'origine est un mouvement ancien. En France au Moyen Age l'origine des produits était contrôlée de façon rigoureuse par chaque corporation ou autorité locale. Mais à la Révolution, la loi du 24 juin 1791, en posant le principe de la liberté du commerce et de la suppression des corporations, a eu pour effet de mettre fin à ce type de garantie. Ainsi à partir du début du XIX^e siècle on a fabriqué des vins qualifiés de « champagne » dans presque toutes les régions viticoles de France et du monde. Mais, dans un arrêt du 9 avril 1894, la Cour de cassation déclara que le « mot champagne ne désigne pas un procédé de fabrication de vin mousseux en général, mais un vin mousseux spécial récolté et fabriqué dans l'ancienne province de Champagne » (ministère de l'Agriculture, 1992). Dès le début du XIX^e siècle des lois ont été prises pour assurer une certaine protection de l'indication de l'origine; ainsi la loi de 1824 institue un délit pour de fausses indications de provenance. Ultérieurement la loi du 6 mai 1919 définira les appellations d'origine, et le décret-loi du 30 juillet 1935 créera les appellations d'origine contrôlée (AOC) pour les vins et eaux-de-vie, étendues en 1955 aux fromages.

Actuellement on peut répartir les certifications de qualité en différents types :

– **la qualité identifiée par l'origine géographique. L'appellation d'origine contrôlée (AOC)** qui lie étroitement les caractères d'un produit au terroir dont il est originaire en est l'exemple le mieux connu. Aujourd'hui, les AOC représentent 45% de la production nationale des vins. Elles concernent aussi 32 fromages, soit plus de 15% de la production nationale, 40 000 producteurs de lait (17% de l'ensemble), 2 000 producteurs fermiers et près de 600 entreprises. Relèvent également du régime des AOC divers autres produits animaux et végétaux. Il existe aussi une mention « appellation provenance montagne » qui s'applique à des produits issus des massifs réglementairement identifiés comme tels.

– **le produit répondant à la notion de qualité supérieure.** C'est le **label national** (label rouge) accordé par un organisme certificateur en fonction d'un cahier des charges : cela concerne notamment les volailles (15% de la production avicole nationale), certaines viandes, produits laitiers, fruits et légumes ainsi que des produits festifs; 30 000 agriculteurs sont impliqués.

– **le certificat de conformité** vise à garantir que certains critères de fabrication ont été respectés : œufs provenant de volailles élevées en plein air, veau sous la mère, poulet de chair, œufs datés, etc.

– **la mention agriculture biologique** : depuis quelques années les produits issus de l'agriculture n'utilisant pas de produits chimiques de synthèse sont homologués par le ministère de l'Agriculture et reconnaissables par un logo.

Ces standards de qualité sont recherchés par les producteurs. Ainsi par exemple dès la première moitié du XVIII^e siècle les fermiers américains «réalisèrent qu'en graduant et classifiant leurs produits ils pouvaient créer des canaux de distribution séparés et accroître leurs profits. Aussi, quand ils s'éloignèrent vers l'Ouest, ils labellisèrent leurs produits selon la région d'origine, et les grossistes utilisèrent ces noms – beurre de Goschen, farine de Genessee et fromage de Herkiner – comme marques de qualité» (OTA, 1992). Le label ou l'appellation contrôlée sont intéressants pour les agriculteurs car ils permettent un meilleur prix des produits. On observe ainsi que depuis trente ans le prix des vins de qualité (AOC ou VDQS (vin délimité de qualité supérieure)) a augmenté en francs constants alors que celui des vins de consommation courante ou du blé chutait : de 1962 à 1992, en monnaie constante, le prix des vins de qualité a crû de 9,6% tandis que celui des vins courants a baissé de 33% et celui de l'ensemble des produits agricoles a chuté de 44%.

Au consommateur le label apporte une forme d'assurance. La réglementation et le standard relèvent alors de l'économie de la confiance étudiée par L. Thévenot (1989). Ces standards de qualité sont particulièrement justifiés en cas d'asymétrie d'information entre le vendeur et l'acheteur (Foray, 1993). Qu'est-ce qui prouverait sinon au consommateur, qui n'a pas individuellement le moyen de faire des analyses, que les légumes biologiques qu'il achète plus cher ont été effectivement produits sans engrais ni pesticides chimiques? L'indication de provenance rassure aussi le consommateur à un niveau plus symbolique : le produit sort de l'anonymat. Une importante littérature a déjà été consacrée à ces conventions de qualité (*cf.* par exemple Eymard-Duvernay, 1989) et divers colloques ont déjà été organisés sur ce sujet dans l'agro-alimentaire (*cf.* par exemple Nicolas, Valceschini, 1993).

Ces appellations et labels font l'objet de convoitise et de contrefaçons et d'aucuns cherchent à s'approprier ce type de rente («champagne» fabriqué en Californie, voire en Angleterre; camembert produit en Hollande, au Japon, etc.). La CEE a adopté en juillet 1992 des textes en ce domaine, qui viennent d'être transposés fin 1993 dans le droit français. Les produits traditionnels devraient être protégés par des «attestations de spécificité» liées à la recette ou au mode de

production; c'est le cas par exemple des rillettes du Mans ou du veau élevé sous la mère. Les étrangers pourront en produire à condition qu'ils respectent scrupuleusement la recette. D'autres produits dont les matières premières, la fabrication et l'élaboration sont liées au terroir relèveront de l'AOP, appellation d'origine protégée. Par exemple les fromages seront protégés au niveau communautaire par l'AOP : d'autres pays pourront fabriquer du camembert (terme générique), mais pas de camembert de Normandie, ni du cantal ou du roquefort. Tous les produits ayant droit à une AOC en France bénéficieront d'une procédure simplifiée pour une reconnaissance rapide en AOP. L'indication géographique protégée (IGP) concernera quant à elle les produits portant le nom d'une région et dotés de qualité ou d'une réputation qui peut être attribuée au milieu géographique (rillettes du Mans, poulets de Loué) (contrairement à AOC les matières premières ne sont pas forcément locales). Comme le notent P. Boisard et M.-T. Letablier (1987) pour la fabrication du camembert, la réglementation devient alors en quelque sorte le gardien de la tradition. On observe aussi une évolution de la notion de qualité dans la CEE, le passage de définitions par la négative à une approche positive : auparavant on spécifiait surtout l'insuffisance de qualité en fixant des dimensions, des teneurs, des taux, etc., minimums ou maximums; désormais on définit aussi positivement la qualité des produits. «De ce point de vue, si l'action de la PAC (politique agricole commune) des premières années a pu donner l'impression d'un rouleau compresseur des originalités, celle des années à venir devrait au contraire permettre le plein épanouissement de ces singularités, parallèlement à celle des productions de masse» (Tabary, 1993).

V. LA CRISE DU MODÈLE TECHNIQUE PRODUCTIVISTE. VERS UN NOUVEAU STANDARD OU VERS UNE DIVERSIFICATION?

Le modèle technique productiviste est actuellement en crise depuis bientôt deux décennies mais celle-ci s'est aggravée ces dernières années. Les facteurs de remise en cause sont notamment :

– le changement du contexte économique avec la croissance du prix de certains intrants dans les années 1974 et actuellement la saturation

des débouchés solvables et le développement des excédents conduisant à des conflits commerciaux sur les marchés internationaux, mais aussi avec les modifications de la demande adressée à l'agriculture;

– les limites du modèle : son fort besoin en capital, les atteintes à l'environnement, la désertification d'une partie du territoire (la production agricole s'étant concentrée dans quelques régions), le coût de soutien des marchés, la stagnation des revenus des agriculteurs, etc.

Un nouveau modèle technique peut-il apparaître, ou du moins comment l'ancien modèle technique peut-il évoluer? Notons tout d'abord qu'il existe un modèle de production qui se veut alternatif, c'est l'agriculture biologique qui n'utilise pas de produits chimiques de synthèse. Mais il est peu probable qu'elle se généralise (à l'heure actuelle elle concerne environ 3 000 à 4 000 agriculteurs seulement en France) car la suppression totale d'engrais et de pesticides chimiques serait difficile à étendre largement du jour au lendemain, en particulier parce qu'elle induirait une chute et une forte incertitude sur les rendements et la production, et donc une hausse des prix que tous les consommateurs n'accepteraient pas.

Par contre le modèle technique dominant va sans doute se transformer (Bonny, 1993a) :

– **en raison de l'évolution technologique** : les applications des biotechnologies et des nouvelles technologies de l'information devraient permettre la substitution progressive d'un nouveau modèle reposant davantage sur la maîtrise du vivant et sur l'information et moins sur les produits chimiques et l'énergie fossile (Bonny, Daucé, 1989; Bonny, 1991);

– **en raison de l'évolution de la demande** : les exigences adressées aujourd'hui à l'agriculture sont multiples :

– la maîtrise de l'offre;

– la qualité, et pas seulement la quantité;

– une agriculture propre, polluant le moins possible (évitant en particulier les fuites de nitrates et de pesticides dans les eaux); diverses réglementations commencent peu à peu à s'appliquer en agriculture en ce domaine;

– une meilleure gestion de l'espace et du territoire. Il faut éviter la concentration de la production agricole en quelques zones se traduisant par la désertification d'autres régions; en effet dans ces dernières, l'espace rural n'est plus entretenu et se dégrade (risque d'incendie, embroussaillage, disparition des sentiers, etc.);

– des techniques socialement acceptées : certaines pratiques agricoles sont mal considérées par le reste de la population : élevage hors

sol, utilisation de pesticides, épandage de lisier, etc. Les façons de produire devraient donc évoluer et l'éthique jouera un rôle croissant;

– la production de services : activités touristiques, services de proximité, entretien du territoire et de la nature, production d'aménités environnementales, etc., devraient se développer dans les prochaines années malgré certains obstacles.

– le maintien d'un certain niveau de population agricole compte tenu de la progression du chômage et des risques de désertification d'une partie du territoire;

– mais l'exigence de compétitivité demeure, ainsi que celle de la production de produits de base à bas prix (les commodités);

– **en raison de l'évolution des réglementations internationales** : depuis le 21 mai 1992 a été adoptée une réforme de la politique agricole commune (PAC) qui entraîne notamment une forte diminution des prix agricoles, compensée par des aides; la nouvelle PAC incite à une certaine désintensification, du moins en certaines régions et pour certaines productions (Blanchet, Carles, 1993; Bonny, Carles, 1993). Les accords du GATT devraient entraîner de leur côté une limitation des exportations. Mais il subsiste de fortes incertitudes, notamment sur la possibilité de durer de la nouvelle PAC, sur l'évolution de la demande sur les marchés mondiaux, sur l'évolution des prix et des parités monétaires. Dans ce contexte d'incertitude un certain nombre d'observateurs insistent sur la nécessité de favoriser la **flexibilité** de l'outil de production agricole (CGP, 1993).

On peut donc penser qu'on observera une évolution du modèle technique dominant car les objectifs et le contexte ont changé depuis la période d'après-guerre qui a vu la naissance du modèle productiviste. **Va-t-il y avoir une nouvelle standardisation qui s'imposera ou bien une diversification?** De nombreux acteurs ont souligné la nécessité d'une agriculture diversifiée (Jollivet, 1988). Compte tenu de la multiplicité des objectifs assignés aujourd'hui à l'agriculture, de la diversité des demandes qu'elle reçoit et des possibilités techniques, on peut penser que **l'agriculture demain sera plurielle et multifonctionnelle; le modèle de production devrait être diversifié, adaptatif, flexible et imaginatif** :

– **multifonctionnel** : certains agriculteurs produiront des denrées de masse à bas prix pour l'alimentation, la végétalochimie ou l'énergie; d'autres des produits de haute qualité organoleptique mettant l'accent

sur le terroir d'origine ou la tradition; d'autres viseront des micro-marchés⁽⁶⁾, d'autres des services : accueil en milieu rural, tourisme, entretien de la nature et du territoire, etc. Un même agriculteur peut bien sûr mettre en œuvre simultanément ces divers types de produits mais la différenciation sera notamment régionale. Il ne faut pas voir dans ces diverses fonctions une entériorisation du dualisme : produire des services doit être considéré en agriculture comme aussi valorisant que produire des aliments;

– **diversifié** : si l'ensemble de l'agriculture devra être moins polluante, certaines formes d'agriculture seront intensives, d'autres extensives, d'autres particulièrement respectueuses de l'environnement (avec des contrats spécifiques pour les biotopes sensibles), quelques-unes seront biologiques. Sur l'ensemble du territoire on aura aussi bien de grandes exploitations sur de vastes surfaces que des moyennes, ou bien encore de petites exploitations à temps partiel.

– **adaptatif et flexible** : la demande adressée à l'agriculture est diverse et varie selon les différents segments de la population; de plus elle évolue selon les modes, les préoccupations diététiques, la médiatisation de certaines découvertes scientifiques. Ainsi on cherche depuis une quinzaine d'années à réduire la teneur en matières grasses, mais un certain « retour de bâton » commence à s'opérer en ce domaine. De même certaines préoccupations en matière d'environnement pourraient peut-être s'avérer versatiles. L'agriculture doit pouvoir s'adapter à cette évolution des différents segments de la demande. Elle doit aussi pouvoir s'ajuster à des modifications de la politique agricole commune et aux incertitudes sur les marchés mondiaux. Il faut encourager au maximum la production dans les pays du Tiers-Monde eux-mêmes pour améliorer leur situation alimentaire; toutefois, compte tenu des incertitudes, l'agriculture doit garder la possibilité et les capacités productives pour répondre à la demande alimentaire mondiale, pour le cas où des exportations importantes seraient nécessaires pour nourrir les 10 milliards d'hommes prévus dans quelques décennies. L'adaptabilité et la flexibilité s'imposent aussi plus généralement en raison de la montée des incertitudes en tous domaines : économiques, politiques, écologiques, culturels, climatiques, etc.

⁽⁶⁾ Les micro-marchés correspondent à des créneaux de diversification de dimension restreinte, comme par exemple la production de laine angora, d'escargots, lamas, bisons, autruches, plantes médicinales, etc. Cette diversification est intéressante pour les agriculteurs, mais assez souvent risquée que ce soit sur le plan technique où les références manquent parfois ou sur le plan économique car les prix peuvent chuter si trop d'agriculteurs choisissent les mêmes créneaux.

– **imaginatif** : le modèle doit être imaginatif pour répondre à la diversité des demandes. Il faut aussi mettre au point des systèmes écologiques, respectueux de l'environnement, des techniques généralisables à l'ensemble du monde, c'est-à-dire qui fassent un usage raisonné et patrimonial des ressources de la planète. L'agriculture doit également viser la mise en valeur des espaces ruraux. Il faut enfin rechercher des techniques et des productions qui respectent la culture (aspects éthiques), qui puissent être motivantes et valorisantes pour les agriculteurs et qui leur assurent une rémunération correcte. Compte tenu de la multiplicité des objectifs le nouveau modèle doit être particulièrement imaginatif.

– Enfin le modèle de production doit avoir pour objectif une **agriculture durable**, c'est-à-dire économiquement viable, écologiquement saine, socialement acceptable, moralement juste. Il s'agit là surtout d'un idéal à viser.

On notera que le modèle de production agricole diversifié et adaptatif que nous venons de définir a certains points communs avec le nouveau modèle productif qui émerge dans l'industrie suite à la crise du fordisme et qui se caractérise par la flexibilité, la recherche de la qualité de la production, une grande réactivité par rapport à la demande, une structure des firmes en réseau, etc. (cf. par exemple Coriat, 1990; Chesnais, 1992; Boyer, Durand, 1993).

Ce modèle diversifié, adaptatif et flexible et cette agriculture multifonctionnelle constituent-ils une rupture, un nouveau modèle? Bien sûr l'agriculture française présente déjà certaines de ces caractéristiques! Mais on peut parler de rupture avec le modèle productiviste sur le plan de la reconnaissance et de l'acceptation de cette diversité et sur le plan des objectifs à viser (n'oublions pas qu'un modèle de production est un paradigme qui détermine les buts à rechercher et les moyens à mettre en œuvre) : l'objectif unique n'est plus de produire et d'accroître les rendements comme dans le modèle productiviste. Il y a également une modification importante à attendre dans les prochaines années et décennies sur le plan technique, une rupture sociologique (cf. Hervieu, 1993), une rupture économique avec de nouveaux modes de rémunération, la valorisation de nouvelles activités et la production de services, enfin la rediversification des fonctions remplies par l'agriculture⁽⁷⁾. Mais le développement d'un nouveau modèle de production est un processus lent et contradictoire qui prend

(7) On peut parler de rediversification car jusqu'au milieu du xx^e siècle l'agriculture assurait d'autres fonctions que la production de masse de biens standardisés : production d'énergie et de force motrice (transport par les chevaux, chauffage au bois), entretien

du temps car il existe divers obstacles. Quoi qu'il en soit le modèle productiviste après avoir été très performant dans les décennies d'après-guerre paraît aujourd'hui devenu en grande partie caduc et obsolète et devrait sans doute laisser la place progressivement à un nouveau modèle.

VI. INNOVATION ET STANDARDISATION EN AGRICULTURE

L'agriculture comme tous les secteurs est caractérisée par d'importantes innovations de procédés et de produits. Comment s'opère l'innovation dans ce secteur compte tenu de sa standardisation et des réglementations qui y sont en vigueur? Le développement des nouvelles technologies en certains domaines pourrait renforcer les normes. En effet l'établissement de ces dernières, par exemple en ce qui concerne l'absence de résidus toxiques ou de pathogènes, pose un problème technique : celui des possibilités de les vérifier; il faut disposer de moyens de mesure pas trop coûteux, ni trop longs à mettre en œuvre ou de tests suffisamment sensibles, sinon l'existence de normes a peu de sens! Or on sera bientôt en mesure d'apprécier de façon de plus en plus précise et rapide la qualité d'un produit, sa composition, son origine. Par exemple la résonance magnétique nucléaire permet de déterminer l'origine d'un produit (le vin notamment), les sondes moléculaires permettent de caractériser les variétés, certains kits basés sur les anticorps monoclonaux ou l'amplification en chaîne de l'ADN rendent possible la détection de pathogènes (présence de moisissures ou de bactéries), etc.

Par ailleurs la mécanisation de certaines tâches, en agriculture comme ailleurs, évolue vers la robotisation : divers robots devraient être commercialisés très prochainement dans ce secteur. Ils renforcent généralement la standardisation : par exemple le robot de traite nécessite une conformation standard des mamelles de la vache, le robot de cueillette des pommes ou des grappes de raisin une certaine forme des arbres ou des plants de vigne.

du territoire (par des terrasses, murets, fossés, taille des haies, etc.), production de matières premières pour l'industrie chimique et textile (colorants, cuirs, laines, fibres, etc.), transformation des produits à la ferme. Avec l'ère du pétrole à bon marché et le modèle productiviste, depuis 1945 les fonctions de l'agriculture se sont davantage réduites à la production de matières premières agricoles.

Une autre innovation technologique accroissant la standardisation est le développement de la multiplication végétative *in vitro* et du clonage pour certaines plantes (et peut-être demain pour certains animaux). C'est l'un des rares cas où il y a production à l'identique d'organismes vivants (puisque'il n'y a pas fécondation qui entraîne une recombinaison génétique (cf. note 1)). Cette technique est utilisée en raison de certains de ses avantages (multiplication de plantes exceptionnelles résistant à certaines maladies, obtention de plants indemnes de virus, gain de temps), mais uniquement pour certains végétaux. Elle permet d'obtenir plus rapidement des plantes horticoles (rosier, œillet, orchidée, gerbera), ou à fruits (fraisier, framboisier), des arbres fruitiers ou forestiers. Mais elle doit être utilisée de façon limitée : il est indispensable que subsiste dans les populations végétales ou animales une certaine diversité génétique qui les rend plus résistantes aux maladies et qui assure une certaine vigueur hybride.

D'autres innovations de procédés vont par contre plutôt dans le sens d'une certaine diversification des pratiques culturales. Ainsi on cherche à adapter de plus en plus finement les apports d'intrants aux besoins des plantes, à faire des épandages d'engrais et de pesticides en fonction des caractéristiques de la parcelle plutôt qu'à effectuer des traitements standards systématiques (Bonny, 1993b). Ces pratiques devraient se développer pour des raisons environnementales, économiques et grâce à certaines avancées techniques. Ainsi les kits de diagnostic des agents pathogènes, le piégeage des insectes nuisibles sur la parcelle pour leur détermination et l'utilisation de modèles de prévision des maladies (avec diffusion des conseils et avertissements phytosanitaires régionaux par Minitel par exemple) pourraient permettre des traitements plus adaptés aux conditions locales. Le développement de la lutte intégrée et de la lutte biologique va également dans ce sens.

Les innovations de produits, quant à elles, sont très nombreuses dans l'agro-alimentaire où elles sont un élément essentiel de la stratégie des firmes : de nouvelles compositions, présentations, conditionnements, emballages, etc., apparaissent souvent sur les linéaires des supermarchés et représentent une diversification d'importance variable, effectuée notamment selon les différents segments de la clientèle. Dans l'agriculture, elles sont un peu moins nombreuses comme il s'agit de produits de base, non de produits transformés. Elles peuvent provenir de l'acclimatation d'espèces sauvages peu cultivées ou élevées jusque-là (cerfs, bisons, lamas, sangliers, autruches, pleurotes, etc.) ou de plantes et fruits exotiques (cf. par exemple la percée du kiwi et de l'avocat ces dernières décennies). On note également assez souvent l'apparition de nouvelles

espèces de plantes déjà connues, de nouvelles variétés du fait des recherches en amélioration génétique.

Mais la grande question pour les prochaines années et décennies concerne les plantes et animaux transgéniques dans lesquels on a inséré des gènes étrangers à leur espèce. Il s'agit là d'une révolution importante (Bonny, Daucé, 1989; Bonny, 1991 et 1993a). Quelques médicaments et vaccins produits par génie génétique sont déjà commercialisés, mais pour l'instant aucune plante ou produit animal, bien qu'un certain nombre de produits soient déjà prêts à être mis sur le marché. Un vif débat a lieu en effet sur ce sujet, certains s'opposant aux manipulations génétiques en raison de leurs risques éventuels pour l'environnement, de certaines de leurs conséquences socio-économiques ou pour des considérations d'ordre moral. Dans ce cas la réglementation, ici celle sur les organismes génétiquement modifiés, ralentit la commercialisation d'une innovation car on tente d'en évaluer les impacts à l'avance. La recherche et les expérimentations en ce domaine sont également réglementées. Certains Etats, suite à la pression de diverses associations, envisagent un étiquetage précisant que le produit a été obtenu avec transfert de gène. Cette technique pourrait entraîner une diversification par création de nouvelles plantes et animaux dont on a modifié certaines caractéristiques génétiques pour les adapter à différents types d'usages ou conditions de milieu : par rapport à l'amélioration des plantes et des animaux par les méthodes classiques de croisement, les possibilités du génie génétique sont plus vastes⁽⁸⁾. Mais elle peut entraîner une diminution de la diversité génétique en raison du clonage et une réduction des possibilités d'accès aux ressources génétiques du fait de la brevetabilité du vivant qui commence à apparaître, ce qui pourrait conduire à un appauvrissement génétique. Ainsi donc en ce domaine le double mouvement de diversification/standardisation paraît se poursuivre.

⁽⁸⁾ Avec la transgénèse les possibilités deviennent théoriquement immenses puisqu'on peut transférer un gène dont on a repéré les propriétés d'une espèce à une autre et faire par exemple exprimer le gène d'une protéine humaine dans un micro-organisme ou un animal, voire une plante, ou transférer un gène d'un végétal à un autre, ou d'un animal à un autre ou à un micro-organisme, etc., alors que le croisement naturel serait impossible. En effet par les méthodes classiques de croisement on est limité par la barrière de la reproduction sexuée : on ne peut croiser ensemble que des plantes ou des animaux de même espèce, ce qui limite les possibilités. Toutefois dans la pratique du génie génétique, le repérage des bons gènes, leur transfert, leur traduction et leur expression dans le nouvel organisme ne sont pas toujours aisés, d'autant plus que les caractères complexes sont déterminés par de nombreux gènes.

CONCLUSION

L'évolution de l'agriculture française durant les quatre décennies d'après-guerre s'est traduite par une certaine standardisation en matière de produits, de procédés et d'organisation. Cela est d'ailleurs souvent critiqué par les consommateurs qui déplorent parfois la standardisation des produits et des paysages, ou par des généticiens ou écologistes qui s'inquiètent de la diminution de la diversité génétique. Toutefois à côté de la production de masse existe depuis plusieurs décennies une fabrication de produits «de qualité» (appellation d'origine contrôlée, labels, agriculture biologique, etc.) qui est très réglementée, mais qui correspond à une certaine diversification. Un standard de qualité peut ainsi permettre d'échapper à la standardisation (au sens de production uniforme).

Le modèle de production des décennies d'après-guerre qui est en crise depuis le milieu des années 70 pourrait évoluer sous l'effet conjugué des applications des nouvelles technologies (biotechnologies, nouvelles technologies de l'information), de l'évolution de la demande adressée à l'agriculture et de l'évolution des réglementations internationales. Bien que certaines tendances aillent vers un renforcement de la standardisation, le nouveau modèle qui pourrait se développer progressivement correspond plutôt à une certaine diversification des processus de production (avec par exemple adaptation fine des apports d'intrants aux besoins des cultures et des élevages) et à une évolution vers une agriculture multifonctionnelle et vers la flexibilité, l'adaptabilité, la diversification et la créativité. On pourrait voir l'agriculture produire simultanément des biens de masse standards à bas prix⁽⁹⁾, des matières premières pour la chimie et l'énergie (on parle de moléculture), des produits de qualité soumis à des spécifications et réglementations particulières (labels, AOC, certifications), des services (entretien du territoire et de la nature, tourisme, etc.). Mais l'évolution du modèle sera progressive, contradictoire et sans doute difficile.

⁽⁹⁾ Production de masse à bas prix ne signifie pas absence de qualité. Cette dernière est au contraire de plus en plus mise en avant. Mais la définition précise de la notion de qualité varie selon les acteurs de la filière agro-alimentaire : pour les uns il s'agit de la composition en protéines ou lipides ou de la teneur en telle ou telle protéine particulière, pour d'autres il s'agit de l'absence de bactéries pathogènes, pour d'autres encore de l'absence de résidus de pesticides, pour les négociants de l'homogénéité et du caractère standard du produit, pour les distributeurs d'une belle apparence et d'une bonne capacité de conservation, pour le consommateur c'est souvent la qualité organoleptique qui est préférée. Compte tenu de cette polysémie, on comprend qu'il ne soit pas toujours aisé de produire de «la qualité».

En définitive, en agriculture comme dans les autres secteurs, la standardisation permet notamment des économies d'échelle, une économie des coûts de transaction, une économie de la confiance. A côté de cela, les standards de qualité rendent possibles des rentes liées à l'appellation et à des modes de fabrication particuliers, et correspondent à une diversification des produits en fonction de divers segments du marché; par ailleurs pourrait se développer à nouveau une rediversification des fonctions remplies par l'agriculture qui en diminuerait la standardisation. La dynamique des relations standardisation – diversification – innovation est ainsi particulièrement riche et complexe en agriculture.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Alphandéry P. et al, *Les formes de l'intensification en élevage bovin*, Paris, INRA-ESR, 1980.
- Blanchet J., Carles R., «Réforme de la politique agricole commune et systèmes de production», INRA, *Actes et communications* (10), juin 1993, 233 p.
- Boisard P., Letablier M.-T., «Le camembert : normand ou normé. Deux modèles de production dans l'industrie fromagère», *Entreprises et produits. Cahiers du Centre d'études de l'emploi* (30), 1987, p.1-29.
- Bonny S., *L'évolution technologique en cours en agriculture et ses conséquences : quelques jalons pour un repérage et une analyse socio-économiques*, Grignon, INRA-ESR, Notes et Documents N° 39, septembre 1991, 91 p.
- Bonny S., *Vers un nouveau modèle de production en agriculture?*, Grignon, INRA-ESR, 1993a, 240 p.
- Bonny S., *Le changement technique en cours et à venir en agriculture : un essai de bilan dans les différents secteurs*, Grignon, INRA ESR; *Etudes économiques*, n°12, 1993b, 134p.
- Bonny S., Carles R., «Perspectives d'évolution de l'emploi des engrais et des phytosanitaires dans l'agriculture française», *Cahiers d'économie et sociologie rurales* (26), 1er trimestre 1993, p. 29-62.
- Bonny S., Daucé P., «Les nouvelles technologies en agriculture», *Cahiers d'économie et sociologie rurales* (13), 1989, p. 5-33.
- Boyer R., Durand J.-P., *L'après-fordisme*. Paris, Syros (Collection alternatives économiques), 1993, 174 p.
- Boyer R., Schméder G., «Division du travail, changement technique et croissance. Un retour à Adam Smith», *Revue française d'économie*, 5(1), hiver 1990, p. 125-194.
- CGP, *France rurale : vers un nouveau contrat*, Paris, La Documentation française, Préparation du XI^e Plan. Commission «Agriculture, alimentation et développement rural», présidée par P. Mangin, 1993, 172 p.
- Chalmin P., El Alaoui A., *Matières premières et commodités*, Paris, Economica, 1990, 141 p.

- Chesnais F. (coord.), *La technologie et l'économie : les relations déterminantes*, Paris, OCDE, 1992, 364 p.
- Chauvet M., Olivier L., *La biodiversité, enjeu planétaire*, Paris, Sang de la terre, 1993, 413 p.
- CNEEJA, *De l'industrialisation à la régression de l'agriculture*. Grenoble, Université des sciences sociales, Centre national d'études économiques et juridiques agricoles, 1971, 133 p.
- Coriat B., *L'atelier et le robot : essai sur le fordisme et la production de masse à l'âge de l'électronique*, Paris, Christian Bourgois, 1990, 303 p.
- David P.A., «Some new standards for the economics of standardization in the information age», in : Dasgupta et Stoneman, *Economic policy and technological performance*, Cambridge University Press, 1987, p. 206-239.
- David P.A., Greenstein S., «The economics of compatibility standards : an introduction to recent research», *Economics of innovation and new technology* 1(1-2), 1990, p. 1-32.
- Dosi G., «Technological paradigms and technological trajectories». *Research Policy* (11), 1982, p. 147-162
- Dosi G., «Sources, Procedures and Microeconomics Effects of Innovation». *Journal of Economic Literature*, XXVI, Sept. 1988, p. 1120-1171.
- Dosi G., Freeman C., Nelson R., Silverberg G., Soete L., *Technical Change and Economic Theory*, Londres et New York, Pinter Publishers, 1988, 646 p.
- Eymard-Duvernay F., «Conventions de qualité et formes de coordination», *Revue économique*, vol. 40, N° 2, mars 1989, p. 329-360.
- Faure C., *Les paysans dans la production capitaliste*, Université Paris VIII Vincennes, travaux sur le capitalisme et l'économie politique (10), 1974, 61 p.
- Ferné G., «Normes, sciences et techniques» in : *La science au présent*, Paris, Encyclopaedia Universalis, 1992, p. 390-392.
- Fischler C., *L'omnivore*, Paris, Odile Jacob, 1990, 440 p.
- Foray D., «Exploitation des externalités de réseau versus évolution des normes : les formes d'organisation face au dilemme de l'efficacité, dans le domaine des technologies de réseau», *Revue d'économie industrielle* (51), 1990, p. 113-140.
- Foray D., «Standard de référence, coûts de transaction et économie de la qualité : un cadre d'analyse», *Economie rurale* (217), sept-oct. 1993, p. 33-41
- Foray D., «Standardisation et concurrence : des relations ambivalentes», *Revue d'économie industrielle* (63), 1er trimestre 1993, p. 84-101.
- Heertje A., *Innovation, technologie et finance*. Paris, Banque européenne d'investissement, 1988, 199 p.
- Hervieu B., *Les champs du futur*, Paris, François Bourin, 1993, 173 p.
- Jollivet M. (dir.), *Pour une agriculture diversifiée*, Paris, L'Harmattan, 1988, 336 p.
- Kindleberger C., «Standards as public, collective and private goods», *Kyklos*, vol. 36, (fasc. 3), 1983, p. 377-396.
- Mendras H., *La fin des paysans : essai*, Actes Sud, 384 p. (1ère édition 1967), 1991.
- Ministère de l'agriculture, «La qualité des produits agro-alimentaires», *Les dossiers de la Politique Agricole Commune*, (6/ 7), septembre 1992, 79 p.
- Nicolas F., Valceschini E. (coord.), «La qualité dans l'agro-alimentaire», *Economie rurale* (217), sept-oct 1993, 68 p.
- OTA, *Global standards. Building blocks for the future*, Washington. Congress of the U.S., Office of Technology Assessment, 1992, 114 p.

- Perez C., «Structural change and the assimilation of new technologies in the economic and social system», *Futures*, 1983, p. 357- 375.
- Rip A., Groenewegen P., «Les faits scientifiques à l'épreuve de la politique», in : Callon M. (dir.) *La science et ses réseaux, genèse et circulation des faits scientifiques*. Paris, La Découverte, Conseil de l'Europe, UNESCO, 1989, p. 149-172
- Salomon J.-J. et Schméder G., *Les enjeux du changement technologique*, Paris, Economica, 1986, 206 p.
- Servolin C., «Aspects économiques de l'absorption de l'agriculture dans le mode de production capitaliste», in : Tavernier Y., Gervais M., Servolin C. (dir.), *L'univers politique des paysans dans la France contemporaine*, Paris, A. Colin, 1972, p. 41-78.
- Sylvander B., *Les conventions de qualité dans le secteur agro-alimentaire : aspects théoriques et méthodologiques*, Communication au Colloque de la SFER «La qualité dans l'agro-alimentaire», 26-27 octobre 1992, 30 p.
- Tabary P., «La politique agricole commune et les goûts», *Revue des deux mondes*, janvier 1993, p. 55-65.
- Thévenot L., «Economie et politique de l'entreprise. Economie de l'efficacité et de la confiance», *Cahiers du Centre d'études de l'emploi* (33), Justesse et justice dans le travail, 1989, p. 135-208.
- Valceschini E., «La qualité des produits agricoles et alimentaires dans le marché unique européen : l'épreuve de la concurrence et de la confiance», in : Déméter 93, *Economie et stratégies agricoles*, Paris, A. Colin, 1993, p. 119-162.

Innovations pharmaceutiques et réglementation : le paradigme de l'explication

Jérôme Dumoulin

CNRS-IREPD
Université Pierre Mendès France, Grenoble

Les réglementations ont été accusées d'avoir freiné les innovations pharmaceutiques dans le monde à partir des années 60. Cet effet est habituellement expliqué par un modèle de l'innovation reposant sur le paradigme de la sélection aveugle. En fait, à partir des années 60, ce paradigme est peu à peu supplanté par le paradigme de l'explication qui est cohérent avec des réglementations renforcées. La crise des innovations pharmaceutiques peut alors être expliquée par les difficultés d'adaptation au nouveau paradigme.

Regulation is often accused of curbing pharmaceutical innovations since the 60's. This process is mostly explained by a model of innovation based on the paradigm of blind selection. Actually, this paradigm has been replaced by the paradigm of legitimation which is consistent with stronger regulations. Then, the crisis of pharmaceutical innovations can be explained by problems of adaptation to this new paradigm.

Je remercie très sincèrement Paul Bouvier-Patron, William Cavestro, Bruno Lamotte et Jean-Paul Laurencin des remarques pertinentes qu'ils ont faites à une version précédente de ce texte et qui m'ont permis de l'améliorer.

«Le médicament n'est pas une denrée comme les autres, a-t-on pu dire. Tout au contraire, il est une denrée comme les autres, au carrefour de la générosité et de la cupidité, de la grâce et de la concupiscence, du moi et du surmoi.»
G.A. Marcel, *Le médicament de l'an 2000*, Masson, 1987, p. 162.

Dans tous les pays industriels, l'industrie pharmaceutique est très réglementée : Bonnes Pratiques de Laboratoire, Bonnes Pratiques Cliniques, Bonnes Pratiques de Fabrication, Autorisations de Mise sur le Marché, Critères éthiques applicables à la promotion des médicaments. Comment comprendre et expliquer ces réglementations ?

Les analyses économiques de la réglementation font référence à de nombreuses justifications (Encaoua, 1986) : la réglementation établit l'infrastructure juridique, elle est un moyen de politique conjoncturelle, elle se substitue au marché défaillant (*market failures*) pour réguler certaines activités. A ce dernier objectif correspondent deux types de réglementations. «La réglementation "économique" d'une activité porte essentiellement sur les prix et sur la structure d'offre de l'activité. Par opposition, la réglementation "sociale" ou "qualitative" concerne les conditions d'exercice dans l'activité et les caractéristiques des produits ou services offerts» (Encaoua, 1986).

L'aspect le plus original de la réglementation pharmaceutique est qu'elle porte sur les caractéristiques des produits offerts; les produits ne sont pas définis directement : la réglementation porte sur les conditions de l'innovation et non sur les produits eux-mêmes. Cette réglementation peut être rapprochée des normes, moyens de gestion de l'innovation (Kirat, 1990), que celles-ci soient d'origine étatique ou industrielle. Les normes ont été étudiées surtout pour leur fonction de compatibilité (par exemple en mécanique, en électronique), là où la concurrence porte sur des combinaisons de produits matériels, de produits intellectuels et de services (Bala, Imai, 1993), ou lorsque la concurrence porte sur les appellations (Thévenot, 1992, Bessy, Chateauraynaud, 1993). En pharmacie, la concurrence porte essentiellement sur l'existence de produits indépendants. La fonction des normes est de garantir leur qualité. Le problème de la qualité provient, d'un point de vue économique, de la situation d'asymétrie d'information : la connaissance limitée par les consommateurs de l'utilité et du danger des médicaments justifie de faire de ces derniers des biens sous tutelle. Les débats aux Etats-Unis, lors de la création de la réglementation pharmaceutique, portaient bien sur la mise en tutelle des consommateurs par l'Etat et les professionnels (Temin, 1979). Selon d'autres terminologies, on fait référence à des objectifs de santé publique, de protection des consommateurs, ou d'éthique. Le médicament est alors considéré

comme un bien de nature spécifique; la réglementation consisterait à le faire échapper en grande partie aux stratégies commerciales pour que les producteurs n'abusent pas de leur pouvoir. L'objectif de limiter les effets de l'asymétrie d'information conduit à celui d'organiser la concurrence (Hancher, 1990), soit pour la favoriser, soit pour la limiter, par exemple en avantageant certaines entreprises (entreprises nationales face aux entreprises étrangères, grandes entreprises face aux petites) par la création de barrières à l'entrée dans l'activité. La réglementation a, de ce fait, des objectifs et des effets multiples (Maynard, Hartley, 1984), sans que l'on puisse facilement distinguer une réglementation économique d'une réglementation « sociale ».

La question la plus débattue à propos de la pharmacie est celle des effets de la réglementation sur les innovations et la compétitivité des entreprises des différents pays. La réglementation s'est développée dans le monde entier à partir de 1962 lorsque, aux Etats-Unis, l'amendement Kefauver a réglementé les essais cliniques et exigé que les nouveaux médicaments fassent la preuve de leur efficacité. Selon une idée largement acceptée parmi les économistes (Comanor, 1986), ces réglementations ont élevé le coût de la Recherche – Développement, et par là limité le flux d'innovation. Cette incidence est considérée souvent comme fort dommageable tant pour l'industrie (qui subit une contrainte sur ses possibilités innovatrices) que pour les consommateurs, qui ne bénéficieraient pas de tous les progrès de la science. Les réglementations seraient ainsi contre-productives. Certains anti-interventionnistes en prennent argument pour soutenir qu'il faut supprimer toute réglementation de l'innovation pharmaceutique. Selon Milton Friedman (1978, cité par Karsenti, 1978), « Les amendements de 1962 doivent être révisés. Ils sont en train de faire plus de mal que de bien. Pour les appliquer, les responsables de la F.D.A. sont obligés de condamner d'innocentes personnes à la mort... D'autres études pourraient même démontrer que la F.D.A. elle-même devrait être abolie. » Cette thèse a été reprise par des auteurs français comme Aftalion (1981), Rosa (1986), Peytavin (1988).

La thèse du caractère contre-productif de la réglementation peut être réfutée par une observation majeure (Thomas, 1993) : les pays où les restrictions à l'innovation sont les plus sévères, comme le Royaume-Uni comparé à la France, sont les pays où les innovations marquantes sont les plus nombreuses et où l'industrie pharmaceutique est la plus prospère. L'interprétation est qu'en élevant le niveau d'exigence pour les nouveaux médicaments la réglementation incite les entreprises à innover de façon plus radicale, à créer des médicaments d'intérêt

thérapeutique plus élevé, qui s'imposent ensuite plus facilement sur le marché mondial. Cette observation nous conduit à centrer l'analyse sur le processus d'innovation lui-même.

La démarche suivie ici consiste à analyser le processus d'innovation, ses transformations, pour comprendre le développement de la réglementation et son rôle. Par réglementation, on désignera les règles et normes qui concernent le processus d'innovation pharmaceutique : recherche, développement, production, commercialisation, que leur origine soit étatique ou professionnelle. Sont exclues du champ d'analyse les réglementations concernant les prix et les brevets.

Le rôle attribué à la réglementation dépend de la représentation que l'on se fait du processus d'innovation. Toutes les thèses évoquées précédemment considèrent que la réglementation crée des entraves à l'innovation, que ces thèses soient favorables ou non à ces entraves. Les études du rôle de frein de la réglementation ne reposent dans la plupart des cas sur aucune analyse du processus d'innovation (Domanor, 1986); dans le meilleur des cas elles utilisent une représentation linéaire du processus d'innovation, selon laquelle celui-ci est un processus séquentiel de *sélection* (par exemple DiMasi *et alii*, 1991). Le flux d'innovations pharmaceutiques est mesuré par le nombre de nouvelles entités chimiques (NEC) qui est fonction des dépenses en Recherche – Développement et de la sévérité de la réglementation (Baily, 1972; Pelzman, 1973; Grabowski et Vernon, 1977; Grabowski, Vernon et Thomas, 1978, Wiggins, 1980; DiMasi, Hansen, Grabowski et Lasagna, 1991). Cette sévérité est mesurée par la longueur des délais d'approbation par les autorités qui est la pierre de touche de l'analyse de la réglementation (par exemple Anderson, 1992; Schweitzer, 1993).

La thèse soutenue ici est que la réglementation peut être favorable à l'innovation, même si elle peut aussi la freiner. L'argumentation est qu'un nouveau paradigme guide le processus d'innovation : le *paradigme de l'explication* s'est substitué au paradigme antérieur de *sélection aveugle*. Si le paradigme de la sélection aveugle pouvait représenter assez bien le processus d'innovation à une certaine période, ce processus s'est transformé au moment même où la réglementation s'est renforcée. Avec le paradigme de l'explication, la réglementation est productive. La réglementation ne cherche pas à soustraire le médicament aux logiques économiques, mais au contraire elle est un investissement de forme (Thévenot, 1986) qui donne au produit un caractère général lui permettant d'entrer dans des relations d'échange et par là elle permet au marché d'exister et de se développer.

Une question préalable est de savoir si le changement de paradigme est bien une réalité (les processus d'innovation ont alors vraiment changé de nature) ou bien s'il ne concerne que la représentation des processus d'innovation. Notre point de vue est le suivant : le processus d'innovation a bien été transformé, transformation lente et non encore achevée complètement, mais les analyses habituelles s'appuient toujours sur l'ancien paradigme, bien que des changements soient reconnus. Le fait que les représentations n'ont pas fondamentalement changé freine lui-même le passage au nouveau paradigme : l'action des entreprises et des Etats est encore guidée par des représentations largement dépassées par les faits.

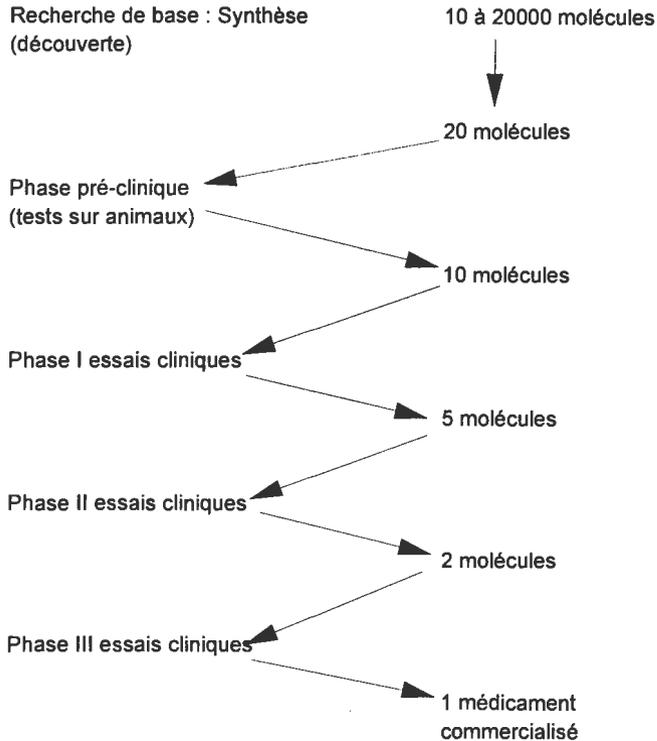
On exposera dans une première partie le ralentissement des flux d'innovations et l'accroissement des coûts de Recherche-Développement tels qu'ils sont habituellement expliqués avec le recours au paradigme de sélection aveugle. Dans une seconde partie, on examinera les transformations du processus d'innovation à partir des années 60 avec l'émergence et le développement du paradigme de l'explication. Il sera montré dans une troisième partie en quoi une réglementation forte est cohérente avec ce paradigme et donc avec une capacité d'innovation accrue.

I. LE PARADIGME DE SÉLECTION AVEUGLE : LA RÉGLEMENTATION FREINE L'INNOVATION

Selon ce paradigme, on décrit le processus d'innovation pharmaceutique comme une série de phases successives au cours desquelles de nombreux tests permettent de sélectionner peu à peu la molécule qui sera commercialisée, à partir de quelque 10 ou 20 000 molécules initiales. La représentation de l'innovation suit la métaphore du pipe-line (Spikler, 1989), ou de l'entonnoir.

Environ 20 molécules sont sélectionnées par des méthodes de *screening* pour le développement pré-clinique (essais sur les animaux), environ 10 molécules pour la phase I des essais cliniques, 5 pour la phase II et 2 pour la phase III (Langle, Ocelli, 1983). Cette représentation repose sur la méthode de *screening* aveugle qui sélectionne les molécules pour leur activité *in vitro*. Mais le qualificatif «aveugle» peut s'appliquer à l'ensemble du processus : c'est la «roulette moléculaire» (Walsh, 1992). Le choix des molécules de départ repose presque entièrement sur le hasard : un effet inattendu ou indésirable d'un médicament ou

Tableau 1
**La Recherche – Développement d'un nouveau médicament
 selon le paradigme de la sélection aveugle
 (métaphore du pipe-line)**



d'une substance existante (poison, drogue, plante,...) est observé par hasard grâce à une découverte heureuse («serendipity»). Cette molécule et d'autres molécules proches sont testées systématiquement, car on espère que, parmi celles-ci, le hasard fera que l'une sera plus active et moins toxique. Le processus de test est également aveugle : on ne sait pas par quels mécanismes une molécule est plus intéressante qu'une autre car, jusque dans les années 60, les relations entre la structure d'une molécule et son activité étaient très peu connues (Pelt, 1969). Ces relations ne pouvaient servir de base à la découverte. De très nombreux exemples de ce processus de découverte heureuse et de sélection aveugle peuvent être cités.

Le caractère de sélection aveugle se retrouvait également dans les phases suivantes du processus d'innovation. L'approbation officielle de nouveaux médicaments par les Etats reposait entièrement sur un rapport d'expert se prononçant sans critère préétabli, principalement au nom de son autorité légitimée par sa compétence reconnue (Jeunemaître, 1984; Hancher, 1991). La diffusion reposait sur des réseaux où les leaders d'opinion jouaient un rôle majeur. Le modèle épidémiologique expliquait la diffusion par une série de contacts (Coleman, Katz, Menzel, 1966) entre les praticiens premiers utilisateurs « innovateurs » et les autres praticiens. La confiance joue un rôle essentiel dans l'explication de la diffusion : elle est la véritable explication. Mais le support obscurcit le sens du message sur l'intérêt du nouveau médicament. Le message n'a pas d'existence autonome. Un médicament nouveau est adopté à cause de sa seule nouveauté (Dupuy, Karsenty, 1974). Dans le processus de sélection aveugle, la réglementation apporte des exigences nouvelles en matière de toxicité, de tératologie, d'effets secondaires, et d'apport thérapeutique qui obligent à multiplier les essais, à allonger leur durée, et à augmenter le nombre des sujets sur qui les produits sont testés. A cela s'ajoutent les délais de l'administration pour approuver un nouveau médicament. L'origine de la réglementation est extérieure au processus d'innovation, elle est attribuée à la demande sociale et au souhait légitime de réduire les accidents. En augmentant le nombre de tests et leur durée, la réglementation augmente la durée totale de Recherche-Développement, donc les coûts : elle réduit le nombre d'innovations et pèse sur la rentabilité de la Recherche-Développement. A l'appui de cette thèse, de très nombreux faits peuvent être mis en avant.

La durée moyenne de développement et d'approbation par les autorités des nouvelles entités chimiques commercialisées aux Etats-Unis est passée de 2,7 années en 1966 à 6,6 années en 1973 (Schwartzman, 1976). Au Royaume-Uni la durée moyenne de développement est passée d'environ 4 ans dans les années 1961 à 1966 à plus de 10 ans à partir de 1977 (Corstjens, 1991). De très nombreux auteurs citent cette durée moyenne de dix ans; une durée de 12 ans est même citée (Langle, Ocelli, 1983; Joglemar, Parterson, 1986). En parallèle à l'augmentation des délais, le coût de la Recherche-Développement s'est également accru. Bien que la mesure des coûts pose de difficiles problèmes et que la comparaison de différentes études soit très hasardeuse (Woltman, 1990), il est incontestable que le coût de la Recherche-Développement des nouveaux médicaments s'est considérablement accru. Les données qui peuvent être citées (DiMasi *et alii*, 1991) pour le coût de Recherche-Développement d'un nouveau médicament sont aux Etats-Unis de 21,8

millions de dollars dans les années 60 (Baily, 1972), 54 millions de dollars pour la période 1963-1975 (Hansen, 1979), et 231 millions de dollars pour la période 1970-1992 (DiMasi *et alii*, 1991). La hausse des coûts est attribuée principalement à la réglementation : même si d'autres facteurs sont évoqués, comme les taux d'intérêt et le taux de succès des recherches, elle accroît les délais de mise sur le marché, le nombre et la complexité des tests exigés. Bien moins difficile à mesurer est l'évolution du nombre annuel de nouveaux médicaments commercialisés. Ce nombre a baissé de moitié dans le monde entre 1961 et 1981 (Reis-Arndt, 1982). Aux Etats-Unis, la baisse a été très forte entre la période 1957-1961 (233 nouvelles entités chimiques) et la période 1962-1968 (93 NCE). Le flux annuel d'innovations s'est ensuite stabilisé aux Etats-Unis après 1970 (Grabowski et Vernon, 1977) et dans le monde entre 1980 et 1984 (Barral, 1990).

Il y a eu croissance considérable des budgets des entreprises consacrés à la Recherche-Développement. En conséquence, la rentabilité moyenne des dépenses de recherche aurait diminué, mais cette baisse est beaucoup plus difficile à évaluer pour les deux raisons suivantes : d'une part la rentabilité d'un nouveau médicament ne dépend pas seulement de la recherche, car elle tient aussi aux conditions de commercialisation (par exemple l'apparition de copies à l'expiration du brevet) et d'autre part il y a des variations considérables d'un médicament à l'autre, d'une entreprise à l'autre, dans l'efficacité des processus de recherche (Joglekar, Paterson, 1986). Il en résulte qu'aux Etats-Unis, l'innovation s'est concentrée davantage dans les grandes entreprises (Grabowski et Vernon, 1977).

La description de l'innovation pharmaceutique par le paradigme de la sélection aveugle peut faire l'objet de trois sortes de critiques :

1) Il ne suffit pas que les méthodes de *screening*, voire de découverte et de diffusion de nouvelles molécules, soient des processus de sélection et que ces processus soient relativement aveugles pour que l'ensemble constitue réellement un *processus de sélection aveugle*. La description est partiellement valable *ex post*, mais est-elle valable *ex ante*? La représentation de l'innovation selon le paradigme de la sélection aveugle suppose en effet qu'il y a indépendance totale entre des phases qui doivent se dérouler séparément les unes des autres : la recherche serait un processus purement mécanique, inintelligent, sans anticipation ni apprentissage. L'ensemble des connaissances de la pharmacie jouerait un rôle très modeste, sinon nul dans ce modèle : il ne porterait que sur le choix des tests à effectuer. Tout le reste serait guidé par le hasard (comment identifier une substance active? pourquoi est-elle active?) ou

par la nécessité (on doit rejeter une substance si un test est défavorable). Peut-on sérieusement considérer que les chercheurs et industriels n'aient aucun rôle propre, aucune initiative, aucune stratégie?

2) Depuis les années 60, les évolutions concernant les délais, les coûts de la Recherche-Développement et le nombre d'innovations ne peuvent pas être expliquées par la seule réglementation. Les délais se sont stabilisés dans les années 80 et ils se réduiraient depuis. Le nombre annuel d'innovations au niveau mondial s'est stabilisé entre 1980 et 1985 et s'est à nouveau accru depuis. Par contre, la croissance des coûts se poursuit encore de façon ininterrompue. Si on peut comprendre que la réglementation ait mis un certain temps à produire tous ses effets, pourquoi les délais et le nombre annuel d'innovations se sont-ils stabilisés dans les années 80, alors que le coût de la Recherche-Développement continue encore de croître?

3) Ainsi que le signalent la plupart des auteurs, l'allongement des délais et des coûts de la recherche peuvent s'expliquer également par d'autres causes que la réglementation. C'est la thèse de l'épuisement des méthodes et des domaines de recherche (Grabowski et Vernon, 1977). Dans les années 60, la recherche s'est portée sur des maladies beaucoup plus complexes pour lesquelles le processus de découverte est plus long et plus coûteux. Le *screening* aveugle perd de son efficacité. Pour ces maladies, le *screening* aveugle est complété par d'autres méthodes de recherche qui le transforment, par exemple, en «*screening* orienté». Il y a donc transformation des processus de Recherche-Développement. Mais avec ces transformations, la représentation de l'innovation par le paradigme de la sélection aveugle devient de moins en moins pertinente. Il faut donc vérifier si l'analyse du rôle de la réglementation résiste à une modification de l'analyse de l'innovation.

Ces limites du pouvoir descriptif et du pouvoir explicatif du paradigme de la sélection aveugle amènent à examiner de plus près le processus d'innovation lui-même et les changements qui l'ont affecté. Le paradigme de l'explication paraît capable d'en rendre compte plus complètement, et de rendre compte aussi du rôle de la réglementation à partir des années 60.

II. LE PARADIGME DE L'EXPLICATION

Les nombreux changements qui ont affecté l'industrie pharmaceutique depuis les années 60 peuvent être qualifiés par le terme de

«scientificité» (Chesnay, 1985) ou de révolution (Scherer, 1993). Il ne suffit pas de les recenser. Il faut qualifier et analyser cette révolution pour comprendre sa rationalité et ses effets. Pour cela, il sera exposé en quoi le processus d'innovation ne peut plus être décrit de façon linéaire, mais en quoi on doit recourir à un modèle systémique : relations enchaînées et réseaux. Selon ce paradigme, le point crucial est alors l'explication de l'action des médicaments.

1. Le caractère plus scientifique de l'innovation

Il est commun d'attribuer aux «biotechnologies» les progrès scientifiques affectant la pharmacie, ce qui est à la fois une tautologie et une erreur. Si par biotechnologies on désigne les sciences et les techniques du vivant, les biotechnologies ne sont pas une nouveauté en pharmacie, qui a toujours été une science et une technique du vivant (tautologie). L'erreur est que, si l'on ne désigne que le génie génétique, très peu de médicaments jusqu'à présent sont issus des «biotechnologies» (Levêque, 1990). Ces progrès doivent être décrits avec plus de justesse.

Depuis les années 60, la recherche pharmaceutique a connu l'émergence de deux grands paradigmes scientifiques qui permettent d'analyser et de prévoir l'action des médicaments, en fournissant des modèles à leurs mécanismes d'action : les récepteurs et la génétique moléculaire. Les récepteurs sont les lieux de fixation des médicaments sur les cellules. La théorie des récepteurs permet de comprendre l'action des médicaments au niveau des inter-relations moléculaires : synergie, potentialisation ou antagonisme de deux médicaments, inversion des effets en fonction de la dose, etc. On a là un des mécanismes d'explication de la relation structure-activité et un des moyens de découverte des médicaments par identification du ou des récepteurs possibles. La génétique moléculaire s'est développée à partir de la connaissance de l'ADN, de sa structure et de ses fonctions de codage. Sur cette base, a pu se développer le génie génétique qui modifie le code génétique de cellules pour que celles-ci produisent la protéine désirée. On a là des théories générales qui donnent des principes gouvernant les phénomènes biologiques et l'action des médicaments.

Ces deux grands paradigmes de la biologie se sont développés en relations étroites avec toute une série d'autres connaissances et d'autres progrès scientifiques, en particulier celles portant sur la structure en trois dimensions et l'action des protéines, la chimie des enzymes, la pharmacodynamique (transformation des médicaments dans

l'organisme), la toxicologie, et aussi la pharmacologie clinique (effet du médicament sur l'organisme humain). Les progrès de la biologie moléculaire se sont accompagnés de progrès au niveau cellulaire et au niveau clinique. Ces différents progrès portent sur la connaissance des mécanismes biologiques, sur les méthodes d'observation et de mesure.

Les changements dans les diverses disciplines ont pour conséquence de permettre la mise en relation de diverses connaissances existantes ou nouvelles. Les nombreuses connaissances pharmaceutiques issues de l'observation et des tests étaient tacites, localisées, contextuelles, privées car appartenant aux entreprises ayant fait les tests; leur utilisation était pratiquement impossible en dehors de l'environnement où elles ont été produites. Comme dans de nombreux autres domaines, elles deviennent plus explicites, plus générales, plus abstraites et aussi plus publiques (Arora, Gambardella, 1992), grâce à des théories biomédicales qui peuvent exprimer des réalités très variées dans un langage unificateur, comme l'immunologie (Moulin, 1991). Ces caractéristiques permettent aux connaissances d'être transférées ailleurs qu'à leur lieu d'émergence, pour expliquer d'autres aspects d'un phénomène, pour étudier d'autres mécanismes biologiques, d'autres pathologies, et utilisées par d'autres chercheurs. Elles peuvent être mises en relation, ce qui leur donne une nouvelle dynamique. Mais il y a toujours des connaissances spécifiques et tacites, car les connaissances théoriques et générales ne permettent pas de tout expliquer, de tout prévoir. L'expérimentation est irremplaçable, le hasard a toujours une place. Les méthodes déductives se combinent aux méthodes inductives.

2. Le processus d'innovation : relations enchaînées et construction d'un réseau de connaissances

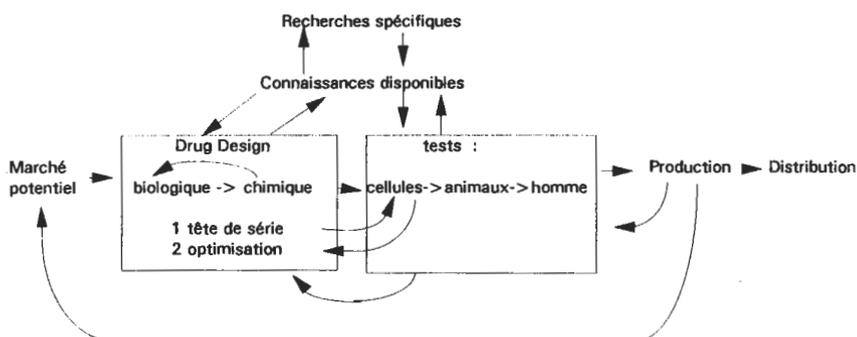
Le processus d'innovation pharmaceutique peut être représenté assez précisément par «ce que l'on appelle communément maintenant le modèle de liaisons en chaîne» (OCDE, 1992) proposé par Kline et Rosenberg (1986). Ce modèle comprend :

- une chaîne centrale de l'innovation qui part d'un marché potentiel, passe à l'invention d'un concept nouveau, puis le teste, le produit et le distribue sur le marché. Cette chaîne centrale est assez proche de la conception linéaire, en pipe-line, de l'innovation;

- des boucles de rétroaction entre phases successives de la chaîne centrale. Ces boucles signifient que l'arrêt du processus à un stade donné n'est pas systématiquement un échec, lorsqu'il permet un processus d'apprentissage;

- une importante boucle de rétroaction entre la distribution sur le marché et l'appréhension du marché potentiel : les firmes cherchent à anticiper, voire à créer des marchés futurs;
- des liaisons avec l'ensemble des connaissances scientifiques et techniques disponibles auxquelles il est fait appel pour résoudre les problèmes identifiés au cours de la chaîne centrale;
- des activités de recherche spécifiques lorsque les connaissances disponibles sont insuffisantes.

Tableau 2
La Recherche-Développement d'un nouveau médicament
 selon le paradigme de l'explication



Le processus d'obtention-sélection de substances actives consiste, comme auparavant, à trouver en premier lieu une molécule «tête de série» ayant l'activité désirée, puis à l'optimiser, c'est-à-dire à rechercher, parmi les molécules proches, celle qui aura l'activité la plus forte, la toxicité et les effets secondaires les plus faibles. Le changement porte sur la manière d'identifier une «tête de série» : celle-ci se fait moins au hasard en s'appuyant davantage sur des connaissances générales et sur le processus d'optimisation; celui-ci comporte de multiples allers et retours très rapides entre phases de conception et d'expérimentation.

La phase de découverte est devenue moins aveugle grâce au «*drug design*» qui permet de créer des «prototypes virtuels» (Arora, Gambardella, 1992) :

- le «*drug design*» d'origine biologique identifie dans quel mécanisme biologique la substance visée doit prendre place pour inhiber ou faciliter ce mécanisme. La recherche décompose la pathologie que l'on veut traiter en mécanismes biologiques qui permettent d'identifier des cibles possibles, par exemple récepteurs membranaires sur lesquels

le médicament pourra se placer. Les bêta-bloquants (le propranolol) ont été les premiers médicaments obtenus par cette méthode. La recherche tire profit de l'avancée des connaissances en médecine et en biologie, qui éclairent de mieux en mieux les processus physiologiques et pathologiques;

– le «*drug design*» d'origine chimique comporte l'ensemble des méthodes qui permettent d'identifier la molécule pouvant avoir telle ou telle action chimique recherchée. Le plus souvent, il s'agit de modifier une molécule déjà connue pour obtenir des propriétés améliorées. Des méthodes informatiques permettent de déterminer par calcul les conformations possibles des molécules aptes à avoir telle ou telle action. Ces calculs portent sur les caractéristiques tridimensionnelles, électroniques, hydrophobiques des molécules. La conception de nouvelles molécules s'appuie sur les progrès des connaissances en chimie. Les tests réalisés au cours de la phase expérimentale portent moins souvent sur des animaux vivants et de plus en plus souvent sur des cellules dont on teste la réaction des récepteurs à la substance étudiée (Ballet *et alii*, 1990). En réduisant la quantité de substances actives à préparer pour les tests, en apportant une meilleure compréhension de leurs résultats, ces changements permettent une plus fine adéquation de la phase expérimentale à la phase conceptuelle. Leur automatisation devient possible. Tout ceci accélère les allers et retours entre la conception de nouvelles voies d'approche, celle de nouvelles molécules et la vérification expérimentale. La connaissance plus précise des mécanismes biologiques permet également une meilleure extrapolation à l'homme des effets observés sur les animaux et sur la cellule. Elle permet une meilleure anticipation de phases de la Recherche-Développement situées en aval.

L'évaluation des effets sur l'homme a fait également des progrès importants avec les méthodes statistiques de la pharmacologie clinique qui permettent de distinguer l'effet placebo de l'effet pharmacologique, de mesurer des améliorations de l'état des patients (car les médicaments ne prétendent pas tous guérir) parfois à long terme, d'observer des effets indésirables même rares, de différencier les effets selon les sous-groupes de patients, etc. Ces méthodes ne permettent pas seulement de mesurer l'intérêt d'un médicament futur ou déjà commercialisé, mais d'aider à mieux comprendre les mécanismes biologiques. Ainsi, l'étude ISIS 3 qui a comparé la streptokinase, le rTPA et l'anistreplase pour l'infarctus aigu du myocarde sur 41 299 cas randomisés (ISIS-3, 1992) n'a pas épuisé la question de l'efficacité de ces médicaments, malgré sa puissance statistique, car ses résultats ont incité à une nouvelle étude

(GUSTO) de même dimension pour mieux comprendre le déroulement des infarctus.

Le modèle de relations en chaîne met l'accent sur le caractère systémique du processus : les différents types de connaissances et leurs progrès sont étroitement liés les uns aux autres. Une avancée pratique se fait en prenant appui sur des connaissances générales; les connaissances empiriques fournies par les tests alimentent à leur tour les connaissances générales. Le résultat du processus est à la fois de trouver un nouveau médicament et de faire progresser les connaissances générales et théoriques, qui peuvent être utilisées ensuite dans le processus. Même si une molécule est écartée par un test, celui-ci peut aider à mieux comprendre les phénomènes en cause, et permettre des progrès ultérieurs. Ainsi, dans une grande entreprise française, de nombreux tests de toxicologie qui étaient sous-traités à des entreprises spécialisées ont été réintégrés dans l'entreprise, dans le but de mieux analyser et comprendre leurs résultats.

Néanmoins, il y a toujours recours à des méthodes empiriques déductives, à des tests très nombreux. A cause de la complexité du vivant, les connaissances générales sont encore limitées. Dans l'innovation pharmaceutique, le hasard joue toujours un rôle non négligeable (Cognata, Schwartzman, 1992), mais il n'est plus le seul maître du jeu. Les innovateurs, chercheurs et industriels ne comptent plus sur la multiplication du nombre de produits testés pour accroître la probabilité d'aboutir; ils cherchent à limiter l'importance du hasard. Le fait que tous les domaines de la biologie, de la médecine, de la pharmacie soient interreliés entraîne qu'une découverte locale provient souvent de progrès réalisés ailleurs; par exemple la découverte des anti-cancéreux a commencé avec des recherches sur les éléments radioactifs, la nutrition des microbes et les substances pour la guerre chimique (Weatherall, 1990). La découverte ne suit pas de route directe, mais sinueuse, avec des retours en arrière, des pauses... L'activité de recherche est risquée, car on ne peut complètement prévoir une découverte : elle n'est pas toujours là où on l'attend et elle se produit souvent là où on ne l'attend pas. Mais la démarche de recherche ne se fait pas pour autant au hasard. Les méthodes déductives, formalisées ont certes leurs limites, mais l'autre branche de l'alternative n'est pas le pur hasard : ce sont des méthodes intuitives, non formalisées qui mobilisent toute une série de connaissances tacites. Les activités de recherche sont sélectives, finalisées et cumulatives (Dosi, 1988), très éloignées d'une série de tirages au jeu à la roulette.

Les deux aspects de l'innovation, que sont la découverte (phases de créativité) et la vérification (phases de sélection) sont étroitement entremêlés. Cela a plusieurs conséquences. Illustration de la caducité du modèle du pipe-line qui sépare ces deux phases, la durée totale du processus n'est plus liée au nombre et à la sévérité des tests, mais à la coordination et à la planification des opérations (Ellul, 1991), à la pertinence de la voie d'approche retenue, au choix des molécules à tester, à l'anticipation des propriétés des molécules conçues. Par exemple, les chimistes ne visent plus à synthétiser le plus grand nombre possible de molécules nouvelles, mais celles supposées avoir l'effet souhaité, en cherchant à écarter celles qui risquent fort ne pas passer les épreuves suivantes. L'organisation de la Recherche-Développement doit donc favoriser avant tout la communication entre équipes, entre disciplines (Bibard *et alii*, 1993). Les phases de tests et d'expérimentation où l'on vérifie l'activité des substances créées participent aussi au processus de découverte. Les exigences méthodologiques (rigueur, scientificité,...) de la vérification doivent être au même niveau que celles de la découverte pour que la vérification puisse être ensuite utile à celle-ci.

La recherche et le développement d'une molécule nouvelle consistent à relier toute une série de faits, connaissances, observations entre eux, ce qui aboutit à construire un réseau. Lorsque cette construction est achevée, les énoncés exprimant les connaissances des différents points de vue (disciplines scientifiques par exemple) sont complètement reliés et compatibles avec tous les autres énoncés (Callon, 1991); leur cohérence produit un énoncé unique qui explique en quoi, pourquoi et comment le nouveau médicament a bien certaines propriétés (Bibart, 1991). Cette construction se fait à travers des traductions d'énoncés en d'autres énoncés. Le processus est très long, il se poursuit au cours de la phase de diffusion. Il n'est d'ailleurs jamais complètement achevé. Cette suite de traductions peut être formulée par le paradigme de l'explication : l'activité d'un nouveau médicament est expliquée (c'est-à-dire traduite et justifiée) selon tous les différents points de vue pertinents (biologie, chimie, biochimie, toxicologie, analyse, pharmacodynamique, pharmacologie clinique, etc.). Et surtout, les différentes explications doivent être cohérentes les unes avec les autres. On a là un véritable paradigme technologique, dans la mesure où il sert de « modèle de solution pour des problèmes technologiques sélectionnés, basé sur des principes sélectionnés dérivés des sciences naturelles et sur des technologies matérielles choisies. (...) Il détermine le champ de la recherche, les problèmes, les procédures et les tâches » (Dosi, 1982).

Depuis longtemps, l'innovation pouvait être décrite en partie selon le modèle des relations enchaînées, car l'innovation n'a jamais suivi complètement le modèle de sélection aveugle. Depuis longtemps, la recherche de substances actives s'est appuyée sur des théories, que celles-ci soient vraies ou fausses, comme par exemple la théorie des signatures. La place des théories s'est réduite dans les années 40 et 50 au profit de méthodes expérimentales, dont la sélection aveugle. Puis, il y a eu un glissement progressif d'un système dominé par la sélection aveugle et les connaissances empiriques à un système où des connaissances générales ont une grande place. L'émergence des paradigmes scientifiques de la théorie des récepteurs et de la biologie moléculaire a accéléré ce changement.

III. LA NATURE ET LA PLACE DES RÉGLEMENTATIONS À LA LUMIÈRE DU PARADIGME DE L'EXPLICATION

Le paradigme de l'explication permet d'interpréter les réglementations de façon nouvelle en considérant que les différentes réglementations permettent d'organiser ces explications. La réglementation remplit deux fonctions : structurer l'information, définir les modalités de la preuve.

1. Structurer l'information

Pour que les différents types de connaissances, d'énoncés puissent être traduits en d'autres énoncés et ne restent pas localisés, ils doivent être mis en forme, organisés selon des règles de généralité. Ainsi, les tests doivent suivre certaines règles pour pouvoir être fiables et reproductibles. Les normes qui remplissent cette fonction peuvent être qualifiées d'*infratechnologies* (Tassey, 1986). «Elles recouvrent l'ensemble des étalons et des instrument de mesure, des méthodes d'essai et d'expérimentation, de codification des propriétés des matériaux, des composants.» La création de ces normes est particulièrement intense lors de l'émergence d'un nouveau paradigme, car elles servent de «support de la communication entre les agents impliqués dans le processus de création technologique» (Kirat, 1990).

2. Définir les modalités de la preuve

L'énoncé qu'un médicament nouveau a certaines propriétés doit être prouvé. A cette étape, se pose le problème de la preuve scientifique. Si l'on retient la thèse de Popper (1970), la connaissance scientifique avance moins par vérification déductive que par réfutation. Dans ce sens, les vérités scientifiques sont toujours provisoires. Mais pour qu'un médicament soit utilisé, la vérité doit être stabilisée dans un énoncé, certes provisoire, mais qui puisse résister aux réfutations les plus évidentes, les plus grossières. Le réseau de l'innovation n'est pas limité aux services de Recherche-Développement de l'entreprise innovatrice, il est complètement ouvert sur l'extérieur, un extérieur d'autant plus vaste que le médicament est commercialisé à l'échelle mondiale. Au stade de la commercialisation, le risque de réfutation s'accroît considérablement, car de très nombreux médecins et pharmaciens peuvent étudier le nouveau médicament.

En indiquant les tests à effectuer et comment les effectuer, les règles définissent comment doivent être apportées les preuves et comment l'énoncé d'efficacité d'un médicament peut être réfuté. Bien que reposant sur de très nombreuses connaissances, ces règles sont la plupart du temps tout à fait arbitraires. Elles permettent d'organiser et de simplifier le débat scientifique. Par exemple, elles définissent les procédures de test toxicologique, comme le nombre d'espèces animales et le nombre d'animaux sur lesquels il faut tester le médicament. Il n'y a pas de garantie scientifique que le médicament ne soit en aucun cas toxique, ce que pourraient montrer, par exemple, des tests non prévus. En l'absence de réglementation, la non-toxicité ne pourrait pratiquement jamais être démontrée. Une autre règle, non écrite, mais très forte, veut que, dans les études randomisées en double aveugle on retienne comme risque adverse le seuil de 5%. Ainsi, le seuil de 5% signifie que, lorsque la probabilité adverse est de 6%, le médicament sera considéré sans discussion comme non supérieur au placebo; si elle est de 4%, il sera considéré comme supérieur. La différence est très faible, voire infinitésimale, mais, dans le processus de l'explication, elle est considérable. La norme de 5% permet de clore le débat, d'interdire la réfutation lorsque la probabilité est de 4,9%. Autre exemple : le code de la commercialisation publié par la Fédération internationale de l'industrie du médicament établit que «le mot "sans danger" ne doit être utilisé que sous réserves». En effet, l'absence totale d'innocuité est pratiquement indémontrable et la règle permet d'éviter des explications sans issue.

Ces règles ont donc un double rôle : le plus évident oblige à l'explication, mais en même temps elles limitent les possibilités pratiquement infinies de réfutation. Elles ne bloquent pas complètement toute réfutation d'une vérité admise provisoirement, mais en interdisent certaines. La réglementation est donc productive en facilitant le processus d'explication.

Cette fonction d'explication n'est pas contradictoire avec la fonction de protection du malade : elles se complètent. Tous les auteurs (par exemple Hancher) soulignent le caractère très réglementé de l'industrie pharmaceutique à cause de nécessités de protection des consommateurs : ce sont des règles de sécurité. Cela signifie que les tests effectués pendant le développement d'un médicament doivent répondre à des exigences situées dans la phase de diffusion. Ceci est vrai, mais ces exigences ne sont pas extérieures au processus d'innovation. Comme l'utilité et le danger d'un médicament ne sont pas directement accessibles aux consommateurs et aux prescripteurs, à partir de la connaissance de leur composition, ou à partir d'un simple essai, l'appréciation de la qualité d'un médicament repose en grande partie sur la confiance des consommateurs, condition de l'existence du marché. Les nombreuses catastrophes qui ont jalonné l'histoire de la pharmacie ont miné chaque fois la confiance des consommateurs : BCG à Lubeck en 1930, Elixir de Sulfanilamide-Messengill aux Etats-Unis en 1937, Stalinon en France en 1954, Thalidomide en Allemagne en 1960, Distilbène dans le monde entier en 1984, etc. (Weatherall, 1990, Jeunemaître 1984). Ces catastrophes ont entraîné des scandales politiques incitant les Etats à intervenir, mais surtout ont entamé la confiance des consommateurs. Il était de l'intérêt des entreprises de rétablir cette confiance pour pouvoir ensuite diffuser les nouveaux médicaments. Les règles sur l'efficacité et la sécurité des médicaments participent à la production et à l'entretien de cette confiance. Ceci explique que l'enregistrement des médicaments a été réalisé en France et au Royaume-Uni à l'initiative et sous la responsabilité de l'industrie elle-même, avant l'intervention de l'Etat (Hancher, 1991). Certains (Thomas, 1992) pensent que, même sans réglementation, la lourdeur des essais cliniques se serait accrue, car la preuve de l'efficacité est devenue plus difficile et la technologie des tests a beaucoup progressé.

La réglementation rend donc possible l'ensemble du processus d'innovation dont les différentes étapes sont reliées : découverte, développement, diffusion.

3. Crise de l'innovation et réglementation

La crise de l'innovation pharmaceutique (réduction du flux annuel d'innovations, accroissement des délais et des coûts) peut s'expliquer par le passage du paradigme de la sélection au paradigme de l'explication. Jusqu'à la fin des années 50, le paradigme de la sélection était la base de l'organisation de la Recherche-Développement pharmaceutique. La méthode de sélection aveugle a été très efficace en permettant de trouver par exemple de très nombreux antibiotiques. Mais vers 1960, cette méthode a commencé à être moins efficace : s'il fallait synthétiser 3000 molécules aux Etats-Unis pour sélectionner un nouveau médicament en 1962, il en fallait 7000 en 1967 et 10 000 en 1973 (Miquel, Doré, 1981). De nouvelles méthodes de Recherche-Développement ont permis, outre leur utilité pour la découverte, de multiplier les modalités d'évaluation des médicaments. Les agences gouvernementales se sont emparées de ces méthodes et ont poussé leur développement. Ainsi, au moment de la catastrophe de la thalidomide, on s'est rendu compte qu'on disposait des méthodes qui auraient permis d'en prévoir les effets tératogènes (Weatherall, 1990). La même année 1962 a connu l'amendement Kefauver aux Etats-Unis et la découverte du premier bêta-bloquant, par utilisation de la théorie des récepteurs.

Le nouveau paradigme ne s'est pas mis en place instantanément, il a nécessité un délai d'adaptation des entreprises et des politiques publiques au cours duquel le flux d'innovation a été ralenti. Une fois les entreprises organisées en fonction de ce nouveau paradigme (début des années 80), le flux d'innovation a pu commencer à croître et les délais à se réduire. Dans les années 80, la durée moyenne du développement des médicaments était de 9 ans environ pour Bristol et Upjohn, mais moins de 5 ans pour Schering, Smith-Kline et Merck, et moins de 3 ans pour Glaxo (Ellul, 1991). La durée de vie effective des brevets, qui était passée de 17,4 ans en 1962 à moins de 7 ans en 1981, s'accroît à partir de 1982 pour dépasser 10 ans en 1989 (Ballance *et alii*, 1993).

Par contre les coûts continuent à croître. Un montant de 200 millions de dollars est cité pour la Recherche-Développement de produits issus du génie génétique (Scotto, 1993). La croissance de ces coûts tient à la complexité des processus biologiques que l'on cherche à modifier et des molécules que l'on cherche à élaborer. Si de nouvelles connaissances permettent d'aborder ces maladies de façon nouvelle, le stock des connaissances publiques, libres, permet rarement de résoudre tous les problèmes rencontrés. Des détours de recherche «fondamentale» sont souvent nécessaires, et les essais empiriques peuvent être encore

plus nombreux. La croissance des coûts ne peut pas être liée à la réglementation considérée comme une variable exogène, car elle est liée à la dynamique d'ensemble de la Recherche-Développement pharmaceutique.

«Le fait que de nombreuses entreprises changent leurs stratégies de Recherche-Développement est la reconnaissance implicite qu'une bonne partie des problèmes est interne aux entreprises» (Ballance *et alii*, 1993). L'adaptation des entreprises a de très nombreux aspects. Une nouvelle organisation de la Recherche-Développement est nécessaire (Bibard *et alii*, 1993), pour coordonner les différentes activités de Recherche-Développement, acquérir des connaissances en toxicologie, en pharmacologie clinique, en biologie moléculaire, disciplines qui ont une place maintenant stratégique. Cet élargissement de la base de connaissances en toxicologie, biologie moléculaire et pharmacologie clinique n'était pas encore réalisé dans les entreprises françaises à la fin des années 80 (Commissariat général du Plan, 1990). L'organisation des entreprises doit également s'adapter en utilisant davantage de chercheurs bien intégrés à la communauté scientifique, et moins de techniciens d'expérimentation (Arora, Gambardella, 1992), pour qu'elles soient capables d'utiliser les connaissances scientifiques générales et de suivre étroitement leurs avancées dans les différentes disciplines concernées. Les stratégies de Recherche-Développement doivent être concentrées sur des thèmes de recherche plus limités, en identifiant les découvertes fondamentales passées qui peuvent être les plus utiles et (Templé, 1989) en sélectionnant davantage les produits au stade du développement. L'organisation de la Recherche-Développement est souvent hésitante et doit souvent être modifiée devant les difficultés rencontrées (Ballance *et alii*, 1993). L'adaptation porte également sur les méthodes d'analyse des marchés qui doivent devenir moins empiriques et sur l'intégration des stratégies de recherche et des stratégies commerciales (Dumoulin, 1992).

De nombreuses entreprises se sont adaptées avec réticence à la réglementation : par exemple en considérant les tests demandés comme de simples moyens pour remplir les dossiers administratifs, et non comme des étapes pour la connaissance du médicament (Legrain, 1986), ou en présentant des dossiers manquant de clarté, ce qui allonge les délais d'approbation (Spikler, 1989). Ce changement suppose également des changements dans les valeurs, par exemple des valeurs éthiques : «l'adhésion aux normes éthiques les plus élevées, dans les domaines scientifiques, médicaux, commerciaux et autres, [...] procure de nombreux avantages aux firmes en termes d'examen de leurs

demandes par l'administration, de capacité de répondre aux questions, critiques et accusations depuis l'extérieur de la firme, fournit en outre un puissant facteur de motivation au personnel» (Spilker, 1989).

L'adaptation est variable selon les entreprises, et selon les pays où elles sont implantées. Ainsi, l'adaptation de l'industrie pharmaceutique britannique au paradigme de l'explication a été plus rapide que l'adaptation de l'industrie française, parce que (Hancher, 1991) l'organisation de l'enregistrement mis en place était différente dans ces deux pays à la fois à cause de la structure de l'industrie et à cause de traditions nationales. En France, l'enregistrement créé avant guerre avait surtout pour but de protéger de la concurrence les producteurs qui étaient surtout de petites entreprises de faible niveau technologique et concurrentiel, et non de protéger les consommateurs. La procédure reposait essentiellement sur des expertises décentralisées sans méthodologie précise. Il s'agissait d'une stratégie défensive. Au Royaume-Uni, l'enregistrement établi en 1964 était offensif, vis-à-vis des consommateurs et des marchés étrangers, au service d'une industrie de grande taille et de grande capacité technologique et concurrentielle. Les expertises étaient centralisées par un organisme national qui a participé au développement des méthodes d'expérimentation.

Cette adaptation est ralentie, en particulier en France, par la persistance, dans les entreprises, de la représentation de l'innovation par le paradigme de la sélection aveugle. Les paradigmes technologiques ont un effet d'exclusion (Dosi, 1984) : les efforts des chercheurs et des gestionnaires sont orientés dans des directions assez précises, tout en étant rendus aveugles aux autres possibilités technologiques. La mise en avant permanente du modèle du pipe-line par les entreprises françaises (comme, encore récemment, *Annales des Mines*, 1991) pour justifier un allongement de la durée de protection des brevets, pour tenter d'obtenir un assouplissement des réglementations, pour obtenir de meilleurs prix des médicaments, ou en justifier les prix élevés, a eu probablement un effet d'auto-intoxication les rendant aveugles aux changements en cours dans les processus d'innovation. N'étant pas complètement conscientes du changement de paradigme, elles n'ont pas cherché à s'y adapter au plus vite.

CONCLUSION

Toutes les réglementations ayant une incidence sur l'innovation n'ont pas été examinées. La réglementation des prix joue un rôle

complètement différent de celui qu'on a vu vis-à-vis de la dynamique de l'innovation. Les pays où les prix sont les plus faibles, du fait de la réglementation (par exemple la France et l'Italie), sont également ceux où l'industrie pharmaceutique est la moins dynamique (Thomas, 1992, Thomas, 1993). La réglementation des prix a pour effet de gommer les différences compétitives entre firmes lorsqu'elle donne à un médicament innovateur un prix peu différent des prix des médicaments anciens : les firmes sont alors moins incitées à innover radicalement. La réglementation des prix a de ce fait un effet inverse de la réglementation de mise sur le marché des nouveaux médicaments. La réglementation des brevets nécessite probablement aussi une analyse spécifique, si on voit par exemple ce qui se passe en matière des semences (Joly, Ducos, 1993).

Si on met de côté les prix et les brevets qui nécessitent des analyses spécifiques, il faut reconsidérer les effets économiques négatifs habituellement attribués à la réglementation dans le processus d'innovation pharmaceutique. Son principe est accepté par les industriels au moins pour des raisons éthiques, à cause du grand danger potentiel des médicaments.

Les changements scientifiques et techniques qui ont affecté l'industrie pharmaceutique depuis les années 60 ne peuvent être analysés simplement par la modernisation ou la «scientifisation» de ses activités. Il s'agit d'un nouveau paradigme technologique, managérial et organisationnel. On est passé du paradigme de la sélection aveugle au paradigme de l'explication. C'est un véritable changement de système technologique, porteur d'une crise d'ajustement structurel (Freeman, Perez, 1988). L'émergence de ce paradigme est le résultat, non seulement des progrès scientifiques et techniques de la pharmacie, mais aussi de l'histoire de cette industrie.

L'exemple de l'industrie pharmaceutique montre que ce changement de système technique est long et complexe pour plusieurs raisons : acquisitions de nouvelles connaissances par les entreprises, création de nouvelles connaissances et nouveaux modes de savoir-faire, réorganisation interne des entreprises pour gérer autrement la Recherche-Développement, création de nouvelles réglementations. La durée de l'adaptation au nouveau système technologique tient non seulement au délai nécessaire pour chacun des changements, mais surtout au fait que les différents acteurs n'ont pas immédiatement conscience du changement profond en train de se produire. Leur représentation de l'innovation reste longtemps attachée au paradigme de la sélection aveugle, et au rôle de frein de la réglementation.

Les réglementations se développent dans des contextes économiques, politiques, scientifiques particuliers : elles sont le produit d'une histoire. Rechercher une justification unique aux réglementations est de ce fait quelque peu réducteur. Dans le cas de la réglementation pharmaceutique, il a souvent été mis l'accent sur la fonction de protection des consommateurs : la réglementation viserait à ne laisser apparaître que les innovations socialement avantageuses. Même si les intérêts des consommateurs et des producteurs peuvent être divergents, la réglementation remplit également une fonction pour les producteurs en facilitant les processus d'innovation. La réglementation aurait alors pour fonction d'aider à organiser les relations entre les différents acteurs de l'innovation. Reste à savoir comment les règles sont interprétées.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Aftalion F., «Le coût social de la réglementation de l'industrie pharmaceutique», *La santé des Français*, Cercles Universitaires, PUF, 1981.
- Anderson F., «The drug lag issue : The debate seen from an international perspective», *International Journal of Health Services*, 22, 1992, p. 53-72.
- Annales des Mines, «Santé : un complexe médico-industriel», *Réalités industrielles*, juillet-août 1991.
- Arora A., Gambardella A., «New trends in technological change : toward an innovative division of labour», *Second International Conference on the Economics of Innovation*, Piacenza, June 12-13 1992.
- Baily M.N., «Research and Development Costs and Returns : The U.S. pharmaceutical industry», *Journal of Political Economy*, 80,1, Jan-Feb 1972, p. 70-85.
- Bala Y., Imai K., «Innovation entrepreneurs et réseaux : l'exemple de la naissance des magnétoscopes», *Revue internationale des sciences sociales*, 135, février 1993, p. 27-39.
- Ballance R., Pogany J., Forstner H., UNIDO, *The World's Pharmaceutical Industries, an International Perspective on Innovation, Competition and Policy*, London, Edward Elgar 1993, 275 p.
- Ballet F., Melcion C., Cordier A., «Nouvelles stratégies en toxicologie», in *L'Europe du médicament, réalités et ambitions*, 4^e colloque DPHM-INSERM, Les Editions INSERM, Paris, 1990, 470 p.
- Barral P.E., *Quinze ans de résultats de la recherche pharmaceutique dans le monde (1975-1989)*, Fondation Rhône-Poulenc Santé, Antony, 1990, 106 p.
- Bessy C., Chateauraynaud F., *Les contraintes de la preuve dans la nouveauté des objets*, Noisy-le-Grand, Centre d'études de l'Emploi, 18 p.
- Bibart L., «Histoire d'une innovation, Le cas d'un médicament», *Revue économique*, 42, 2, 1991, p. 273-299.
- Bibart L., Byé P., Dumoulin J., Monateri J.-C., «Recherche, développement, stratégie : Rhône-Poulenc», in *Stratégie technologique et avantage concurrentiel*, Grenoble, IREPD, 1993, 310 p.

- Callon M., «Réseaux socio-techniques et irréversibilité», in Boyer R., Chavance B., Godart O., *Les figures de l'irréversibilité en économie*, Editions de l'Ecole des Hautes Etudes en sciences sociales, 1991.
- Chesnay F., «L'industrie pharmaceutique dans la crise», *Revue d'économie industrielle*, 31, 1^{er} trimestre, 1985, p. 79-99.
- Cognata A., Schwartzman D., «Has Pharmaceutical Research Become More Scientific?», *Second International Conference on the Economics of Innovation*, Piacenza, June 12-13 1992.
- Coleman J., Katz E., Menzel H., *Medical Innovation : A Diffusion Study*, Indianapolis, Bobbs-Merrill, 1966.
- Comanor W.S., «The political Economy of the Pharmaceutical Industry», *Journal of Economic Literature*, XXIV, September 1986, p. 1178-1217.
- Commissariat général du Plan, Groupe de travail «Les opportunités et risques de 1992 pour l'industrie pharmaceutique», in *L'industrie chimique et pharmaceutique française dans la perspective de 1992*, Paris, La Documentation française, 1990.
- Corstjens M., *Marketing Strategy in the Pharmaceutical Industry*, London, Chapman & Hall, 1991, 243 p.
- DiMasi J.A., «Rising Research and Development Costs for New Drugs in a Cost Containment Environment», *PharmacoEconomics*, 1, suppl 1, 1992, p. 13-20.
- DiMasi J.A., Hansen R.W., Grabowski H.G., Lasagna L., «Cost of innovation in the pharmaceutical industry», *Journal of Health Economics*, 10, 1991, p. 107-142.
- Dosi G., 1982, «Technological Paradigms and technological trajectories : a suggested interpretation of the determinants and directions of a technical change», *Research Policy*, 11(3), 1982, p. 147-162.
- Dosi G., *Technical Change and Industrial transformation. The Theory and a application to the semi-conductor industry*, Mac Millan, 1984, 338 p.
- Dosi G., «The nature of the innovative process», in Dosi G., Freeman C., Nelson R., Silverberg G., Soete L., *Technical Change and Economic Theory*, London, Pinter Publisher, 1988.
- Dumoulin J., «L'évaluation économique du médicament par les firmes pharmaceutiques : de l'approche empirique à l'approche scientifique», *Journal d'économie médicale*, 4/5, 1992, p. 191-201.
- Dupuy J.-P., Karsenty S., *L'invasion pharmaceutique*, Paris, Editions du Seuil, 1974.
- Ellul E., «Réduire les délais de développement d'un produit pharmaceutique», *Réalités industrielles*, juillet-août 1991, p. 40-42.
- Encaoua D., «Réglementation et concurrence : quelques éléments de théorie économique», *Economie et Prévision*, 76, 1986, p. 7-46.
- Freeman C., Perez C., «Structural crisis of adjustment is, in Dosi G., Freeman C., Nelson R., Silverberg G., Soete L., *Technical Change and Economic Theory*, London, Pinter Publisher, 1988.
- Friedman M., «Frustrating Drug Advancement», *Newsweek*, 8/1/1973.
- Grabowski H.G., Vernon J.M., «Innovation and Invention, Consumer Protection Regulation in Ethical Drugs», *American Economic Review*, 67, Feb. 1977, p. 359-364.
- Grabowski H.G., Vernon J.M., Thomas L.G., «Estimating the effects of regulation on innovation : an international comparative analysis of the pharmaceutical industry», *The Journal of Law and Economics*, n° 21, 1, 1978, p. 133-163.

- Hancher L., *Regulating for Competition. Government, Law, and the Pharmaceutical Industry in the United Kingdom and France*, Oxford, Clarendon Press, 1991, 429 p.
- Hansen R.W., «The pharmaceutical development process : estimates of development costs and times and the effect of proposed regulatory changes», in Chien R.I. ed., *Issues in Pharmaceutical Economics*, Lexington, D.C. Health, 1979.
- ISIS-3, «A randomised comparison of streptokinase vs tissue plasminogen activator vs anistreplase and of aspirin plus heparin vs aspirin alone among 41 299 cases of suspected acute myocardial infarction», *The Lancet*, vol. 339, n° 8796, 28 March 1992, p. 753-770.
- Jeunemaître A., *Le contrôle social du médicament : le paradoxe de la pharmacologie clinique*, Thèse pour le doctorat de 3^e cycle, Université Paris 1, 1984.
- Jeunemaître A., *Régulation et instrumentation, La politique du médicament en France et aux USA*, Paris, Centre de recherche en gestion, Ecole Polytechnique, 1985, p. 121.
- Joglemar P., Paterson M.L., «A closer look at the returns and risks of pharmaceutical R & D», *Journal of Health Economics*, 5, 1986, p. 153-157.
- Joly P.B., Ducos C., *Les artifices du vivant. Stratégies d'innovation dans l'industrie des semences*, INRA Editions, Economica, 1993, p. 422.
- Karsenti R., «La recherche dans l'industrie pharmaceutique, moteur de développement», *Cahiers de sociologie et de démographie médicales*, XVIII, 2, p. 59-58.
- Kirat T., «Les normes, instrument de gestion de l'innovation?», *Economies et Sociétés*, Série Sciences de Gestion n°15, mai 1990, p. 165-191.
- Kline S.J., Rosemberg N., «Innovation : An Overview», in R. Landau & N. Rosemberg eds., *The Positive Sum Strategy*, National Academy Press, 1986.
- Langle L., Ocelli R., «Le coût d'un nouveau médicament», *Journal d'économie médicale*, 1, 2, 1983, p. 77-106.
- Legrain M., «La mise sur le marché du médicament, contraintes scientifiques et réglementaires», in Chatain René éd., *Marketing pharmaceutique*, Paris, Technique et documentation Lavoisier, 2 vol., 1986.
- Levêque F., «Les produits biotechnologiques à l'épreuve des faits», *La Recherche*, 217, janvier 1990, p. 118-127.
- Maynard A., Hartley K., «The Regulation of the Pharmaceutical Industry», in Lingren B. ed., Arne Ryde *Symposium on Pharmaceutical Economics*, Swedish Institute for Health Economics and Liber Förlag, 1984, p. 123-137.
- Miquel J.-F., Doré J.-C., «L'industrie pharmaceutique et la recherche de nouveaux médicaments», *Revue d'économie industrielle*, 18, 4, 1981, p. 237-255.
- Moulin A.M., *Le dernier langage de la médecine : histoire de l'immunologie de Pasteur au Sida*, Paris, Presses Universitaires de France, 1991.
- OCDE, *La technologie et l'économie, Les relations déterminantes*, Paris, 1992.
- Pelt J.-M., *Les médicaments*, Les Editions du Seuil, coll. Le rayon de la science, 1969, 1990 p.
- Peltzman S., «An Evaluation of Consumer Protection Legislation : The 1962 Drug Amendments», *The Journal of Political Economy*, vol. 81, 1973, p. 1049-1091.
- Peytavin J.-L., «L'industrie pharmaceutique malade de l'Etat», *Le Figaro*, 18 novembre 1988.
- Popper K., *La logique de la découverte scientifique*, Paris, Payot, 1970.

- Reis-Arndt E., «25 Jahre Arzneimittelentwicklung, Neue Pharmazeutische Wirkstoffe, 1961-1985», *Pharma Dialog*, 96, cité par Barral, 1990.
- Rosa J.- J., «Le médicament en économie de marché : contraintes industrielles, sociales, et administratives», *Prospective et santé*, 1986, 159 p.
- Scherer F. M., «Pricing, Profits and Technological Progress in Pharmaceutical Industry», *Journal of Economic Perspectives*, 7, 3, 1993, p. 97-115.
- Schwartzman D., *Innovation in the Pharmaceutical Industry*, Baltimore, The Johns Hopkins University Press, 1976, 399 p.
- Schweitzer S.O., «Comparaison France Etats-Unis de l'autorisation de mise sur le marché de nouveaux produits pharmaceutiques», *Journal d'économie médicale*, 11, 1, 1993, p. 33-44.
- Scotto M., «Bâle mise sur le génie génétique», *Le Monde*, 20 avril 1993.
- Spikler B., *Multinational Drug Companies, Issues in Drug Discovery and Development*, New York, Raven Press., 1989.
- Tassey G., *The role of the National Bureau of Standards in supporting industrial innovation*, IEEE Transactions on Engineering Management, 33, 3, août 1986, cité par Kirat, 1990.
- Temin P., «The origin of compulsory drug prescription», *Journal of Law and Economics*, 22 (1), April 1979, p. 91-105.
- Templé P., «La recherche-développement dans l'industrie pharmaceutique : une appréciation des résultats», *Economie et Prévision*, n°88-89, 2/3, 1989, p. 89-94.
- Thévenot L., «Les investissements de forme», in *Conventions économiques*, Paris, Centre d'études de l'emploi, Presses Universitaires de France, 1986, p. 21-71.
- Thévenot L., «Des marchés aux normes», in Boyer R. et Allaire G., *Régulation et conventions dans l'agriculture et l'agro-alimentaire*, 1992.
- Thomas L.G., «Price Regulation, Industry Structure and Innovation, An International Comparison of Pharmaceutical Industries», *PharmacoEconomics*, 1, suppl. 1, 1992, 9-12.
- Thomas L.G., *Implicit Industrial Policy : The Triumph of Britain and the Failure of France in Global Pharmaceuticals*, Atlanta, Emory University, School of Business, 1993, p. 35.
- Walsh V., «Firm strategies, globalisation and new technological paradigms : the case of biotechnology», *Entreprises et industries européennes face à la mondialisation*, Saint-Malo, 25-26 juin 1992.
- Weatherall M., *In Search of a Cure, A History of Pharmaceutical Discovery*, Oxford University Press, 1990, 298 p.
- Wiggins S.N., «Product quality regulation and new drug introductions : some new evidence from the 70's», *The Review of Economics and Statistics*, LXVIII, 4, 1981, p. 615-619.
- Wiggins S.N., *The cost of developing a new drug*, Washington, Pharmaceutical Manufacturers Association, 1987.
- Woltman H.R., «Reviewing the Bidding : Recherche et développement Costs and Profitability of New Chemical Entities», *Journal of Research in Pharmaceutical Economics*, 1, 3, 1989, p. 49-65.

Spécialisation et échanges intra-branche : reflet de l'absence de politique industrielle commune dans le cas des neuf Etats-membres de la CEE

Dimitri Mardas *

Maître assistant, département de Sciences économiques,
Université Aristote de Thessalonique, Grèce,
Expert national détaché,
Commission des Communautés européennes,
Direction générale du marché intérieur et des services financiers, **

Cet article a pour objet de mettre au point les déterminants des échanges intra-branche intra-communautaires en fonction de certains aspects spécifiques de la politique industrielle. Il cherche à démontrer en particulier que la convergence des investissements sectoriels suffit à expliquer assez bien les échanges intra-branche. En outre il vise à montrer que des efforts semblables en innovation technologique tendent à favoriser également la même forme d'échanges.

* Je tiens à exprimer mes remerciements à madame le professeur C. Nème de l'Université Paris II, à messieurs K. Mehta et C. Martinez, de la Direction Générale de l'industrie de la Commission de la Communauté européenne et aux référés anonymes de la revue *Economie Appliquée*, pour avoir bien voulu me donner leurs points de vue sur le présent article.

** Cet article ne reflète pas la position officielle de la Commission des Communautés européennes.

This article aims to examine specific determinants of intra-industry trade which are mainly associated with industrial policy. It points out more precisely that same efforts related to sectoral investments could sufficiently explain the intra-industry trade flows. In addition similar efforts regarding the R&D expenditure by sector aim to reinforce the differentiation process and hence this pattern of trade.

INTRODUCTION

La mise en place d'une union douanière implique des gains et des pertes suivant la similitude ou la complémentarité des structures économiques des pays en faisant partie (cf. Byé, 1970, p. 81-90). En effet, l'union douanière entre économies concurrentes, contrairement aux économies complémentaires, tend à promouvoir l'efficacité de leur appareil productif (Leibenstein, 1966), à accentuer le processus de leur spécialisation et enfin à améliorer leur bien-être (Viner, 1950; Lipsey 1957; Bhagwati, 1971). Dans ce cas, les aspects positifs de l'union douanière tels que la «*création d'échanges*» et «*l'effet de réduction du coût*» permettent de contrebalancer toute répercussion négative qui est en principe associée au «*détournement des échanges*» (cf. Corden, 1972).

La libéralisation des échanges au sein de l'Union qui provoque un effet statique lié à la «*création de l'échange*» met également en relief un processus dynamique (cf. Krauss 1972). Dans ce cas, d'après l'expérience acquise, la spécialisation *horizontale* s'accroît par la réduction des variétés produites par chaque firme, et parallèlement par le développement de la production à grande échelle d'un nombre limité de variétés restantes (cf. Balassa, 1963, 1986); une évolution qui tend à intensifier l'échange intra-branche, c'est-à-dire les échanges croisés entre produits finaux proches substitués dans la production et la consommation (différenciation *horizontale*), ou seulement proches substitués dans la consommation (différenciation *verticale*). Le même processus tend à favoriser, par la production de biens intermédiaires dans des établissements différents (spécialisation *verticale*) les échanges croisés de pièces détachées et de composants.

Les structures industrielles similaires des Etats à l'intérieur de l'Union sont susceptibles de se manifester par la croissance de la formation brute de capital fixe (FBCF), ou de la production au profit des mêmes

branches d'activité. Ce qui tend à accentuer encore la spécialisation et les échanges intra-branche.

De tels résultats semblables peuvent apparaître par l'intermédiaire d'une deuxième action, liée cette fois à la technologie. Ainsi, les mêmes efforts en Recherche et Développement (R&D) industrielle tendent à accentuer la différenciation au sein de l'Union et donc à provoquer un processus similaire au précédent.

Les relations proposées dans l'espace communautaire entre les échanges intra-branche et le profil industriel des neuf Etats-membres vont être vérifiées dans le cadre de cette recherche. La Belgique et le Luxembourg sont examinés comme une seule unité. L'année de référence dans le cas de l'analyse statistique est 1987, ou l'année la plus proche pour laquelle les données statistiques sont disponibles.

En particulier, l'évolution des échanges intra-branche dans l'espace communautaire est développée brièvement dans le cadre d'une première section. Ensuite, une deuxième section étudie l'impact des politiques industrielles et de la technologie sur la spécialisation des Etats-membres. Enfin, la répercussion de ces dernières sur les échanges intra-branche est présentée.

I. RAPPEL DES ÉCHANGES INTRA-BRANCHE AU NIVEAU INTRA-COMMUNAUTAIRE

Les indicateurs de Grubel et Lloyd (1975, p. 20-23), cités à l'annexe, sur les échanges intra-branche, permettent de conclure en premier lieu que le commerce intra-branche intra-CEE en produits industriels, bien que relativement étendu entre les neufs Etats-membres, n'a pas enregistré d'évolution frappante de 1978 à 1990. Par contre, durant la décennie précédente (de 1968 à 1978), il avait connu une croissance plus nette (Jacquemin et Sapir, 1986, p. 37).

Dans la période récente, les indices du tableau 1 (ci-après) montrent le taux de croissance modéré du commerce intra-branche dans la plupart des neuf Etats-membres (à l'exception du Danemark et des trois nouveaux adhérents). Ils désignent en outre la réduction de son intensité en Allemagne et au Royaume-Uni, en raison du renforcement de la position concurrentielle de la première et de l'affaiblissement de celle du deuxième.

La quasi-stabilité des indices ci-dessus s'est également accompagnée d'une réduction de la dispersion des indices d'échanges intra-branche

Tableau 1
Echanges intra-branche intra-communautaires ^a

Année Pays	1978 ^b	1980 ^b	1982 ^b	1985 ^b	1987 ^b	1987 ^c	1990 ^c	78-90 Δ max
France	0,75	0,76	0,78	0,78	0,76	0,75	0,76	+ 3
Belgique-Lux.	0,64	0,65	0,67	0,68	0,67	0,67	0,68	+ 4
Pays-Bas	0,64	0,66	0,66	0,64	0,70	0,70	0,71	+ 7
Allemagne ^d	0,74	0,72	0,70	0,70	0,70	0,70	0,68	- 6
Italie	0,58	0,58	0,59	0,59	0,61	0,61	0,62	+ 4
Royaume-Uni	0,75	0,77	0,72	0,71	0,71	0,71	0,69	- 8
Irlande	0,52	0,52	0,52	0,55	0,56	0,57	0,58	+ 6
Danemark	0,48	0,51	0,51	0,55	0,59	0,59	0,60	+ 12
Grèce	0,20	0,20	0,23	0,23	-	0,26	0,28	+ 8
Espagne	-	0,53	-	0,58	-	0,65	0,69	+ 16
Portugal	-	0,33	-	0,38	-	0,40	0,43	+ 10

^a. indicateur ($\bar{B}i$) de Grubel et Lloyd (1975)

^b. CEE des Dix

^c. CEE des Douze

^d. République fFédérale d'Allemagne

Tableau 2
Ecart type (σ) des échanges intra-branche intra-communautaires ^a

	1978	1980	1982	1985	1987	1990
σ	0,105	0,101	0,095	0,082	0,064	0,060

^a. CEE des neuf Etats-membres

($\bar{B}i$) des «neuf» de 1978 à 1990. L'écart type (σ), indiqué au tableau 2 (page suivante), vérifie cette tendance, laquelle pourrait souligner une convergence de leurs structures industrielles. Cette évolution permettrait de mener à la conclusion que l'union douanière entraîne, dans une deuxième phase de son intégration, le rapprochement des structures industrielles de ses membres.

II. LES POLITIQUES INDUSTRIELLES NATIONALES CONVERGENTES ET LEUR RÉPERCUSSION VIS-À-VIS DE LA SPÉCIALISATION

Les études empiriques sur les échanges intra-branche montrent que les facteurs déterminants de ces derniers paraissent dépendre à la fois de la stratégie des entreprises et des mesures visant à orienter le comportement des industriels. Ces mesures, qui s'intègrent au sein de la politique industrielle, cherchent à promouvoir l'expansion de la production, se répercutant explicitement sur toutes les activités au niveau de la firme. Par la voie d'objectifs bien définis en matière d'adaptation des structures ou, alternativement, en mettant l'accent sur des actions globales capables de créer un climat propice au développement industriel, elles s'efforcent d'améliorer l'appareil productif, ce qui se répercute sur les échanges commerciaux.

Ces instruments, associés à la politique sectorielle sélective (ou, en son absence, aux mesures équivalentes d'une politique économique peu volontariste), aux dépenses en Recherche et Développement (R&D), à la politique de concurrence et à la politique commerciale, constituent en dernier ressort des éléments déterminants des échanges intra-branche du point de vue de l'offre. Les deux premières politiques ont un caractère purement offensif, tandis que les deux dernières revêtent un aspect défensif.

Parmi les quatre instruments ci-dessus de la politique industrielle, les actions au niveau sectoriel et sur la Recherche-Développement (R&D) peuvent être considérées comme les principaux facteurs explicatifs des échanges intra-branche. En particulier, la restructuration de l'appareil productif d'un pays par des actions d'ordre sectoriel, en liaison avec les dépenses de technologie, créent un potentiel capable d'infléchir l'importance des tendances protectionnistes de toute mesure d'ordre défensif.

Des actions convergentes dans ce domaine au niveau de l'Union réussissent à créer des avantages comparatifs dynamiques multiples intra-régionaux. En effet, le même choix, entre les États, en matière d'investissements sectoriels et de dépenses de Recherche et Développement (R&D) en vue d'obtenir des structures industrielles compétitives, peut être considéré comme la source primordiale de la différenciation et, par conséquent, des échanges intra-branche.

1. Les politiques d'adaptation structurelles : les contraintes

Dans l'espace communautaire, des approches très diverses ont été adoptées en matière de politiques industrielles. Cette différence conceptuelle dérive de multiples désaccords entre les Etats-membres, sur les mérites de la planification indicative et sur le rôle de la protection (Pearce et Sutton 1985, p. 151-153); ce qui constitue le facteur prédominant empêchant la mise en place d'une politique industrielle commune. L'absence de coordination des politiques industrielles en Europe peut être considérée comme un facteur d'aggravation des crises au sein de chaque économie de la Communauté (CEPII, 1983, chap. V). Les Etats-membres en difficulté ne sont pas capables de trouver, de la part de leurs partenaires, un soutien aussi étendu que celui offert par la politique agricole commune (PAC). Ainsi chacun privilégie la solution qui lui est la plus profitable.

Néanmoins des efforts en commun en matière de politique industrielle auraient probablement pu favoriser une spécialisation traditionnelle dans l'espace communautaire. Par contre, les politiques nationales industrielles, en l'absence de tels efforts, ont créé un tissu productif caractérisé par une spécialisation intra-sectorielle; ce qui résulte des stratégies au niveau des firmes et des contraintes concurrentielles extra- et intra-communautaires. La première contrainte est due à la volonté des firmes multinationales de s'implanter ailleurs, en vue de bénéficier des avantages du pays d'accueil, ou de tourner les obstacles douaniers qui frappent les importations. La tendance à l'internationalisation du capital crée un effet en retour, qui pourrait affecter les pays exportateurs en capital par le bais du commerce international, selon Michalet (1981).

D'autre part, en ce qui concerne les contraintes intra-communautaires, ou mieux les contraintes résultant de la concurrence entre les pays industriels, la pression des demandes mondiale et communautaire (Commission des Communautés européennes, 1980), les avantages fournis par la révolution technologique et l'ouverture progressive du marché intérieur par la suppression des barrières non tarifaires, arrivent à exercer des restrictions analogues sur les choix stratégiques au niveau sectoriel. Les investissements sont concentrés en priorité sur les secteurs qui peuvent concurrencer le défi du grand marché pour satisfaire la demande communautaire et parvenir à incorporer de manière spectaculaire les inventions technologiques. Cette adaptation reflète à son tour «... la qualité de la spécialisation internationale des pays...» selon Rollet (1981 p. 5).

La contrainte extérieure examinée en liaison avec les contraintes intra-communautaires incite par conséquent à la mise en place d'une politique sectorielle plus ou moins sélective ou de mesures semblables pour y faire face, capables de favoriser la croissance de nouveaux secteurs ou de guider les investissements vers les secteurs les plus lucratifs.

L'évolution de la production ou de la formation brute du capital fixe (FBCF) entre les Etats-membres fait preuve depuis 1960 une convergence au niveau sectoriel dans l'espace communautaire; ce qui est susceptible d'expliquer la spécialisation intra-branche des Etats-membres les plus développés, mise en lumière par des taux accrus d'échanges croisés de produits similaires. Des tendances communes en matière d'investissement par secteur caractérisent également les pays développés de l'organisation de coopération et de développement économique (OCDE). (Page, 1980; Fouquin, 1985).

Plus précisément, cette convergence s'explique par le fait que les Etats-membres, surtout les plus développés, concentrent leurs efforts d'investissement dans les mêmes produits. Ainsi les données disponibles sur la production et la formation brute de capital fixe (FBCF) par pays montrent qu'une hausse importante et simultanée a caractérisé les moyens de transport et les machines électriques. Pour les produits chimiques et mécaniques, les augmentations sont moins significatives. En revanche, dans le cas des minerais métalliques ou non, et des produits métalliques, les investissements ont relativement stagné ou fléchi. Les produits alimentaires qui occupaient une position dominante en 1960 ont relativement perdu du terrain. Une baisse plus accentuée est également à signaler pour les produits textiles-habillement-cuirs. Les investissements en machines de bureau n'enregistrent pas de hausse notable, à l'exception de la France, de l'Allemagne et de l'Irlande.

Cette évolution nous permet de déduire que des conceptions différentes en matière de politique industrielle conduisent cependant à des stratégies semblables en matière de spécialisation au niveau communautaire. Ainsi, les Etats-membres tendent à améliorer, rationaliser ou geler leurs efforts d'investissement dans les mêmes secteurs.

En particulier les planifications indicatives *françaises* par la voie de «prévisions» ou de «scénarios» évoluant dans le contexte d'une politique sectorielle traditionnelle (du premier au septième plan) et d'une politique d'innovation technologique (depuis le huitième plan) fournissent des indications semblables à celles de la Belgique et des Pays-Bas ainsi qu'aux politiques allemande, britannique ou italienne.

Ces dernières se trouvent à mi-chemin entre la programmation et le principe de non-intervention.

En effet, aux *Pays-Bas*, à la suite d'une politique d'industrialisation de 1948 à 1960, le pays a introduit une politique industrielle prévisionniste de 1965 à 1977 qui visait, dans le cadre d'une planification décentralisée, à améliorer l'efficacité de l'industrie. Cette action a été suivie depuis 1978 par une politique d'innovation technologique. Dans le même contexte la politique industrielle *belge* visait, par la voie de mesures sélectives de 1970 à 1980, à réaliser des objectifs sectoriels bien définis. Elle se fondait sur trois objectifs : la planification, la conduite d'une politique industrielle active et la décentralisation économique. (Descamps, 1980, p. 21). Depuis 1980 de mesures d'ordre fiscal sont de plus en plus orientées vers l'aide à l'innovation et des nouvelles technologies.

En *Allemagne*, l'industrie a été construite dans un climat qui mettait l'accent sur la fonction vitale de la concurrence et sur une structure en faveur de la cartellisation, depuis 1929. Les divers gouvernements cherchaient à superviser les affaires industrielles plutôt qu'à les contrôler par le truchement d'une concurrence la plus grande possible. L'intervention publique *italienne*, souvent peu rentable, exprimée par le poids important du rôle de l'Etat dans les activités productives de divers secteurs, ne diffère pas beaucoup de celle de la France et du Royaume-Uni, selon Price (1981, p. 65). L'intérêt de l'Etat et sa participation dans les affaires industrielles s'accroissent de 1957 à 1973 parallèlement aux aides sectorielles. A la suite des «opérations de sauvetage» lancées par cette politique durant la période de la crise des années 70, son profil change et une nouvelle politique d'innovation s'établit depuis 1978.

Les changements d'ordre politique au *Royaume-Uni*, qui ont introduit des révisions radicales dans le contexte de la «politique industrielle» (de 1960 à 1970) ou de la «politique d'innovation» (de 1970 à 1980), n'ont pas permis la mise en place d'une structure industrielle solide. «L'anathème politique (...), par les conservateurs sur la planification...» selon Snaks (1977, p. 20) ou l'absence d'une politique libérale cohérente dans le temps, ont réduit l'efficacité de mesures d'ordre structurel. Le ton de cette politique a changé depuis 1980. En effet, depuis cette période, une «politique technologique» par le biais d'une politique de l'offre cherche à créer un environnement peu interventionniste pour les investissements industriels. Au *Danemark*, le gouvernement évite, depuis l'après-guerre, toute intervention étendue dans l'industrie. Il cherche à créer un environnement propice susceptible de faciliter la croissance industrielle. «Les stratégies de politique industrielle

prévisionniste (...) n'ont pas seulement évolué, elles n'ont jamais été considérées dans un agenda...» note Koungstad et Olsen (1985, p. 13). Les politiques industrielles danoises ne privilégient pas de stratégies visant à sélectionner les gagnants ou à sauver les perdants, car elles les considèrent peu efficaces. Enfin, l'idée de base de la politique industrielle *irlandaise* est l'orientation de la production envers les exportations et l'attraction des sociétés étrangères. Le processus de la croissance industrielle a été basé en principe sur le plan à moyen terme débutant en 1958. La politique industrielle vise au moyen d'incitations de plus en plus sélectives à mieux servir les choix stratégiques au niveau sectoriel en faveur des industries de haute technologie, notamment depuis 1980.

En conclusion, l'esprit volontariste partisan de la programmation coexiste donc avec l'esprit libéral dans l'espace communautaire. Ils opèrent tous deux par la voie de mécanismes nationaux divers, sous les mêmes contraintes et les mêmes défis, et parviennent à renforcer les mêmes tendances au niveau sectoriel.

Un mécanisme principal de la politique industrielle, commune à tous les Etats-membres est celui de la politique des achats publics. Cette dernière met en place un traitement similaire en faveur des mêmes produits (Atkins, 1988). En effet, dans ce contexte l'Etat, par un traitement préférentiel, tend à favoriser le développement de productions qui sont en général les mêmes dans les neuf Etats-membres. Cette tendance convergente est une expression représentative de choix communs bien que non coordonnés au niveau des politiques industrielles nationales dans l'espace communautaire

2 . Production, stock de capital et spécialisation

Les exportations ou les importations sont les variables de base convenables pour mesurer la tendance des spécialisations (Balassa, 1965; Rollet, 1976). Néanmoins, on peut aboutir à des résultats très intéressants avec des variables relatives aux investissements ou à la production, utilisées aussi dans le cadre de cette étude, pour explorer le phénomène. En effet, on met en évidence deux séries de spécialisation du type Balassa (1965), la première se référant au niveau de groupes majeurs de produits et la deuxième à un niveau plus désagrégé, c'est-à-dire au niveau de la branche.

En particulier, pour capter les tendances en matière de spécialisation par la voie des investissements ou la production, on part de l'hypothèse

que les entrepreneurs intra-Union sont rationnels, informés, et qu'à l'aide des incitations offertes ils cherchent à améliorer leur position concurrentielle. Ils opèrent dans un environnement où l'entrée est potentiellement libre. Dans ces conditions, la mise en place d'un monopole physique dans une variété donnée ayant un prix élevé ne peut se produire qu'en cas de produits identiques bénéficiant d'une concurrence au niveau de leurs spécifications (différenciation *horizontale*).

En cas contraire, toute firme s'efforce de réduire les possibilités d'entrée d'une nouvelle variété identique à sa propre production par le mécanisme des prix. Selon cette deuxième solution, l'augmentation de la longueur de la série de production en variétés limitées et à prix concurrentiels (spécialisation *horizontale*) a lieu. L'évolution du stock de capital ou de la production peut refléter cette évolution. Ces deux variables peuvent donc être utilisées comme expression de la spécialisation.

Selon cette option, la thèse qu'on cherche à démontrer est la suivante. Les investissements, au sein d'un secteur visent à développer la production, à grande échelle d'un nombre limité de variétés à meilleur marché, par firme/pays. L'accentuation de ce processus dans une union douanière implique un accroissement des variétés intra-Union. De ce fait, la spécialisation *horizontale* est la source des échanges intra-branche en variétés, différentes en qualité et en prix (différenciation *verticale*). En cas de différenciation *horizontale*, les investissements par firme visent à assurer une position de monopole physique à une variété donnée, en quantités suffisantes, pour satisfaire les besoins accrus dans l'Union; évolution qui aboutit également à des échanges intra-branche entre variétés de qualité similaire.

Il reste à préciser la liaison entre les investissements ou la production et les échanges intra-branche. D'après les hypothèses proposées, on s'attend à des échanges intra-branche intenses en cas de flux d'investissements sectoriels ou de productions similaires entre firmes/pays de l'Union; ou encore, en termes de spécialisation, on prévoit des échanges intra-branche élevés dans les secteurs bénéficiant d'une spécialisation moyenne. De ce fait, les répercussions des politiques en matière de réformes sectorielles présentent un grand intérêt.

En conséquence, la formule de Balassa (1965) a été utilisée, de manière toutefois transformée. En particulier, des variables correspondant au stock du capital et à la production ont été introduites, à la place des variables exprimant les flux d'exportations par produits et pays,

selon la version de Balassa (1965). De ce fait l'indicateur à utiliser revêt la forme suivante :

$$SI = \frac{\frac{A_{ij}}{A \cdot j}}{\frac{A_{icee}}{A \cdot cee}} \quad (1)$$

où

A_{ij} = le stock du capital ou la production du produit (i) du pays (j);

$A \cdot j$ = le stock du capital ou la production totale des produits manufacturés (·) du pays (j);

A_{icee} = le stock du capital ou la production, du produit (i) de la Communauté économique européenne (CEE);

$A \cdot cee$ = le stock du capital ou la production totale des produits manufacturés (·) de la Communauté économique européenne (CEE).

Si cet indicateur est entre un et deux, la spécialisation est plus intra-branche. Un indice supérieur à deux ou proche de zéro montre une tendance en faveur d'une spécialisation inter-branche ou l'absence de spécialisation.

Pour mesurer la spécialisation d'une manière adéquate, on doit utiliser les données statistiques les plus désagrégées possibles. Néanmoins, les contraintes relatives aux données ne permettent pas d'établir une série d'indices de spécialisation à un niveau très fin. De ce fait, en fonction des données disponibles, nous présentons des résultats à deux niveaux d'agrégation. En premier lieu, l'analyse se situe au niveau de douze grands secteurs par la voie des données sur le stock du capital. En deuxième lieu, pour éviter toute illusion statistique provenant des choix ci-dessus, l'analyse se situe sur 110 branches d'activité industrielle classées selon le code NACE et appartenant aux douze secteurs considérés. Le manque de données analytiques sur les investissements au niveau de la branche explique le choix de la deuxième série d'estimations basées sur les données de la production.

Cette dernière série fournit d'une manière plus précise le profil des spécialisations sans toutefois aboutir à des conclusions contradictoires par rapport à celles résultant de la première série d'indices. De ce fait, on analyse la spécialisation selon des méthodes alternatives suffisamment compatibles.

3. L'interaction entre la spécialisation et les échanges intra-branche intra-communautaires

L'élaboration du phénomène des spécialisations, au niveau des douze secteurs choisis ainsi qu'au niveau de 110 branches, montre clairement la répercussion des politiques sectorielles nationales, non coordonnées par une politique commune, sur la spécialisation intra-branche. Ce qui explique à son tour le niveau relativement élevé des échanges intra-branche intra-CEE, notamment dans le cas des spécialisations moyennes. Plus précisément, en ce qui concerne le profil de la spécialisation considérée et ses répercussions sur les échanges intra-branche intra-communautaires, le tableau 3, construit à l'aide de données statistiques se référant aux douze grands secteurs choisis, met en lumière les faits suivants :

Tableau 3
Indicateurs de spécialisation (SI), par grand secteur ^a

Produits	France	Belgique	Pays-Bas	Allem. ^b	Italie	R.-Uni	Irlande	Danemark
M/c bureau Prod. trait.info. instr. de préc. et d'optique	1,85	1,153	1,39	1,18	0,634	0,635	1,19	0,756
Produits en métaux	1,15	<i>0,541</i>	0,75	1,21	0,90	0,781	0,651	0,857
Produits alimentaires	1,12	1,131	1,54	0,930	0,747	1,11	2,14	1,61
Moyens de transport	1,10	<i>0,56</i>	<i>0,487</i>	1,07	1,03	0,862	<i>0,237</i>	0,683
Min.non métalliques	1,00	1,13	0,77	0,884	1,315	0,942	1,90	1,07
Papier/prod. en papier	1,06	0,948	1,154	0,965	0,793	1,078	0,640	1,37
Min.et métaux ferreux	0,936	1,55	<i>0,545</i>	0,785	1,223	1,062	<i>0,419</i>	<i>0,288</i>
Prod.text.-cuir-chaussures et habillement	0,919	1,212	0,684	0,848	1,631	0,938	1,45	0,81
Prod. en caoutchouc/plast.	0,893	0,837	<i>0,548</i>	1,08	1,113	0,972	<i>0,549</i>	0,707
Mat.élect.et électronique	0,848	0,832	1,30	1,303	0,86	0,924	0,596	0,650
Produits chimiques	0,711	1,25	1,64	0,918	1,119	1,20	0,95	0,724
M/c indust. et agricoles	0,681	<i>0,552</i>	<i>0,55</i>	1,24	0,844	1,177	<i>0,308</i>	1,41

^a. du type de Balassa (1965)

^b. République fédérale d'Allemagne

– Premier constat : un certain nombre d'Etats-membres, à savoir la Belgique, les Pays-Bas, l'Irlande, le Danemark, détiennent des positions

relativement fortes ou faibles dans peu d'activités ($SI \geq 1,41$ et $SI \leq 0,56$); à l'inverse, les indicateurs de spécialisation des autres Etats-membres, à savoir l'Allemagne, la France, le Royaume-Uni et l'Italie, sont moins dispersés (SI est autour de l'unité). Ce qui est susceptible de signaler, en principe, l'absence de positions dominantes sélectives.

En particulier, l'écart type (σ) construit à travers des spécialisations au niveau de 110 branches d'activités industrielles du code NACE fournit des informations plus précises sur cette tendance. En effet, les indices sur l'écart type, présentés au tableau 4, montrent également que la dispersion des indicateurs de spécialisation est plus élevée dans le cas du premier groupe de pays que dans le cas du second.

Tableau 4

Écart type (σ) des indicateurs de spécialisation (SI) intra-branche ^a

France	Belgique	Pays-Bas	Allemagne ^b	Italie	R.-Uni	Irlande	Danemark
0,39	0,87	0,88	0,37	0,67	0,45	1,02	1,12

^a sur 110 branches

^b République fédérale d'Allemagne

Par ailleurs, le tableau 5 (ci-après), qui retrace les positions les plus fortes de spécialisation par pays/branche, permet de conclure que même à un niveau plus agrégé les positions dominantes sont très limitées, ce qui caractérise, en principe, les pays disposant d'écart types relativement élevés.

– Deuxième constat : les indices sur la spécialisation, et leur interaction avec les coefficients des échanges intra-branche à deux niveaux d'agrégation choisis montrent par ailleurs que les échanges intra-branche modérés ou bas correspondent à des positions de spécialisations fortes ou faibles. En outre, cette interaction précise que les valeurs relativement plus élevées des échanges intra-branche correspondent à des indicateurs de spécialisation proches de l'unité.

En effet le tableau 6 (ci-après), construit sur des données peu agrégées correspondant aux grands secteurs examinés, confirme ce qu'on anticipait. D'autre part, le tableau 7 (ci-après), construit toutefois à l'aide de données d'un niveau d'agrégation plus fin correspondant aux 110 activités du code NACE, mène également à des conclusions

Tableau 5
Indicateurs de spécialisation. Positions fortes ($SI \geq 2.00$)

NACE Produits	Fran.	Belg.	P.-Bas	Allem. ^a	Ital.	R.-Uni	Irl.	Dan.
221 Ind. de fer acier		2,4						
223 Prod. laminés		2,9						
224 Prod. mét.non ferreux		2,0						
243 Chaux, ciment							3,1	
245 Travaux prod. non mét.						2,5		
248 Man. de prod. céramiq.					2,0			
255 Peintures, vernis			2,0	2,4				
257 Prod. pharmaceutiques							2,8	
259 Prod.chem.usagé ménag.			2,7	2,2				
316 Outil. prod. finis en mét.								2,6
330 Machines de bureau							5,3	
342 Matériel élect. d'équip.				2,3				
345 Equip. électronique	2,0							
348 Assembl. d'équip. élect.					2,5			
351 Véhicules automobiles					2,3			
361 Construction navires			2,4					3,9
363 Motocyclettes					2,4			
364 Const. aéronotique	2,1					2,3		
365 Autre équip.de transp.			4,5					
372 Equip. médical				2,4				
374 Horlogerie	2,1							
411 Corps gras vég. anim.		2,2	2,5					
412 Abattage, prép.viande							3,3	3,3
413 Produits laitiers							3,4	2,3
415 Préparat.de poisson								6,4
417 Pâtes alimentaires					3,5			
421 Chocolat-confiserie							3,0	
422 Nourriture animaux		3,0						
429 Préparation tabac						2,0		
431 Industrie lainière					2,7			
432 Industrie de soie					2,0			
434 Prép.de lin, chanvre.		2,6						
436 Bonneterie					2,2			
437 Prod. finis de textile					3,1			
438 Tapis		5,5				3,1		3,6
441 Tannerie					3,1			
451 Indus. de chaussures					2,5			
456 Fourrure		2,3						
462 Préparation de bois		2,5						
463 Prép. bois pour sol							2,6	
491 Bijoux		4,4						
492 Instrum. de musique					2,2			
494 Jouets et art. de sport								5,4

^a République fédérale d'Allemagne

semblables par Etat-membre examiné. Le même tableau signale en outre l'impact significatif de trends en faveur des spécialisations suffisamment dispersées (cas de Irlande et du Danemark en principe, dont l'écart-type est supérieur à l'unité) sur les échanges intra-branche; ce qui est traduit par leur niveau suffisamment faible correspondant aux positions fortes et faibles de spécialisation.

La thèse selon laquelle le comportement rationnel des entrepreneurs cherche à créer un monopole physique pour une variété donnée (concurrence par la spécification) ou à mettre l'accent sur des économies d'échelle pour des variétés limitées (concurrence par le prix) conduit aux conclusions suivantes quant à l'interaction des indicateurs de spécialisation et des échanges intra-branche intra-communautaires.

Tableau 6
**Intéraction entre l'indicateur de spécialisation (SI)
et celui des échanges intra-branche ($\bar{B}i$)
au niveau des grands secteurs dans l'ensemble
de neuf Etats-membres de la CEE**

	$SI \geq 1,41^a$	$0,55 < SI < 1,41^b$	$SI \leq 0,55^c$
$\bar{B}i_{CEE9}$	0,601	0,681	0,474
$\bar{B}i_{CEE}^d$	0,683	0,692	0,652

a. Positions de (SI) fortes correspondant à dix grands secteurs

b. Positions de (SI) moyennes correspondant à soixante-quatorze grands secteurs

c. Positions de (SI) faibles correspondant à dix grands secteurs

d. Moins l'Irlande et le Danemark

En particulier, la stratégie d'investissement au sein des neuf Etats-membres a favorisé un tissu de spécialisations intra-branche. Dans peu de cas, l'indicateur de spécialisation de type Balassa (1965) est supérieur à 2, ce qui est susceptible de prouver l'absence de positions dominantes étendues. De ce fait, la vague des spécialisations intra-communautaires n'a pas réussi à créer de fournisseurs dominants, capables de promouvoir des échanges intersectoriels.

Néanmoins, malgré cette homogénéisation du marché communautaire par des mouvements convergents de spécialisation, des différences partielles se sont manifestées entre les Etats-membres, notamment pour l'Irlande et le Danemark.

Tableau 7
**Intéraction entre les indicateurs de spécialisation (SI)
 et des échanges intra-branche (\bar{B}_i) par Etat-membre examiné
 au niveau de 110 branches d'activité industrielle**

	Positions (SI) fortes ^a	Positions (SI) moyennes ^b	Positions (SI) faibles ^c
\bar{B}_i	France		
	2,17 < SI < 1,57 0,772	1,57 < SI < 0,47 0,774	SI < 0,47 0,715
\bar{B}_i	Belgium		
	5,52 < SI < 2,04 0,619	2,04 < SI < 0,29 0,797	SI < 0,29 0,590
\bar{B}_i	Pays-Bas		
	4,59 < SI < 1,70 0,637	1,70 < SI < 0,50 0,772	SI < 0,50 0,674
\bar{B}_i	Allemagne		
	2,4 < SI < 1,60 0,636	1,20 < SI < 0,54 0,653	SI < 0,54 0,604
\bar{B}_i	Italie		
	3,51 < SI < 2,07 0,396	2,07 < SI < 0,54 0,648	SI < 0,54 0,580
\bar{B}_i	Royaume-Uni		
	3,13 < SI < 1,68 0,650	1,68 < SI < 0,54 0,721	SI < 0,54 0,648
\bar{B}_i	Irlande		
	5,30 < SI < 1,01 0,492	1,01 < SI < 0,41 0,725	SI < 0,41 0,517
\bar{B}_i	Danemark		
	6,48 < SI < 1,82 0,463	1,82 < SI < 0,23 0,746	SI < 0,23 0,406

^a. Correspondant aux dix branches qui enregistrent les indicateurs (SI) les plus élevés par Etat-membre

^b. Correspondant aux quatre-vingt-dix branches qui enregistrent des indicateurs (SI) moyens

^c. Correspondant aux dix branches qui enregistrent les indicateurs (SI) les moins élevés par Etat-membre.

En général les choix concernant la production et les investissements sectoriels ont créé, entre les neuf États-membres de la CEE, des structures favorables aux spécialisations et aux échanges intra-branche. De même, les efforts en Recherche et Développement (R&D) dégagent aussi clairement cette tendance.

III. EFFORTS CONVERGENTS EN INNOVATION TECHNOLOGIQUE

La politique d'innovation technologique s'efforce d'orienter la recherche vers les secteurs considérés comme prioritaires. Elle privilégie en principe les activités qui se trouvent au centre d'intérêt de la politique industrielle, ou mieux de la politique sectorielle en vigueur.

De ce fait, la politique d'innovation technologique est cohérente avec la politique sectorielle ou avec toute action d'effet équivalent. Elle semble donc faire partie des politiques visant à restructurer le tissu industriel. Les plans français depuis 1980, le plan belge à moyen terme (1971-1975), la loi 675 italienne de 1983, la politique de l'Autorité pour le développement industriel (IDA) irlandais, les actions du Département du commerce et de l'industrie (DTI) et surtout de l'ancien Conseil national de l'entreprise (NEB) du Royaume-Uni, etc., permettent d'illustrer ce point de vue.

Les gouvernements des États-membres, par diverses mesures d'envergure différente, parviennent toutefois à des résultats exprimant une convergence au niveau de la Recherche et Développement (R&D) industrielle, et plus particulièrement au niveau de la stratégie commune en innovation technologique sectorielle. En particulier les gouvernements financent la recherche technologique par des mesures d'ordre fiscal ou par des crédits et subventions ou des taux d'intérêt réduits, par des recherches communes, par des contrats de recherche avec les instituts privés ou publics (ces derniers servant en principe aux petites et moyennes entreprises-PME).

A titre d'exemple en matière de mesures fiscales (telles que les amortissements accélérés, la possibilité de constituer des réserves pour investissement en exonération d'impôts, etc.), *L'Allemagne* montre une préférence marquée pour les aides indirectes plutôt que directes, considérant qu'elles sont plus compatibles avec le principe d'économie sociale du marché. Par ailleurs, les crédits et les subventions constituent

une batterie alternative d'instruments efficaces en matière d'objectifs spécifiques.

En *France*, bien que la politique des aides-fonds d'emprunts ait été remise en cause au cours du septième plan, pour la chimie par exemple, et que les mesures d'effet indirect, telles que les déductions fiscales sur l'accroissement des dépenses de recherche ne soient pas exclues, les aides directes constituent en fait l'instrument majeur de la politique d'innovation du pays. En *Belgique*, des prêts sans intérêt peuvent être accordés pour la mise au point de prototypes ou de procédés de fabrication de nouveaux produits. Aux *Pays-Bas*, un système de taux d'intérêt particulier, des mesures fiscales et des subventions sont prévus pour stimuler la recherche industrielle. Au *Royaume-Uni*, la politique d'innovation de 1970 et la politique technologique des années 80 (par la voie de subventions favorisant les innovations spécifiques et les efforts visant à stimuler la diffusion d'une technologie particulière) cherchaient à encourager les investissements en faveur des secteurs de haute technologie. Au *Danemark*, un programme de développement technologique est en vigueur afin de convaincre les entreprises d'utiliser la technologie de l'information. Cette mesure traduit la petite taille des entreprises danoises et la faiblesse qui en découle. Néanmoins, aucun support important n'est prévu en faveur des industries modernes. Une assistance générale est offerte aux entreprises qui cherchent à introduire les nouvelles technologies. Enfin *l'Irlande*, et plus précisément l'Autorité pour le développement industriel (IDA), prévoient des subventions en faveur des firmes pour favoriser des projets spécifiques de recherche. En outre, par le biais de sa politique de «chasse aux gagnants», elle offre des aides accrues aux entreprises les plus dynamiques.

1. Relation entre la politique d'innovation technologique et les échanges intra-branche

L'exposé de la section précédente relatif à la politique des réformes structurelles a permis de dégager des tendances similaires quant aux choix en matière d'investissements industriels aboutissant en conséquence à une spécialisation intra-branche avec peu de positions dominantes.

Dans le cadre de cette stratégie, les investissements en produits de haute technologie sont privilégiés de manière semblable dans tous les Etats-membres, phénomène clairement exprimé par les diverses formes

de politiques d'innovation technologique. Ces dernières, selon des plans indicatifs, ou selon une politique de l'offre, mettent principalement l'accent sur les produits électroniques, les ordinateurs, la robotique, la pharmacie et les produits électriques.

La tendance convergente au sein des neuf Etats-membres, par le renforcement synchrone des dépenses en Recherche et Développement (R&D) pour les mêmes groupes de produits, crée des avantages technologiques multiples en faveur de ces produits par le biais de la différenciation. Ainsi les échanges intra-branche sont susceptibles de se développer.

Cette tendance limite la création d'«écarts technologiques» isolés par pays/produit du type de Posner (1961), susceptibles de favoriser des échanges intersectoriels. Le tableau 8 permet de vérifier la répercussion des mêmes choix en matière de dépenses technologiques sur les échanges intra-branche. L'indicateur (R) du tableau relatif à l'intensité technologique exprime le rapport des dépenses de Recherche et Développement (R&D) à la production (Q) de chaque grand secteur examiné. L'indicateur (\bar{B}_i) est celui de Grubel et Lloyd (1975).

L'absence des données statistiques analytiques, quant aux dépenses technologiques au niveau de la branche, n'a pas permis de construire des variables (R) plus précises. La même contrainte, même à un niveau peu agrégé, n'a pas permis la construction de la variable considérée pour les Pays-Bas.

Il est clair que les indicateurs (R) les plus élevés, pour les cinq Etats-membres industriels, en principe, correspondent aux mêmes produits, c'est-à-dire aux produits pharmaceutiques, électroniques, aux machines de bureaux et aux produits électriques. Au deuxième rang, se situent les instruments de précision et les véhicule routiers ainsi que les produits chimiques. Les machines spécialisées pour l'industrie et les produits en caoutchouc et en plastique suivent, puis viennent les métaux non ferreux, les produits métalliques et les métaux ferreux, de même que les produits dérivés du pétrole. Le même choix des partenaires communautaires caractérise également les produits à faible technologie tels que les textiles-vêtements, les produits alimentaires, les minéraux non métalliques et les produits en papier et en bois.

La dernière colonne du tableau 8 montre la moyenne des indicateurs \bar{B}_i d'échanges intra-branche des Etats-membres examinés. Des valeurs \bar{B}_i élevées correspondant aux secteurs prioritaires en R&D visent à illustrer l'écart technologique étendu au niveau communautaire. De même, les choix stratégiques quant aux produits de technologie moyenne

Tableau 8
Indicateur d'intensité technologique (R)^a
et des échanges intra-branche (\overline{Bi})^b pour huit Etats-membres
de la Communauté économique européenne^c

PRODUITS	France	Belgique	Allem.	Italie	R.-Uni	Irlande	Danemark	\overline{Bi}
Produits pharmaceutiques	6,69	10,50	nd	7,18	18,1	1,30	7,14	0,87
Equipement électronique	10,90	15,90	8,35	6,29	13	6,81	3,54	0,73
M/c de bureaux ordinateurs	4,94	0,34	nd	5,06	11,3	nd	nd	0,81
Produits électriques	1,83	6,13	4,89	2,47	5,66	1,76	2,03	0,74
Inst. de précision/d'optique	3,07	2,4	3,26	1,47	4,31	1,01	17,4	0,68
Véhicules routiers	2,17	0,43	3,11	2,88	3,78	1,85	nd	0,61
Produits chimiques	2,80	3,15	5,58	1,71	3,95	0,83	1,62	0,66
M/c pour l'industrie	1,52	1,81	2,95	0,95	1,75	1,15	1,76	0,65
Prod. en caoutchouc/plast.	1,63	0,66	1,17	0,75	0,57	0,52	0,32	0,77
Métaux non ferreux	1,42	0,62	0,75	0,55	0,70	0,55	nd	0,75
Produits métalliques	0,31	1,17	0,87	0,62	0,45	0,99	nd	0,70
Métaux ferreux	0,91	0,80	0,69	0,46	0,71	0,19	nd	0,78
Dérivés du pétrole	0,60	nd	0,24	0,24	0,39	nd	nd	0,50
Textiles-vêtements	0,26	0,49	0,28	0,02	0,18	0,44	0,06	0,66
Produits alimentaires	0,22	0,25	0,21	0,12	0,33	0,23	0,25	0,56
Minéraux non métalliques	0,69	1,07	0,78	0,11	0,53	0,30	0,64	0,64
Produits en papier	0,13	0,30	0,15	0,01	0,22	0,19	0,02	0,74
Bois et meubles	0,07	0,33	0,42	0,04	0,06	0,26	0,12	0,59

^a. L'indicateur d'intensité technologique (R) est construit par le rapport des dépenses en Recherche et Développement (R&D) à la production (Q) par secteur.

^b. \overline{Bi} : indicateur de Grubel et Lloyd (1975).

^c. Des statistiques analytiques en dépenses technologiques ne sont pas disponibles pour les Pays-Bas

^d. République fédérale d'Allemagne

nd : Données non disponibles.

sont susceptibles de fournir des explications adéquates, compte tenu de l'absence de positions technologiques dominantes.

Enfin, la figure (ci-après) désigne l'absence d'écart technologique posnérien qui aurait pu favoriser le développement des échanges intersectoriels en cas de positions dominantes technologiques sélectives par pays/produit. Dans ce cas, nous devrions anticiper des indicateurs Bi bas, notamment en position de fortes dépenses technologiques.

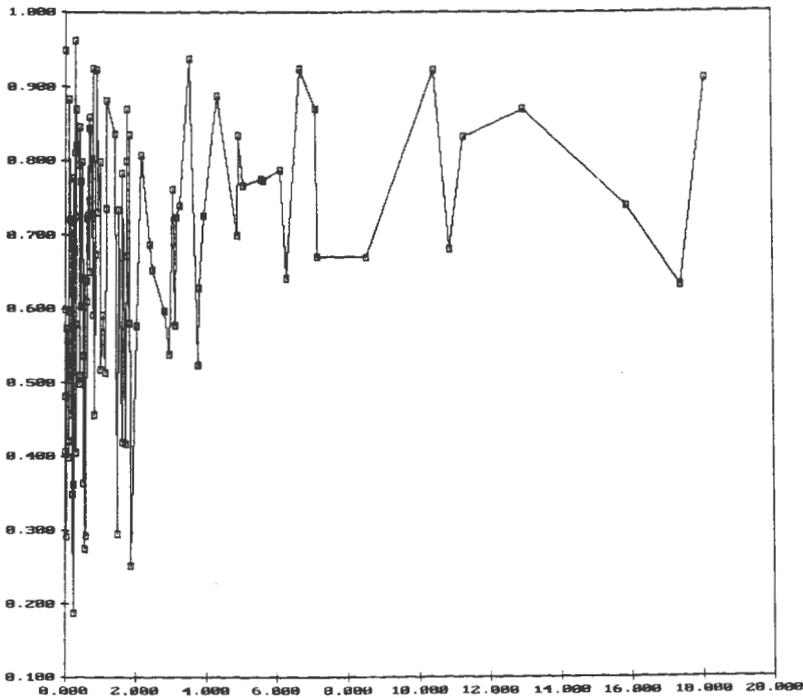


Figure
**Intéraction entre 115 indicateurs d'intensité technologique (R)
 et des échanges intra-branche (B_i)^a.
 Le cas de huit Etats-membres de la CEE^b**

^a. L'indicateur d'intensité technologique (R) est construit par le rapport des dépenses en Recherche et Développement (R&D) à la production (Q) par secteur examiné. B_i est l'indicateur des échanges intra-branche de Grubel et Lloyd (1975)

^b. Pays-Bas exclus par manque de statistiques analytiques sur les dépenses en Recherche et Développement (R&D)

La figure a été construite par l'interaction entre les 115 indicateurs d'intensité technologique (R) du tableau 8 et les variables respectives des échanges intra-branche (B_i) pour les mêmes produits et pays. Le degré de corrélation entre les deux variables, positif, mais relativement modéré, $R^2 = 0,278$ (significatif avec un risque d'erreur de 1%), conduit enfin à la conclusion suivante : les dépenses technologiques similaires dans un environnement qui accentue la différenciation créent un tissu industriel favorable aux échanges intra-branche. Ce dernier est

suffisamment intense pour les produits à haute technologie, mais non négligeable pour les biens de technologie moyenne.

CONCLUSIONS

La présente recherche visait à mettre au point les déterminants des échanges intra-branche intra-communautaires spécifiques au sein de neuf Etats-membres de la CEE, ou, selon le terme récemment introduit par le traité de Maastricht (1992), de l'Union européenne. On a cherché à démontrer que des efforts convergents en investissements sectoriels accentuent les échanges intra-branche. Ces efforts, qui créent un tissu de spécialisation intra-branche dans l'espace communautaire, ne permettent pas le développement de fournisseurs dominants forts. Les indicateurs de spécialisation du type de Balassa (1965) construits à partir de deux séries de données statistiques – sur les investissements et la production – ont permis de défendre ce point de vue. Cette évolution permet d'expliquer en partie le niveau élevé des échanges intra-branche intra-communautaires.

Sur le plan technique, les échanges examinés sont suffisamment élevés dans la majorité des produits. Ces derniers bénéficient d'un indicateur de spécialisation proche de l'unité; ce qui illustre clairement la convergence au niveau de la production. Enfin, l'évolution décrite permettrait de déduire que les Etats-membres, par les spécialisations *horizontales* ou *verticales*, ont privilégié une structure d'échanges en produits différenciés ou intermédiaires, fabriqués en quantités suffisante pour qu'ils soient exportés.

L'influence du progrès technique dans les échanges intra-branche intra-communautaires fut le deuxième aspect examiné. On s'est efforcé d'étayer la thèse selon laquelle des efforts semblables en innovation technologique accentuent les échanges croisés des mêmes produits. L'évolution des dépenses en innovation technologique dans l'espace communautaire montre que les neuf Etats-membres préfèrent, en général, innover dans les mêmes activités industrielles. Ce qui tend à conduire à la multiplication de nouveaux produits, à la croissance de la différenciation dans l'espace communautaire et par conséquent à l'intensification des échanges intra-branche.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Atkins W. S., *The Cost of Non-Europe in Public Sector Procurement*, Brussels : Commission of the EC, vol. 5, 1988.
- Balassa B., *Theory of Economic Integration*, Homewood Illinois : Irwin, 1961.
- Balassa B., «European Integration : Problems and Issues», *American Economic Review*, Papers and Proceeding, vol. 53, May 1963, p. 175-184.
- Balassa B., «Trade Liberalisation and Revealed Comparative Advantage», *The Manchester School of Economics and Social Studies*, vol. 33, May 1965, p. 99-123.
- Balassa B., «The Determinants of Intra-industry Specialisation in United States Trade», *Oxford Economic Papers*, vol. 38, June 1986, p. 220-233.
- Bhagwati J. N., «Trade-Diverting Custom Unions and Welfare Improvement : A clarification», *Economic Journal*, vol. 81, September 1971, p. 580-587.
- Byé M., *Les problèmes économiques européens*, Paris, Cujas, 1970.
- CEPEII – Centre d'études prospectives et d'informations internationales, *L'économie mondiale : la montée des tensions*, Paris, Economica, 1983.
- Commission des Communautés Européennes, *La spécialisation des douze pays européens avant et après 1973*, Série Fast, n°1, Direction générale XII, Bruxelles, 1980.
- Commission des Communautés Européennes, «L'impact sectoriel du marché intérieur sur l'industrie : les enjeux pour les Etats-membres», *Economie européenne*, N° spécial, novembre 1990.
- Corden W. N., «Economies of Scale and Custom Union Theory», *Journal of Political Economy*, vol. 80, March-April 1972, p. 465-475.
- Décamps I., *Planification et décentralisation économiques en Belgique*, Bruxelles, U.G.A., 1980.
- Fouquin M., «Modalité de restructurations industrielles», *Revue d'économie industrielle*, n° 31, 1er trimestre 1985, p. 21-32.
- Grubel H. G., Lloyd P.J., *Intra-industry Trade*, London, MacMillan, 1975.
- Jacquemin A., Sapir A., «La part des échanges intra-CEE dans le commerce communautaire : une analyse sectorielle», *CEPS Working Document*, May 1986.
- Koungstad J., Olsen O.J., *Danish Industry Policy*, Institute of Economics and Planning, Roskilde, Roskilde University Center, 1985.
- Krauss M. B., «Recent Developments in Customs Union Theory : An Interpretive Survey», *Journal of Economic Literature*, vol. 10, n° 2, 1972, p. 413-463.
- Leibenstein H., «Allocative Efficiency vs X-Efficiency», *American Economic Review*, vol. 56, June 1966, p. 392-414.
- Lipsey R. E., «The Theory of Customs Unions : Trade Diversion and Welfare», *Economica*, vol. 24, February 1957, p. 40-46.
- Michalet C.A., «Une nouvelle approche de la spécialisation internationale», *Revue d'économie industrielle*, n°17, 3e trimestre 1981, p. 61-75.
- Page J.-P., «Les structures industrielles des pays de l'OCDE face à la crise», *Revue d'économie industrielle*, n° 13, 3e trimestre 1980, p. 1-22.
- Price V. G., *Industrial Policies in the European Community*, London, MacMillan Press, 2^e éd, 1981.
- Pearce J., Sutton J., *Protection and Industrial Policy in Europe*, London : Routledge & Kegan Paul, 1985.

- Posner M. V., «International Trade and Technical Change», *Oxford Economic Papers*, vol. 13, October 1961, p. 323-341.
- Rollet P., «Les phénomènes de spécialisation à l'intérieur de la Communauté européenne : Le cas des produits manufacturés et des machines», *Revue du Marché commun*, n° 96, mai 1976, p. 251-263.
- Rollet P., *Spécialisation internationale et orientation géographique du commerce extérieur : une analyse des phénomènes de spécialisation bilatérale de la Communauté européenne à Neuf; de la France, de l'Allemagne, du Royaume-Uni, des USA et du Japon, de 1963 à 1979*; Rapport effectué pour la Commission des Communautés européennes, Bruxelles, réf. II/107/82-Fr., 1981.
- Snaks M., *Planning and Politics : the British Experiences 1960-1976*, London : G. Allen & Unwin, 1977.
- Viner J., *The Customs Union Issue*, New York : Camerle Endowment for International Peace (ed), London, Steven & Sons, 1950.

ANNEXE

L'indicateur de Grubel et Lloyd (1975, p. 20-24) sur les échanges intra-branche pour une branche (i) est tel que :

$$Bi = \frac{(Xi + Mi) - |Xi - Mi|}{Xi + Mi} \quad 0 \leq Bi \leq 1$$

Plus l'indicateur (Bi) est proche de 1, plus le courant des échanges intra-branche est élevé.

Le même indicateur pour (n) branches revêt la forme suivante :

$$Bi = \frac{\sum_{i=1}^n [(Xi + Mi) - |Xi - Mi|]}{\sum_{i=1}^n (Xi + Mi)}$$

Xi = les exportations de la branche (i)

Mi = les importations de la branche (i)

Monnaie, « finance » et intérêt

Jean-Luc Bailly

Centre d'études monétaires et financières
Université de Bourgogne

Dans le présent article nous vérifions qu'il ne peut y avoir d'intérêt purement monétaire au sens où on l'entend habituellement, c'est-à-dire portant sur une pure monnaie bancaire. La raison première en est que la monnaie n'est pas un actif net, et n'est, de ce fait, pas porteuse d'intérêt. Pourtant, cela ne signifie pas qu'il ne puisse se former un intérêt net destiné à rémunérer les banques prêteuses de liquidités qu'elles ont elles-mêmes créées. L'introduction du concept de « finance » dans le champ de l'analyse permet de saisir le processus par lequel se forme, dans l'économie, un capital liquide qui reste la propriété des banques. Dès lors, il devient possible de comprendre que ce qu'il est convenu d'appeler « l'intérêt monétaire » servi aux banques entre, en vérité, dans la catégorie des coûts macroéconomiques de production et qu'il est donc net pour l'ensemble de l'économie, banques comprises.

The purpose of this paper is to show that there is no such thing as a pure money interest in the usual meaning of the phrase, that is, interest bearing on pure bank money. Basically, money cannot bear interest because it is not a net asset. This does not however imply that banks cannot charge interest on liquidities they have created. The concept of « finance » brings to light the process through which a liquid capital owned by the banks is formed in the economy. It then becomes clearer that what is commonly called « money interest » paid to banks is in fact a macroeconomic cost of production and therefore a net asset for the whole of the economy, including the banks.

INTRODUCTION

Le principe de la demande effective repose sur l'idée selon laquelle la production est initialement financée sans l'apport d'une épargne préalable, mais grâce à la monnaie émise par le système bancaire. Selon les théories post-keynésiennes du circuit, la monnaie, qui résulte de «la création d'un flux de crédit demandé par les entreprises» (M. Lavoie, 1987, p. 71), est endogène et n'a d'autre origine que les avances effectuées par les banques à l'économie. En conséquence, les banques étant supposées prêter la monnaie à l'économie, on en déduit qu'elles doivent recevoir un *intérêt* purement monétaire, dont le montant est proportionné à la quantité de monnaie qu'elles ont mise en circulation. La définition du taux d'intérêt de la monnaie, donnée par Keynes, ne semble pas contredire ce point de vue. Il écrit en effet : «Le taux de l'intérêt de la monnaie, rappelons-le au lecteur, n'est rien d'autre que le pourcentage d'excès d'une certaine somme de monnaie stipulée livrable à terme, par exemple un an plus tard, sur ce que nous pouvons appeler le prix comptant ou "cash" de ladite somme livrable à terme.» (Keynes, 1971b, p. 229.) L'intérêt monétaire serait donc constitué d'une certaine quantité de monnaie qui viendrait s'ajouter aux sommes initialement prêtées par le système bancaire. Si l'on retient cette définition on peut dire alors que l'intérêt monétaire (i) porte sur la monnaie créée et prêtée par les banques, (ii) consiste en un surplus de monnaie qui gonfle la quantité de monnaie initialement lancée.

Dans le chapitre XVII de la *Théorie générale*, Keynes présente le taux d'intérêt de la monnaie comme la mesure du rendement d'une «richesse durable». Le concept de taux d'intérêt spécifique de la monnaie ne serait alors qu'une «extension de celui d'efficacité marginale du capital» (P. Lévy, 1991, p. 156). A. Barrère relève en effet (A. Barrère, 1990a, p. 302) que, en plaçant le taux d'intérêt de la monnaie sur le même plan que les autres taux d'intérêt spécifiques, Keynes donne à la monnaie les attributs d'un actif et fait de l'intérêt monétaire un «concurrent» des intérêts propres aux autres actifs. De son côté, P. Norel note que «dans un premier temps Keynes définit le concept de taux d'intérêt spécifique et l'utilise pour homogénéiser la monnaie aux autres richesses» (P. Norel, 1990, p. 61). Or une telle démarche, contradictoire avec les conceptions de la monnaie antérieurement présentées par Keynes dans *A Treatise on Money*, nous paraît logiquement fondée sur un glissement conceptuel confondant monnaie et liquidité. Dans la *Théorie générale* la notion de liquidité est assez mal définie puisqu'elle s'applique tantôt à la monnaie,

tantôt à la forme monétaire du revenu. Partant, les théories de l'intérêt monétaire qui s'appuient sur la notion de monnaie «richesse durable», telle qu'elle est présentée dans le chapitre XVII, se développent sur une ambiguïté conceptuelle qui les empêche de présenter une explication des revenus perçus par les banques en raison de leur activité monétaire.

Si l'on considère le système bancaire comme un tout ⁽¹⁾ et que l'on retient comme définition *a priori* de l'intérêt qu'il est un surplus de monnaie, alors le modèle traditionnel du circuit est confronté à une difficulté : la logique interdit que les banques émettent elles-mêmes l'intérêt monétaire, en sus de la monnaie nécessaire à la monétisation du produit. C'est dire que l'intérêt doit prendre forme dans le processus économique, sinon on admettrait cette explication aberrante selon laquelle les banques auraient le pouvoir de créer leur propre revenu en émettant de la monnaie sur elles-mêmes.

L'intérêt monétaire étant supposé rémunérer l'activité monétaire des banques, son analyse doit s'inscrire dans la perspective de la définition d'un revenu macroéconomique. La question n'est pas alors seulement celle de la création de monnaie, mais aussi celle de la formation d'un revenu, donc d'un pouvoir d'achat constitutif des profits bancaires. Par ailleurs, le caractère macroéconomique de l'intérêt monétaire impose que celui-ci ne soit pas alimenté par une épargne liquide préalablement constituée, mais qu'il trouve sa source dans le produit courant. Il y va de la cohérence de l'analyse fondée sur le principe de la demande effective. Il s'ensuit que les sommes qui alimentent l'intérêt monétaire net devraient être dégagées de la monnaie émise dans la période courante. L'analyse est donc soumise à une contrainte absolue : *le système bancaire ne peut retrouver plus de monnaie qu'il n'en lance dans l'économie.*

De cette contrainte, J. Léonard tire que : «Dans ces conditions surgit une première conclusion fondamentale : dans le cadre des hypothèses du système envisagé *la masse globale de l'intérêt monétaire est nécessairement et logiquement nulle...*» (J. Léonard, 1987, p. 50.) L'existence même d'un surplus de monnaie ne peut être prouvée. Tout semble indiquer au contraire que l'intérêt monétaire ne peut être alimenté par une création additionnelle de liquidités par le système bancaire. De cela J. Léonard déduit alors que l'activité des banques s'inscrit dans une situation paradoxale génératrice de crise, où sont stipulés «des *taux* d'intérêt positifs au sein d'une configuration

(1) Nous considérons bien entendu que la Banque centrale fait partie intégrante du système bancaire.

contrainte par la nullité de la *masse* de l'intérêt net» (J. Léonard, 1987, p. 150). Or, il nous semble que cette contradiction n'est pas le fait du système économique, mais bien plutôt le résultat d'un glissement analytique qui n'intègre pas la véritable nature de la monnaie.

Si l'on entend par intérêt monétaire net un surplus de monnaie, il est certain que sa «masse» est nulle, pourtant cela ne signifie pas que l'intérêt net qui résulte de l'activité monétaire des banques soit lui-même nul. Partant de l'idée que le *taux* d'intérêt est une *variable* déterminée sur le marché monétaire, il est abusif d'en déduire implicitement que *l'intérêt* doit être lui-même une *grandeur* purement monétaire. Autrement dit, le caractère monétaire du *taux* d'intérêt n'implique pas nécessairement que la monnaie bancaire soit porteuse d'intérêt et que l'intérêt soit constitué de pure monnaie. En vérité, la question du surplus de monnaie s'appuie sur l'idée qu'il existe dans l'économie une «pure monnaie» créée par les banques et mise à la disposition des entreprises⁽²⁾. Le cœur du problème est bien la nature de la monnaie bancaire et, par voie de conséquence, le mode de financement de la demande effective et le rôle du système bancaire dans le processus de financement de la production et de la circulation des marchandises.

La question de l'intérêt monétaire relève, en vérité, de celle plus vaste du financement de la production et notamment de l'investissement. On sait les critiques auxquelles a dû faire face Keynes à la suite de la publication de la *Théorie générale*⁽³⁾ où il développe l'idée selon laquelle les entreprises peuvent se procurer des moyens de production sans puiser dans une épargne préalablement formée. Dans deux articles de juin et décembre 1937⁽⁴⁾, publiés dans *The Economic Journal*, il répond à ces critiques grâce à la notion de «motif de finance», comme quatrième motif de détention de monnaie. Or, l'introduction de «la finance» dans le champ de l'analyse du financement de la production nous permet de comprendre que ce qu'il est convenu d'appeler «l'intérêt monétaire», bien que macroéconomiquement net, ne suscite la création d'aucune monnaie supplémentaire par rapport à ce qui est nécessaire au financement de la production courante.

(2) On peut voir, à ce sujet, A. Parguez, «Ordre social, monnaie et régulation», *Economie Appliquée*, tome XXXIV, 1981, n°2-3, Librairie Droz, Genève.

(3) Nous pouvons notamment citer Hicks J.R., «Mr Keynes theory of employment», *The Economic Journal*, juin 1936, p. 238-253; Ohlin B., «Some notes on the Stockholm theory of savings and investments II», *The Economic Journal*, juin 1937, p. 221-240; Robertson D.H., «Alternative theories of the rate of interest», *The Economic Journal*, septembre 1937, p. 428-436.

(4) «Alternative theories of the rate of interest» et «The "ex ante" theory of the rate of interest». Cf. Keynes, 1937a et 1937b.

I. MONNAIE ET INTÉRÊT MONÉTAIRE

C'est un fait établi de longue date : tout intérêt résulte du prêt d'un actif. Aussi, si l'on imagine que les banques doivent percevoir un intérêt en raison de la monnaie qu'elles ont créée, c'est que l'on admet que la monnaie bancaire est de la nature d'un actif. Autrement dit, les banques auraient par avance le pouvoir de créer une richesse positive sous forme liquide. Construit sur cette base, le modèle du circuit se présente comme celui de la circulation d'une pure monnaie que les agents économiques pourraient se transmettre sur les différents marchés, et ce, pendant le temps nécessaire à son retour à son point de départ.

De cette conception circulatoire de la monnaie, on tire alors que l'intérêt, portant sur cette pure monnaie, court entre le moment où celle-ci est avancée aux entreprises et celui où ces dernières la restituent aux banques, au terme de la période de prêt. Ainsi, les entreprises sont réputées endettées monétairement et, partant, redevables d'un intérêt monétaire aux banques. Le versement d'un intérêt au système bancaire semble justifié par la formation d'une dette de l'économie vis-à-vis des banques d'une part, et en raison du temps qui s'écoule entre la date du prêt et celle du remboursement d'autre part. Or, M. Zerbato affirme avec juste raison que « c'est la réduction de la dimension monétaire du circuit à celle d'une pure monnaie endogène qui interdit d'expliquer le profit monétaire et, *a fortiori*, l'intérêt » (M. Zerbato, 1990, p. 98). Toutefois la solution exogène de M. Zerbato n'est pas non plus recevable. Elle s'inscrit en contradiction avec le principe de la demande effective qui fonde la théorie keynésienne du circuit, puisqu'elle admet qu'il puisse se former un pouvoir d'achat en dehors de la production. Le véritable problème qui se pose ici n'est pas celui de l'exogénéité ou de l'endogénéité de l'intérêt, mais bien plutôt celui de l'existence d'une pure monnaie qui ne soit pas directement attachée au produit et qui néanmoins serait dotée d'un pouvoir d'achat.

En fait, le schéma proposé repose sur la confusion analytique des deux fonctions d'intermédiation des banques : intermédiation monétaire et intermédiation financière⁽⁵⁾. Autrement dit on confond création monétaire et octroi de crédits ou, en d'autres termes encore, création de monnaie et prêt d'actifs liquides. Or, grâce à la théorie des

(5) Dans la deuxième partie du chapitre 31 de *A Treatise on money*, Keynes avait déjà fait ressortir les deux grandes fonctions bancaires. On peut voir aussi à ce propos J. Riboud, « Controverse sur la banque et la monnaie », *Revue politique et parlementaire*, diffusion PUF, Paris, 1986, p. 17 et suiv.

émissions développée par B. Schmitt, nous pouvons montrer que les deux opérations ne sont pas confondues et que les banques ne peuvent recevoir d'intérêt sur la seule base de la monnaie qu'elles créent.

1. Le circuit de la monnaie est instantané

Par nature la monnaie bancaire est de crédit puisqu'elle n'a d'autre origine que l'engagement des banques à effectuer des paiements pour le compte de leurs clients qui, de leur côté, s'engagent à restituer la monnaie dont ils ont pu disposer. Il nous faut clairement distinguer l'opération, le crédit, de l'objet de l'opération, la monnaie.

Dès *A Treatise on money*, Keynes a découvert que, bancaire, la monnaie n'a pas d'existence hors de son association au produit physique. Il montre en effet que chaque banque « crée une demande d'émission de monnaie vis-à-vis d'elle-même. » (Keynes, 1971a, p. 20). Mais il est certain que les banques ne peuvent émettre la monnaie sur elles-mêmes pour ensuite la prêter. De fait, par simple jeu d'écriture, les banques lancent une promesse de payer, mais tant que le paiement n'a pas lieu et donc que la promesse n'est pas réalisée, la monnaie n'existe pas. Autrement dit, tant que les banques n'exécutent pas un ordre de paiement, pour le compte d'un agent non financier sur un autre, la monnaie n'est pas émise. On nous accordera que l'ouverture d'une ligne de crédit, par exemple, ne vaut pas endettement des particuliers puisque « l'avance en compte n'est comptabilisée par la banque qu'au fur et à mesure de son utilisation par le bénéficiaire pour effectuer ses paiements. C'est à ce moment seulement qu'a lieu la création monétaire. » (F. Poulon, 1982, p. 185.) Puisque la monnaie naît dans l'opération même de la rémunération des facteurs de production, c'est-à-dire dans un paiement, il serait contradictoire d'affirmer ensuite qu'il puisse exister une « pure monnaie » formée avant ou en dehors de ce même paiement.

Il apparaît clairement que la monnaie est créée grâce à l'intermédiation des banques. C'est dire que l'émission de la monnaie requiert l'association de trois agents, ou trois groupes d'agents : les banques, les entreprises et les travailleurs. Le processus de l'émission est triangulaire. Intermédiaires, les banques ne peuvent se rendre monétairement créditrices de l'économie sans en être simultanément débitrices.

Lorsqu'elles concrétisent leur promesse de payer en créditant les travailleurs, les banques lancent une monnaie qui est simultanément

positive et négative. Le paiement des travailleurs se traduit en effet par une double écriture comptable. La monnaie apparaît comme grandeur négative à l'actif des banques, c'est la dette des entreprises, et positivement à leur passif, comme créance des travailleurs. «Par définition la monnaie est un actif et identiquement un passif, autrement dit un *actif-passif*» (B. Schmitt, 1975, p. 13).

La monnaie inscrite au compte des travailleurs n'est pas pour eux-mêmes à la fois positive et négative. Au contraire dès que leur compte est crédité, ils se retrouvent détenteurs d'une créance nette, dont l'objet est le produit physique auquel ils ont donné naissance par leur travail. En d'autres termes, les titulaires de revenu détiennent le produit sous forme monétaire.

De leur côté les entreprises ne reçoivent positivement aucune monnaie de la part des banques. Une telle occurrence est logiquement impossible, puisque, au moment du versement des salaires, il n'existe encore aucune monnaie positive. La monnaie «prêtée» aux entreprises ne peut donc apparaître qu'à l'actif des banques en tant que dette, c'est-à-dire négativement puisque la dette est formée à partir de zéro. Il est toutefois bon de noter que cette dette des entreprises est purement nominale, puisqu'elle correspond exactement à la valeur du produit qu'elles détiennent physiquement.

A l'issue de l'opération, la monnaie n'étant pas positive et négative sur le même agent, la créance et la dette ne s'annulent donc pas pour l'ensemble de l'économie. Le résultat de l'émission est positif puisque le revenu est maintenant formé, monétairement sur les titulaires de revenus et physiquement dans les entreprises.

Mais, si l'opération est bien de la nature d'un crédit, il ne peut cependant en naître aucun intérêt monétaire puisqu'il ne s'écoule pas de temps entre le moment où les banques prêtent la monnaie (aux entreprises) et celui où elle leur est restituée (par les titulaires de revenus). Nous pouvons encore exprimer cela autrement en disant que l'émission de la monnaie ne suscite la formation d'aucune créance *monétaire* nette des banques sur l'économie.

En effet, la monnaie lancée par les banques pour le compte des entreprises leur est restituée sans délai par son inscription positive au compte des titulaires de revenu. La monnaie n'a à aucun moment circulé hors des banques. L'inscription positive au compte des travailleurs est bien le signe que la monnaie n'est pas sortie de l'espace bancaire. Le circuit de la monnaie ne s'inscrit pas dans l'écoulement du temps, les banques étant débitées (flux) et créditées (reflux) dans le même mouvement. De cela on infère que la monnaie n'a pas de durée positive

dans le temps et que son émission ne se traduit pas par le prêt d'un actif à l'économie par les banques. Il ne s'est donc formé, en raison de l'émission monétaire, aucune dette de l'économie vis-à-vis des banques. Partant, celles-ci ne détiennent aucune créance *monétaire* sur les agents économiques et de ce fait ne peuvent prétendre percevoir un intérêt. La théorie moderne des émissions nous permet de vérifier une profonde intuition de Thomas d'Aquin qui dès le XIII^e siècle écrivait : « Recevoir un intérêt pour l'usage de l'argent est de soi injuste, car c'est faire payer ce qui n'existe pas... Pour s'en convaincre il faut se rappeler que l'usage de certains objets se confond avec leur consommation. Dans les échanges de cette nature on ne devra donc pas compter l'usage de l'objet à part de sa réalité même... » (cité par A. Lapidus, 1987, p. 101).

De ce qui précède nous pouvons tirer une conclusion d'ordre général : puisqu'il n'y a pas de dette monétaire de l'économie vis-à-vis du système bancaire, la question du versement d'un intérêt monétaire aux banques au titre de l'émission de la monnaie ne se pose pas, faute d'objet. *La monnaie n'est pas porteuse d'intérêt.*

Pourtant il est indéniable que les banques stipulent un taux d'intérêt lorsqu'elles effectuent des opérations de crédits au bénéfice des entreprises qui rémunèrent des services producteurs. C'est qu'en réalité l'émission de la monnaie se double d'une opération financière. Simultanément à leur fonction d'intermédiation monétaire, les banques assurent une intermédiation financière. Il est donc vrai qu'en faisant appel aux banques pour financer la production les entreprises contractent une dette *financière* d'un montant rigoureusement égal à celui de la monnaie émise.

2. La formation de la dette financière des entreprises

Il est certain que les titulaires de revenu ne peuvent dépenser leur revenu dans l'instant même où ils le reçoivent. N'étant pas consommé, nous pouvons affirmer que le revenu est intégralement épargné. Or, cette épargne est immédiatement prêtée aux banques et forme les dépôts ordinaires, jusqu'au moment de son retrait par les ménages. Cependant, dans l'instant même de l'émission monétaire, la dette et la créance initialement *monétaires* des entreprises et des ménages sont détruites.

Nous avons dit ci-dessus que la monnaie est instantanément restituée aux banques (donc détruite) par son inscription au compte des ménages et que cela valait pour le désendettement monétaire des entreprises à l'égard des banques. Mais les entreprises ne peuvent se libérer de

leur dette monétaire inscrite en banque que parce qu'elles sont dans la position de leur restituer la monnaie qui a été lancée à leur demande. Elles ne peuvent le faire qu'en *empruntant* la monnaie détenue par les titulaires de revenus, puisque c'est la seule monnaie qui ait été formée positivement. Or les ménages sont maintenant titulaires d'un actif monétaire et c'est cet actif liquide qui fait l'objet du prêt aux entreprises.

Lorsque nous écrivons que la monnaie est restituée aux banques, c'est bien *la forme* qu'elles ont émise qui leur est rendue, mais pas le pouvoir d'achat qui reste fixé dans l'économie, aux comptes des titulaires de revenus. Lorsqu'ils restituent la monnaie aux banques en lieu et place des entreprises, les titulaires de revenus *prêtent leurs revenus* et de ce fait se retrouvent créanciers des entreprises. A la dette monétaire se substitue, grâce à l'intermédiation maintenant ⁽⁶⁾ financière des banques, une dette financière des entreprises, dont l'objet n'est pas la monnaie, mais le produit courant. Ainsi, à l'issue de l'opération de crédit bancaire, les entreprises sont endettées vis-à-vis des ménages à travers les banques, et non pas, à proprement parler, vis-à-vis des banques qui ne sont que des intermédiaires.

Il est bon de noter ici, pour la clarté de l'exposé, que la créance des ménages porte sur le produit lui-même et non pas sur les entreprises proprement dites. Ces dernières en effet empruntent, non pour acquérir le produit pour leur propre compte (elles le détiennent déjà sous sa forme physique), mais pour financer la production. Nous pouvons dire avec B. Schmitt que le prêt qui double l'émission de la monnaie est un «prêt absolu», car il n'engage aucun déplacement de marchandises et, bien qu'ils prêtent leurs revenus, les ménages ont toujours le pouvoir d'obtenir le produit. Ils détiennent des droits de créances libellés sur les banques et portant sur le stock des produits physiques déposés en entreprises. Ces dernières ne pourront sortir de leur dette qu'en vendant les biens qu'elles ont en stock.

Il n'y a donc rien d'étonnant à ce que, durant la période où le revenu n'est pas consommé par les ménages, on puisse noter, comme M. Lavoie, que les entreprises «ne versent des intérêts aux banques que sur l'équivalent moyen de l'épargne conservée sous forme de dépôt par les ménages» (M. Lavoie, 1987, p. 84). S'ils sont d'abord versés aux banques, ces intérêts ne sont pas destinés à alimenter leurs propres

⁽⁶⁾ Lorsque nous écrivons «maintenant», cela ne signifie pas qu'il y ait décalage dans le temps continu entre les deux intermédiations. La dissociation est purement analytique.

revenus, mais en principe devraient rémunérer les comptes de dépôts des titulaires de revenus ⁽⁷⁾.

A nouveau ici, nous voyons que l'économie n'est pas à proprement parler endettée vis-à-vis des banques. Aussi, quand bien même il prête un revenu épargné par les ménages, le système bancaire ne peut-il prétendre recevoir pour lui-même un intérêt. Cela n'interdit en rien, bien entendu, que les banques soient rémunérées pour les services rendus, mais ces rémunérations ne sont pas de la nature d'un intérêt. Il est vrai que les banques captent pour elles-mêmes une partie de ces intérêts. Ce n'est pas en l'occurrence parce qu'elles prêtent leurs propres liquidités, mais qu'en tant qu'entreprises elles produisent un service d'intermédiation.

De ce qui précède nous pouvons tirer que les banques ne perçoivent pas d'intérêt macroéconomique en raison de leur activité en matière de monétisation initiale de la production. Certes, la monétisation de la production est l'aspect fondamental de l'activité bancaire en matière monétaire. Mais la fonction monétaire du système bancaire ne se limite pas à cela; les banques contribuent aussi au financement de la demande effective.

II. « FINANCE » ET INVESTISSEMENT

Dans ses articles de juin et décembre 1937, Keynes introduit, dans sa théorie de la préférence pour la liquidité, le motif de « finance » comme quatrième motif de détention de monnaie. Mais ce qui distingue le « motif finance » des trois autres motifs de détention de liquidités, c'est qu'il est directement associé aux dépenses des entreprises en matière d'équipements. La demande de liquidités auprès des banques, au titre du « motif de finance » par les entreprises, correspond à leur volonté de se procurer des moyens de production avant même d'avoir capté l'épargne nécessaire à leur financement. De ce point de vue la « finance » peut être interprétée comme une avance du système bancaire au secteur productif.

Keynes définit la « finance » comme une avance de liquidités faite par les banques aux entreprises qui souhaitent investir avant même d'avoir capté les fonds nécessaires pour cela sur les différents marchés. Il écrit notamment : « Pendant la période d'interregne – et pendant cette période

⁽⁷⁾ C'est bien là la source de la rémunération des comptes courants pratiquée dans certains pays.

uniquement – entre la date où l'entrepreneur aménage sa finance et la date où il réalise son investissement, il y a une demande additionnelle de liquidités sans que, pour autant, il y ait nécessairement un accroissement de l'offre de celle-ci» (J. M. Keynes, 1937b, p. 218). Mais, bien qu'il indique : «pour l'ensemble de la communauté, elle est seulement une transaction comptable» (J. M. Keynes, 1937b, p. 219), il précise que la «finance» «requis pendant la période d'interrègne entre l'intention d'investir et son accomplissement est principalement offerte par des spécialistes, en particulier par les banques, qui organisent et gèrent un fonds rotatif de finance liquide» (J. M. Keynes, 1937b, p. 219). Bien que n'existant à proprement parler que dans la comptabilité des établissements financiers, il n'en reste pas moins que la «finance» s'inscrit dans le processus d'écoulement du produit comme un stock de liquidités mises à la disposition des entreprises. Mais il est bon de noter avec A. Graziani que la «finance» intervient dans le financement *initial* de la production et ne peut de ce fait être confondue avec le financement final des biens capitaux à l'aide de l'épargne des titulaires de revenus⁽⁸⁾. Ainsi, le rejet de la théorie des fonds prêtables se trouve confirmé par le concept de «finance» qui permet de résoudre le problème de chronologie entre investissement et épargne soulevé par les critiques de Keynes. A. Graziani écrit : «Le problème de l'ainsi dit financement des investissements ne pourra se présenter qu'après [le financement initial], quand, le revenu ayant été créé, il est possible d'avoir une formation d'épargne correspondante, qui pourra être absorbée par le placement de titres» (A. Graziani, 1985, p. 169).

Dans un précédent article nous avons montré (J.-L. Bailly, 1992) que la finance avancée par les banques joue le rôle d'une épargne, sans pour autant s'y substituer définitivement. Cela signifie qu'il se forme au niveau bancaire un flux de liquidités qui ne sont pas puisées dans l'épargne des agents. C'est sur l'existence de ce flux que s'appuie Keynes pour contester la théorie des fonds prêtables. Il écrit notamment : «Afin de nous aider à lever tout malentendu, nous pouvons remarquer que tandis que l'épargne se forme simultanément à l'investissement (en raison de la première acquisition d'un bien capital par un entrepreneur), le mouvement de fonds (*i.e.* de monnaie) disponible pour l'investissement (en raison de la première acquisition de ce bien capital par un détenteur définitif) a lieu ensuite; la couverture de ce délai par la "finance" (*i.e.* par l'offre de monnaie) étant le rôle

(8) Cf. «Le financement de l'économie dans la pensée de J.M. Keynes», *Cahiers d'économie politique*, «La théorie générale de John Maynard Keynes : un cinquantenaire», éd. L'Harmattan, Paris, 1988, p. 150-166.

du système de crédit (qui s'occupe uniquement de finance et jamais d'épargne)» (Keynes, 1939, p. 284). C'est dire qu'une partie de la monnaie émise par les banques peut se constituer en capital qui, bien que non directement attaché au produit physique, s'insère dans le processus général du financement de la demande effective et donc de la mise en œuvre de la production. Propriété des banques, ce capital liquide justifie le versement d'un intérêt à leur endroit.

1. La formation de la « finance »

Le concept de « finance » s'inscrit dans la perspective keynésienne selon laquelle l'investissement autonome des entreprises n'est pas *initialement* financé par une épargne préalablement constituée. Toutefois, la « finance » doit être saisie dans le respect du principe général de l'équivalence des revenus des services producteurs et du produit. Autrement dit, la « finance » n'est pas une addition de pouvoir d'achat qui serait créé par les banques, bien qu'elle permette aux entreprises de se procurer les biens de production dont elles ont besoin, sans attendre l'écoulement final de leurs produits. La question de la formation de la « finance » revient donc à s'interroger sur la formation d'un capital liquide non additif au pouvoir d'achat⁽⁹⁾ disponible dans l'économie.

Considérons, dans une période donnée, une économie dotée d'un système bancaire et composée de deux entreprises quelconques E_1 et E_2 , et de leurs salariés. Peu importe la forme des produits fournis par ces entreprises, puisque la notion de « finance » ... recouvre tout autant l'utilisation de la masse de fonds tournants pour financer la production de biens capitaux que la production de biens de consommation...» (Keynes, 1939, p. 283).

Supposons que, à l'aide de crédits bancaires en « finance » d'un montant de 20, l'entreprise E_1 se procure des biens de production auprès de l'entreprise E_2 . E_1 ne peut faire cette acquisition que parce que les produits de E_2 n'ont pas encore été vendus. C'est dire que l'équivalent des revenus formés lors de la production des biens encore stockés par E_2 est épargné par les ménages⁽¹⁰⁾ et que ni E_1 ni E_2 n'ont encore pu capter cette épargne. A première vue on peut penser qu'il

⁽⁹⁾ C'est-à-dire qui ne s'ajoute pas au revenu courant et donc à l'épargne courante.

⁽¹⁰⁾ L'identité des détenteurs de ces revenus ne nous importe pas ici, qu'ils soient originaires de E_1 ou E_2 est indifférent. Par la suite, et afin de clarifier l'analyse, nous supposerons que ce sont les salariés de E_2 qui épargnent les 20 considérés.

existe dans l'économie 40 unités de monnaie correspondant à un produit de 20. Or il n'en est rien.

Grâce aux 20 qui lui sont versés par E_1 , E_2 va pouvoir se dégager de sa dette formée lors du versement des salaires et encore inscrite en banque. Ainsi, la monnaie offerte au titre de la « finance » est, comme toute monnaie bancaire, créée et détruite dans l'instant. Pourtant, la monnaie émise par la banque suscite une dette financière de E_1 , cette dernière n'a pas encore payé les biens dont elle dispose maintenant. E_1 paiera plus tard, après avoir capté les revenus des ménages. Notons qu'aucune monnaie supplémentaire ne vient « concurrencer » le pouvoir d'achat qui est à la disposition des ménages et a été formé dans la production des équipements achetés par E_1 . En face des produits détenus maintenant par E_1 , seuls subsistent dans l'économie les 20 détenus par les titulaires de revenus. L'échange interindustriel n'a provoqué la formation d'aucune unité additionnelle de pouvoir d'achat. De l'autre côté, aucune monnaie ne fait défaut dans l'économie pour écouler le produit global de E_1 et E_2 . Mais il nous faut analyser plus avant le phénomène.

Si E_2 vend pour 20, c'est que ces 20 ont été produits et ont occasionné la formation de revenus. De cela on tire que les salariés de E_2 détiennent 20 sous forme d'épargne. Par hypothèse ils n'ont pas encore dépensé cette dernière dans l'achat des produits de E_1 , sinon E_1 n'aurait pas emprunté 20 à la banque pour acheter des moyens de production à E_2 . Ainsi, alors que les biens sont achetés par E_1 , les ménages rémunérés par E_2 restent détenteurs d'un pouvoir d'achat équivalent au produit de E_2 . « (...) En son premier état, le bien capital est constitué de *biens salaires* stockés face à l'épargne monétaire. » (B. Schmitt, 1984, p. 165). Les biens achetés par E_1 sont stockés productivement et les revenus versés par E_2 sont épargnés.

E_1 étant maintenant propriétaire des biens produits par E_2 , on en infère que E_1 a permis à E_2 de se désengager aussi bien vis-à-vis des banques que des ménages. Mais, pour autant, l'ensemble des entreprises $\{E_1, E_2\}$ n'a pas éteint sa dette financière formée vis-à-vis des ménages⁽¹⁾, et a de plus contracté une nouvelle dette financière vis-à-vis du système bancaire. Les ménages restent détenteurs d'un titre sur l'équivalent du produit de E_2 et les banques sont créancières de E_1 . Il ressort de cela que :

– au regard des titulaires de revenus, E_1 s'est substituée à E_2 . Autrement dit, E_1 reprend à son compte l'opération de versement des

⁽¹⁾ Cf. ci-dessus, 2. *La formation de la dette financière des entreprises.*

salaires aux travailleurs de E_2 , donc les engagements réels et financiers de E_2 , en plus des siens propres;

- E_1 , endettée pour 20 vis-à-vis de la banque, devra rembourser celle-ci, avec intérêt, après avoir écoulé finalement ses propres produits;
- la reprise d'émission par E_1 , grâce à la « finance », fait que le système bancaire devient un partenaire dans le financement de la production.

L'analyse ne concerne donc plus E_2 , mais seulement les banques, E_1 et les titulaires de revenus. Que les revenus aient été originellement formés sur E_1 ou sur E_2 est maintenant tout à fait indifférent pour la question qui nous intéresse ici. Les transactions effectuées dans l'espace des entreprises ne remettent pas en cause le processus d'intégration de la monnaie et l'équivalence des salaires et du produit. Répétons-le, dans l'opération aucun pouvoir d'achat additionnel n'a été créé par les banques et aucun paiement final n'a été effectué. Certes E_2 ne détient aucune créance sur E_1 , mais E_1 n'a pas encore payé les biens dont elle dispose puisqu'elle a pu se les procurer en contractant une dette.

Etant donné que E_1 emprunte en banque pour obtenir les 20 produits par E_2 , le phénomène n'est pas exactement de même nature et l'opération n'a pas les mêmes conséquences que si E_1 avait acheté les 20 après avoir capté les revenus versés par E_2 en vendant, par exemple, des titres sur le marché financier. Du fait de l'intermédiation bancaire, les échanges interindustriels ne sont pas neutres; en reprenant l'émission des salaires de E_2 , E_1 intègre au produit la monnaie lancée au titre de la « finance » et donne un contenu réel à la créance bancaire, ou si l'on préfère à la « finance ». Ainsi, la « finance » existe bien en tant que capital financier détenu par les banques et est porteuse d'intérêts.

Nous pouvons maintenant saisir toute l'importance de la notion de « finance » dans la description du processus général de financement de la production, et la rupture radicale qu'elle marque vis-à-vis de la théorie des fonds prêtables. Elle conforte l'idée selon laquelle la production peut être engagée et élargie sans recourir à une épargne préalablement constituée. Elle permet, comme l'avait pressenti Keynes⁽¹²⁾, d'expliquer la dynamique de l'accumulation du capital, dans le respect rigoureux du principe fondamental selon lequel le financement courant de l'économie est réalisé *in fine* grâce à la production courante. Cela signifie encore que la « finance » ne participe pas de la *monétisation* initiale du produit. La « finance » existe parallèlement à l'épargne des ménages. Mais, pour autant, elle ne s'y substitue pas dans l'espace des entreprises et, pour reprendre une expression de Keynes, nous pouvons dire qu'elle en

(12) Cf. Keynes, [1939].

est le parent et non le jumeau. Les 20 de notre exemple ne sont pas stockés sous forme liquide par le système bancaire et ne constituent pas un capital *monétaire* macroéconomique qui pourrait être à nouveau prêté aux entreprises pour réaliser des investissements avancés. Les 20 prêtés par le système bancaire sont dépensés et existent maintenant sous la forme d'un capital *financier*. D'ailleurs, ils n'apparaissent comme «finance» que parce qu'ils sont dépensés. Ils ne sont donc pas de la catégorie des «fonds prêtables» puisqu'ils ne s'ajoutent pas à l'épargne.

Dans notre exemple, après formation de la «finance», la situation se présente alors de la manière suivante.

Les salariés de E_2 ont un titre (leur revenu monétaire de 20) sur le produit de l'ensemble $\{E_1, E_2\}$. Mais la banque détient une créance d'un montant équivalent sur E_1 . Ainsi, en achetant le produit de E_2 à l'aide d'une avance bancaire, E_1 se trouve devant deux engagements, l'un, vis-à-vis des titulaires de revenus, l'autre, vis-à-vis de la banque. Or, E_1 ne peut se dégager du second sans se dégager du premier. Elle ne pourra rembourser la banque qu'en vendant son propre produit avec profit.

Nous l'avons dit, l'achat des moyens de production équivaut à une reprise d'opération de la part de E_1 . C'est dire que tout se passe comme si E_1 avait elle-même rémunéré les salariés de E_2 . Mais ici le phénomène comporte une particularité : *pour investir, E_1 a dépensé par avance son profit sur le marché des produits*. Donc en reprenant à son compte le coût des salaires versés aux travailleurs employés en E_2 , E_1 transforme le paiement des travailleurs en un achat direct du produit. C'est dire que d'emblée les titulaires de revenu ne pourront exercer leur pouvoir d'achat pour eux-mêmes sur l'équivalent du produit de E_2 . Leur revenu est immédiatement constitué en épargne macroéconomique, et c'est cette épargne que devra capter E_1 pour rembourser la banque.

2. Le remboursement de la «finance»

Si l'on considère la période qui sépare le prêt de la «finance» de son remboursement, on peut avoir l'impression qu'il existe dans l'économie un stock autonome de liquidités, ou si l'on préfère un capital liquide sans contrepartie, créé par les banques. On comprend alors pourquoi de nombreux auteurs, faute de percevoir le processus de formation de la «finance», pensent que les banques ont le pouvoir de créer un capital *ex nihilo*. Mais, en réalité, la «finance», «qui peut saisir un investissement après l'autre» (J. M. Keynes, 1937a, p. 209), reste enfermée dans

l'espace des institutions financières et des banques ⁽¹³⁾, et ne circule pas entre les autres agents économiques.

Pendant la période «d'interrègne» il est vrai que la «finance» fonctionne comme un capital macroéconomique liquide dans le sens où sa formation ne réduit pas la richesse disponible dans la société, «Elle n'emploie pas d'épargne» (J. M. Keynes, 1937b, p. 219). Ce qui de plus distingue la «finance» d'une épargne ordinaire, c'est qu'elle est *nécessairement* investie et que, cela étant, elle permet aux entreprises de déterminer par anticipation la fraction du produit courant qui est consacrée à l'investissement. Toutefois, il est bon de garder à l'esprit que la création de monnaie au titre de la «finance» est *postérieure* à la formation du revenu courant et donc, à l'épargne courante, et que de ce fait elle ne peut être la cause de l'épargne.

Dans notre exemple, E_1 , qui achète des biens d'équipement à E_2 , reprend à son compte l'engagement contracté par E_2 de fournir un produit, sous forme de biens ou de titres, aux titulaires de revenus. Elle ne peut le faire que grâce au crédit bancaire. Les banques, d'autre part, ont prêté 20 à E_1 , parce que celle-ci n'avait encore pu capter l'équivalent des revenus formés sur E_2 . Ainsi, si E_1 a produit pour 80, elle se retrouve engagée à fournir aux titulaires de revenus la totalité de son produit plus l'équivalent de celui de E_2 , soit 100 au total. Dans le même temps E_1 s'est aussi engagée à restituer 20 aux banques (sans compter l'intérêt), or ces 20 existent dans l'économie.

De l'autre côté, les titulaires de revenus disposent de 100 unités de revenu, dont 20 sont *nécessairement* épargnées. Or, ces 20, qui correspondent aux biens d'équipement consommés productivement par E_1 , sont épargnées de façon *définitive*. Autrement dit, et bien qu'elles soient encore inscrites aux comptes des ménages, ces 20 sont *déjà dépensées productivement* par E_1 pour l'achat de biens d'investissement qui sont sortis de l'espace économique pour entrer dans celui des valeurs d'usages. Les 20 épargnées par les ménages ne pourront donc être utilisées finalement qu'à l'achat des produits offerts par E_1 .

Ainsi, E_1 doit 20 à la banque en raison de la «finance» et 20 aux titulaires de revenus qui détiennent un titre sur la production globale. Mais ces deux dettes ne sont pas additives, puisque les 20 détenues par les titulaires de revenus sont rigoureusement la contrepartie des 20 de «finance». Il s'ensuit que E_1 ne pourra se désengager de sa dette bancaire qu'autant qu'elle pourra capter les 20 détenues par les

(13) Les banques et institutions de crédit peuvent bien entendu s'échanger les titres de créance qu'elles détiennent sur les entreprises.

ménages, sur le marché des produits. En d'autres termes, E_1 ne peut se désendetter vis-à-vis de la banque, et donc effectivement payer les biens d'équipement, qu'à la condition expresse qu'elle capte pour elle-même le pouvoir d'achat correspondant à ces biens.

Notons que, si E_1 vend des titres aux épargnants pour rembourser la banque, elle entretient de fait sa dette en repoussant le remboursement effectif. La vente de titres ne vaut pas achat final du produit par l'entreprise, mais emprunt de fonds. Certes, en émettant des titres, l'entreprise peut annuler sa dette bancaire, mais elle la remplace par une autre, formée vis-à-vis des ménages. En dernier ressort, et quel que soit le créancier de l'entreprise, le paiement final de l'investissement sera réalisé à l'aide de ses profits. C'est la raison pour laquelle, afin de ne pas compliquer inutilement l'analyse, nous retenons l'idée que l'entreprise se désengage de sa dette correspondant à la «finance», en remboursant la banque dans la période courante à l'aide de son profit courant. Le processus d'investissement achevé, le revenu courant apparaît comme la source unique de financement de la production courante.

Nous pouvons voir alors que les 20 qui, pendant la période de transition, constituent un capital inscrit en banques, ne s'ajoutent ni ne s'alimentent au produit courant. La «finance» «n'absorbe ni n'accroît les richesses» (J. M. Keynes, 1937a, p. 209). La «rotation» de la «finance» prépare les conditions pour que se forme une épargne macroéconomique, mais elle ne s'y substitue pas à proprement parler. La fin de la période «d'interrègne» est marquée par le fait que l'épargne chasse la finance de l'espace économique. Tout le produit est finalement écoulé sur les titulaires de revenus.

Toutefois, grâce à leur «endettement spontané», les banques permettent aux entreprises de disposer de moyens de production en temps voulu. C'est-à-dire encore qu'en tant qu'intermédiaires financiers les banques s'endettent vis-à-vis d'elles-mêmes, grâce à leur fonction d'intermédiation monétaire. Mais, à travers la «finance», cette dette est prise en charge par l'économie qui a pu produire de nouveaux biens pendant la période d'interrègne. Il paraît donc raisonnable d'affirmer que, pour cette raison, les banques doivent percevoir un intérêt qui, la dette étant nette pour l'ensemble, doit être lui-même macroéconomiquement net.

III. L'INTÉRÊT DIT MONÉTAIRE EST ORIGINAIRE DE LA PRODUCTION

La «finance» fonctionne comme un fonds liquide net pour l'ensemble de l'économie. A ce titre, elle apparaît comme une contribution financière des banques à la production et à la circulation des marchandises. Aussi, l'intérêt versé par l'économie au système bancaire doit-il être interprété comme le «prix» à payer par la communauté pour que les entreprises disposent d'un financement transitoire de substitution. C'est donc en vertu de ce rôle spécifique de la «finance» que doit être réinterprétée la question de l'intérêt macroéconomique perçu par le système bancaire.

Nous l'avons vu, l'intérêt ne peut être constitué de pure monnaie. La question qui se pose à nous maintenant est de définir l'origine et donc la nature de l'intérêt formé au titre du prêt de liquidité par les banques au secteur productif. Pour cela il nous faut reconsidérer la place de la «finance» dans le processus économique, et d'abord l'objet du prêt des liquidités.

1. «Finance» et production

Les banques prêtent la monnaie aux entreprises pour que celles-ci puissent se procurer des moyens de production. Cette monnaie n'est pas *a priori* porteuse de pouvoir d'achat, ce n'est qu'autant qu'elle se substitue provisoirement à l'épargne des ménages dans l'espace des entreprises qu'elle est dotée d'un tel pouvoir⁽¹⁴⁾. Nécessairement dépensée, la «finance» inscrite comme dette des entreprises auprès du système bancaire est *effectivement* représentative des équipements investis, mais sans priver l'épargne «oisive» des ménages de son pouvoir d'achat. Néanmoins, en tant que forme financière des équipements pendant la période d'interrègne, la «finance» est insérée dans le processus productif général, et à travers elle les banques apparaissent comme des acteurs de la production. De cette manière elles contribuent à la formation de richesses au même titre que n'importe quel autre détenteur de capital financier. Ainsi, nous pouvons dire que, dès l'instant où la «finance» est prêtée aux entreprises, elle entre dans la catégorie du capital productif.

(14) C'est le sens de la reprise d'émission qui fixe la «finance» au sein des entreprises.

L'investissement de la «finance» permet un élargissement de la production. Grâce à sa dépense sur le marché des moyens de production, les travailleurs donnent naissance à une quantité de valeurs d'usages supérieure à ce qu'elle aurait été si les entreprises n'avaient pas investi. Ainsi, une partie du produit formé pendant la période d'interrègne peut bien être imputée à la participation du système bancaire à la production courante.

Par la «finance», les banques disposent d'un titre de propriété sur les biens d'équipement. Elles ne détiennent ce titre de propriété *qu'en raison de ce que la monnaie prêtée a été détruite* lors de l'achat des moyens de production. Il n'existe donc comme liquidité dans l'économie que l'épargne des titulaires de revenus qui correspond aux équipements nouvellement intégrés dans le processus productif. Ainsi, pour honorer sa dette bancaire, l'entreprise doit réaliser un profit en vendant son produit courant. Lors du remboursement, le revenu monétaire restitué au système bancaire est immédiatement détaché des biens auxquels il correspond, il est donc détruit et la monnaie est définitivement annulée. Du point de vue de l'économie considérée dans son ensemble, les 20 restituées par E_1 sont détruites et le produit, maintenant investi, est repris par l'ensemble des entreprises grâce à la dépense du revenu courant.

Nous pourrions dire que la boucle est bouclée, puisque les banques ne créent aucun pouvoir d'achat. Elles ont prêté de la monnaie qui a été «chargée» dans l'espace des entreprises, et elles retrouvent de la monnaie qui a été «vidée» dans l'espace des entreprises. A aucun moment la «finance» n'est intégrée au flux du revenu, elle n'est jamais mise à disposition des titulaires de revenus puisqu'elle est immédiatement dépensée dans l'achat de biens de production. De l'autre côté, lorsque les entreprises remboursent les banques à l'aide de leurs profits, elles en détruisent du même coup le pouvoir d'achat. C'est à ce moment qu'a lieu le paiement final des équipements. Il ne s'agit donc pas d'un transfert de revenu, mais bien d'une dépense définitive. Le remboursement de la «finance» n'apporte donc aucun pouvoir d'achat aux banques, celui-ci est définitivement détruit, c'est dire que l'intérêt versé aux banques prêteuses de «finance» ne peut être additif à celle-ci.

La «finance» joue de façon transitoire le rôle de l'épargne que n'ont pu encore capter les entreprises pour financer leurs investissements. Faisant temporairement fonction d'épargne en matière de financement, elle en revêt également les attributs en matière de rémunération du prêteur. La reprise d'émission réalisée par E_1 , grâce au prêt de la «finance», a pour effet de générer un flux d'intérêt vers les banques.

De même que tout placement d'épargne rapporte un intérêt à son propriétaire, le prêt de la « finance » rapporte un intérêt au système bancaire.

Dans la période d'interregne la monnaie prêtée au titre de la « finance » par le système bancaire a indéniablement accéléré la mise en œuvre d'une nouvelle production. On infère de cela que le secteur productif doit verser un revenu aux banques de la même façon qu'il verse un revenu aux autres catégories de prêteurs. Devant verser des intérêts aux banques en raison du prêt dont elles ont bénéficié, les entreprises intègrent ceux-ci dans leurs coûts de production. Il s'ensuit que l'intérêt formé au titre du prêt de la « finance » est rigoureusement de même nature et donc de même origine que celui formé pour rémunérer les placements en épargne ordinaire dans le système productif. C'est dire que l'intérêt versé au système bancaire, en raison du prêt de la « finance », est de la catégorie des revenus. Il doit donc logiquement trouver sa source dans la production courante elle-même. Pour reprendre une expression utilisée par F. Perroux à propos du profit, nous pouvons dire de l'intérêt portant sur la « finance » qu'il est « originaire » de la production. Et, c'est précisément de cela qu'il tient son caractère macroéconomique.

2. L'intérêt dit monétaire est logé dans les coûts de production

Dire que l'intérêt est « originaire » signifie qu'une fraction du produit courant est soustraite du pouvoir d'acquisition déposé sur les ménages. C'est donc dans l'analyse du circuit du revenu que nous trouverons la réponse relative à l'origine de la monnaie qui alimente l'intérêt dit monétaire.

Les théoriciens du circuit en conviennent, dans une économie monétaire de production, toute production est immédiatement monétaire et réelle. La monnaie étant intégrée lors du versement des salaires, il s'ensuit que la rémunération des facteurs de production fonde l'équivalence des salaires et du produit. Si l'on respecte cette contrainte pour expliquer l'intérêt versé aux banques au titre de leur fonction monétaire, force est d'admettre qu'il est logé dans les salaires nominaux. « Etant donné que les intérêts sont positifs, l'émission de x unités de salaires comprend l'achat, égal à y , de biens intérêts, pour le compte de leurs titulaires... » (B. Schmitt, 1984, p. 560), ici les banques. C'est donc que sur le marché des produits est effectuée une modification de la répartition du revenu nominal.

Ce sont les entreprises qui versent les intérêts monétaires aux banques. Or elles ne peuvent le faire qu'en raison de ce qu'elles

réalisent un profit. Partant, l'explication de l'intérêt monétaire ressortit de celle du profit des entreprises.

Le principe de la formation du profit peut être résumé brièvement. Tout pouvoir d'achat se formant initialement sur les salariés, le profit va se constituer par captation de pouvoir d'achat sur le marché des produits. Autrement dit, en dépensant la totalité de leurs salaires, les travailleurs ne pourront obtenir la totalité du produit réel. On en infère qu'une partie du revenu courant ne peut être dépensée par les travailleurs sur le marché des produits. De fait, la réalisation d'un profit positif sur le marché des produits révèle une dépense de revenu sur le marché des services producteurs. En effet, suivant le principe de la demande effective, lors du versement des salaires, les entreprises anticipent un certain niveau de profit. Or, elles ne peuvent espérer réaliser celui-ci qu'à la condition qu'elles *le dépensent par avance sur le marché des services producteurs*. Ainsi, en raison même de cette dépense, une fraction du produit est *dès l'origine* soustraite du pouvoir d'acquisition des salaires. Ce qui semble être un revenu de transfert formé sur le marché des produits se révèle pour être en fait une réponse à la formation d'un *revenu* qui a été dépensé de manière anticipée, par les entreprises, sur le marché des services producteurs. Puisque le revenu a été dépensé, le produit correspondant doit logiquement rester aux mains des entreprises, c'est dire que l'anticipation du profit ne peut être confondue avec une prévision plus ou moins exacte. Certes, le fait que les entreprises dépensent leur profit sur le marché des services producteurs ne préjuge en rien de ce qu'elles vont *finalement* le réaliser sur le marché des produits. Il est vrai que le profit ne sera véritablement *réalisé* qu'à la condition que les titulaires de revenus achètent les biens qui leur sont offerts aux prix proposés et qu'ainsi les entreprises retrouvent les sommes qu'elles ont lancées dans la rémunération des facteurs de production. Or c'est précisément grâce à cette fraction du revenu qu'elles vont pouvoir payer les intérêts qu'elles doivent au système bancaire.

Ainsi, catégorie du profit au sens large, l'intérêt n'est pas véritablement un revenu de transfert réalisé sur le marché des produits, mais bien au plus profond un revenu de production. Révélé sur le marché des produits, il est initialement inscrit dans les rémunérations des services producteurs. L'intérêt dit «monétaire» ne fait pas exception, «... les entrepreneurs financent eux-mêmes leur profit de risque par le crédit, comme ils financent les autres revenus et l'investissement. Il est donc intégré avec son propre coût, le montant de la charge d'intérêt à payer aux banques, dans le prix de l'offre globale répercuté dans le prix

des produits unitaires» (A. Barrère, 1990b, p. 28). Nous voyons bien maintenant que la question relative à l'origine de la monnaie qui vient alimenter l'intérêt dit «monétaire» ne se pose pas. Tout revenu est émis lors de la rémunération des services producteurs et aucune monnaie ne fait défaut dans le circuit pour payer les intérêts bancaires.

L'intérêt versé aux banques apparaissant comme un revenu de production, nous pouvons dès lors constater qu'il n'est pas constitué de «pure monnaie». Prélevé sur les profits bruts des entreprises, il est partie intégrante du circuit du revenu et son existence ne fait pas appel à une branche «spéciale» ou «exogène» du circuit de la monnaie.

C'est en raison même de ce que la «finance» contribue à l'accroissement de la productivité physique du travail que l'intérêt dit «monétaire» inscrit la participation du système bancaire à la formation de la structure productive dans le produit courant. Finalement, il ressort que les revenus des banques entrent dans la catégorie des coûts macroéconomiques de production. A ce titre, l'intérêt versé aux banques est supporté par l'ensemble des entreprises et des ménages. Il n'est donc pas additif au revenu monétaire global et ne réduit en rien le pouvoir d'achat disponible dans l'économie.

CONCLUSION

La monnaie ne trouve pas son objet dans le système bancaire non plus que sur le marché monétaire, elle n'est donc pas en elle-même un actif. Aussi, bien que naissant dans des opérations de crédit, elle n'est pas porteuse d'intérêt. Il est vain dans ces conditions de rechercher un surplus de monnaie qui viendrait alimenter le profit des banques.

La «finance», comprise comme un capital liquide qui reste enfermé dans l'espace des banques, permet de saisir la possibilité d'un intérêt positif, c'est-à-dire net pour l'ensemble des banques, dont le taux est déterminé sur ce qu'il est convenu d'appeler le marché monétaire. De fait, la «finance» joue, pendant la période d'interrègne, le rôle que joue toute épargne prêtée aux entreprises. L'affirmation de Keynes dans ses deux articles de juin et décembre 1937, notamment, est vérifiée, la demande de monnaie pour «motif de finance» va indiscutablement influencer le niveau des taux d'intérêt.

En outre, l'absence de «paradoxe de l'intérêt» n'interdit pas la possibilité de crises. Au contraire le fait de saisir un intérêt monétaire positif permet de mieux comprendre les mécanismes monétaires et

financiers qui peuvent entraîner une perturbation de l'économie. En effet, lorsqu'elles forment la «finance», les entreprises anticipent non seulement la captation de l'épargne des ménages, mais dépensent un profit qu'elles n'ont pas encore réalisé. C'est dire que tout déficit de l'écoulement aura des conséquences encore plus dramatiques du fait que les entreprises devront rembourser les liquidités empruntées, mais aussi les intérêts correspondants. Il est manifeste que les banques ne sont pas neutres et que leur rôle ne se limite pas à la gestion des fonds qui leurs sont confiés par les particuliers et les entreprises. Catégorie du revenu courant, l'intérêt dit monétaire est néanmoins une grandeur nette pour l'ensemble de l'économie. Il ressort du financement initial de la production et répond au rôle actif que jouent les banques, au premier chef de par leur fonction monétaire, dans le processus de formation et d'accumulation des richesses. Leur action a des conséquences déterminantes sur l'activité économique puisqu'elles agissent sur la répartition du produit courant entre consommation et épargne par l'intermédiaire des taux d'intérêt stipulés sur la «finance».

Enfin, la question de l'intérêt dit monétaire nous renvoie, à travers la notion de «finance», à la clarification du principe de la demande effective. Ce dernier est en effet généralement appréhendé du point de vue exclusif des entreprises qui forment leurs décisions en matière de production. Or, le concept de «finance» nous permet de présenter sous un nouvel éclairage le processus de financement de la production, en distinguant nettement la monétisation initiale du produit des opérations financières qui sous-tendent l'acquisition de biens d'investissement par les entreprises. Cela devrait nous permettre d'analyser plus avant la logique du paiement des biens d'équipement et, suivant de nouvelles voies, mesurer l'influence que peut avoir l'activité monétaire des banques dans le développement des phénomènes de l'inflation et du chômage.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Bailly J.-L., «Nouvelles considérations sur le motif de «finance» de John Maynard Keynes», *Economie Appliquée*, tome XLV, n°1, 1992, p. 105-127.
- Barrère A., *Macroéconomie keynésienne*, éd. Dunod, Paris, 1990a.
- Barrère A., «Signification générale du circuit : une interprétation», *Economies et Sociétés*, Monnaie et Production, n°6, février 1990b, p. 9-34.
- Graziani A., «Le débat sur le "motif de financement" de J. M. Keynes», *Economie Appliquée*, tome XXXVIII, n°1, 1985, p. 159-175.

- Graziani A., «Le financement de l'économie dans la pensée de J. M. Keynes», *Cahiers d'Economie Politique, La théorie générale de John Maynard Keynes : un cinquantenaire*, éd. L'Harmattan, Paris, 1988, p. 150-166.
- Graziani A., «La théorie du circuit et la théorie macroéconomique de la banque», *Economies et Sociétés*, Monnaie et Production, n°6, février 1990, p. 51-62.
- Hicks J.R., «Mr Keynes theory of employment», *The Economic Journal*, juin 1936, p. 238-253.
- Keynes J. M., *A Treatise on Money, The Collected Writings of J. M. Keynes*, vol. V et VI, Macmillan, Londres, 1971a.
- Keynes J. M., *La théorie générale de l'emploi, de l'intérêt et de la monnaie*, Petite Bibliothèque Payot, Paris, 1971b.
- Keynes J. M., «Alternative theories of the rate of interest», *The Economic Journal*, June 1937a, réimprimé dans *The Collected Writings of J.M. Keynes*, vol. XIV, Macmillan, Londres, 1973, p. 201-215.
- Keynes J. M., «The "ex ante" theory of the rate of interest», *The Economic Journal*, December 1937b, réimprimé dans *The Collected Writings of J.M. Keynes*, vol. XIV, Macmillan, Londres, 1973, 1937b, p. 215-223.
- Keynes J. M., «The process of capital formation», *The Economic Journal*, September 1939, réimprimé dans *The Collected Writings of J.M. Keynes*, vol. XIV, Londres 1973, p. 278-285.
- Lapidus A., «La propriété de la monnaie : doctrine de l'usure et théorie de l'intérêt», *Revue économique*, n° 6, novembre 1987, p. 1095-1109.
- Lavoie M., «Monnaie et production : une synthèse de la théorie du circuit», *Economies et Sociétés*, Monnaie et Production, n° 9, 1987, p. 65-101.
- Léonard J., «Le paradoxe de l'intérêt et la crise de l'économie monétaire», *Economies et Sociétés*, Monnaie et Production, n° 9, 1987, p. 149-168.
- Lévy P., «L'utilité de la monnaie : des hésitations marshalliennes aux résolutions keynésiennes», *Economies et Sociétés*, Série Œconomica, P.E. n° 15, mai 1991, p. 139-165.
- Norel P., «Flexibilité financière et crise économique : une approche keynésienne», *Economies et Sociétés*, Monnaie et Production, n° 6, 1990, p. 60-79.
- Ohlin B., «Some notes on the Stockholm theory of savings and investments II», *The Economic Journal*, Juin 1937, p. 221-240.
- Parguez A., «Ordre social, monnaie et régulation», *Economie Appliquée*, tome XXXIV, n° 2-3, Droz Genève, 1981, p. 383-448.
- Poulon F., *Economie générale*, éd. Dunod, Paris, 1982.
- Riboud J., *Controverse sur la banque et la monnaie*, Revue politique et parlementaire, diffusion PUF, Paris, 1986.
- Robertson D.H., «Alternatives theories of the rate of interest», *The Economic Journal*, September 1937, p. 428-436.
- Schmitt B., *Théorie unitaire de la monnaie nationale et internationale*, éd. Castella, Albeuve, Suisse, 1975.
- Schmitt B., *Inflation, Chômage et Malformations du Capital*, éd. Economica-Castella, Paris, 1984.
- Zerbato M., «Intérêt, profit et bouclage monétaire du circuit», *Economies et Sociétés*, Monnaie et Production, n° 6, février 1990, p. 97-106.

Notes et commentaires

Disintegration of the USSR : Economic Implications

Vladimir Kollontai

The time that has passed since December 1992 has given some very interesting and important material to assess the possible consequences of the disintegration of the former Soviet Union. Of course any such assessment has to be preliminary, since not all the latent forces have as yet come into play. But several lines of approaching this topic are already apparent. Quite evidently it is necessary to analyze the disintegration of the USSR in the context of the overall economic transformation taking place in that part of the world. Next an appraisal has to be made of the internal consequences of the changes that have taken place in the former Soviet Union. Lastly it is imperative to dwell on the international implications of what has happened with the Soviet Union. But first of all – a few introductory remarks.

1) The former Soviet Union and Russia as such have a historically accumulated, complex ethnic structure. The percentage of ethnically mixed families is very high. The prevailing pattern is a very patchy one with a smeared territorial distribution of ethnic grouping among the population. This leads to a constant interaction between different cultures, religions, value systems (sometimes very fruitful, in other cases – conflictual). Imperial relations, to the extent that they existed had a great specificity since the colonies were not overseas but on the same territory. The ruled and the ruling were not territorially separated. Metropolitan power and state rule often become intermingled. This did not lighten the burden for the oppressed but it did lead to a more subordinated status for the whole population. In the Russian intelligencia this fostered two conflicting tendencies: one Western-oriented, modernizing, the other nationalistic and traditionalistic (slavianofil).

2) Another feature of Russian development that needs special mention is the historically limited role of law and social institutions in restricting the use of political power. Hierarchical structures and decision-making processes are a very strong historical heritage in Russia. Most often this tends to aggravate the struggles for political positions, strengthens illusions that political decisions can solve everything, fosters sullen passiveness by the majority of the population and creates a climate of overall lack of socio-economic initiative.

3) It is increasingly doubtful whether in today's circumstances even the analysis of stable societies can be carried out on a basis of separated social sciences: economics, by itself, politics by themselves, ideology, culture, public attitudes, etc. also. In periods of violent, profound transition this is even more so. Perestrojka and post-perestrojka have led to an unprecedented interaction between political, economic, cultural, ideological, moral change. As in all periods of fundamental change this interaction is far from always mutually reinforcing; more often it is multidimensional, conflictual, even incompatible with developments in adjacent fields. Accordingly this article will attempt to focus attention not just on economic questions and not just on individual actions and decisions but on the broader socio-economic panorama of the changes taking place and their consequences.

4) Last but not least it must be stressed that the events in the former USSR reflect (as we will see) not only a process of disintegration but also a very tangible trend towards reintegration in new forms and areas. (Economic imperatives, cultural and intellectual ties, creation of coordinating groups and liaison institutions in various fields, etc.)

*

* *

In a certain sense the disintegration of the Soviet Union was an unpremeditated even unexpected but logical result of the socio-economic transformations initiated in the country. In the years preceding Perestrojka industrial development and growing interdependence were increasingly coming into conflict with the prevailing overcentralized system of decision-making and management. The imperative need for

decentralization was one of the major factors, stimulating Perestrojka. In this context the prospect of market formation gave a serious additional impetus to centrifugal forces in the country.

The complex interaction between economic reform and the disintegration of the Soviet Union can be observed in a narrower and a broader perspective.

At a purely economic level the very prospect of creating a market economy in a very inefficient and highly differentiated society could not but lead to a breakup and disintegration of that society. Numerous uncompetitive economic units inevitably saw the danger that competition could create for them and tried to find an escape from the existing central powers that were committed to introducing the market. At the same time regional and local authorities expected through a strengthening of their political bargaining power to acquire a larger role in the inevitable privatization process, which promised to distribute enormous state riches. Later the numerous mistakes and miscalculations of the central authorities created an additional stimulus for the urge towards independence – the desire to avoid unpopular central policies and decisions.

In a broader sense many of the non-economic features of Perestrojka also tended to foster separatist movements. Perestrojka and Glasnost were in many ways attempts to overcome wide-spread fear in a post-totalitarian society and to stimulate democracy; this was primarily interpreted as a right to demand things that were previously banned. Increased local autonomy was high upon the list of such demands. Perestrojka wanted to overcome alienation in society by stimulating a feeling of identity but this very often found its expression in nationalistic and religious communities. Last but not least Perestrojka intended to bolster initiative, but in many cases this was translated into political ambitions at national and/or local levels.

As often happens in Russia the process of decentralization proceeded in a very unorderly, violent and unpredictable fashion. It constantly interacted with other factors in the breakup of the old system and at the same time became a prominent new element in the growing loss of governance (which is emerging as a central issue of the post-perestrojka era).

This is a dilemma that has to be faced by most attempts at fundamental socio-economic transformations. The borderline between the breaking up of the old system (loss of governance by the old regime) and the loss of governance in general is very fragile and difficult to perceive. The more developed and interdependent a society, the more difficult it is to define this borderline. Events in the former Soviet Union are the first attempt at drastic changes in a highly industrialized and interdependent society, moreover after many decades of totalitarian rule and non-market, even anti-market development.

The previous regimes had achieved a very high level of mutually reinforcing interaction between all the major components of the social system, its structures, institutions, ideology, legal framework, etc. Hence the impossibility of economic reform without a serious breakup of old system-economic, social, political, ideological, moral, etc. Several attempts at economic reform in the 60-ies, 70-ies and 80-ies failed because they were restricted purely to the economic field.

One of the major difficulties of Perestrojka stems from the obvious fact that the larger the scope of change the greater the possibility that the transformation process might lead up to a partial or total loss of governance. As a result of the success of perestrojka such a loss of governance in certain areas is materializing – most evidently in the disappearance of the Soviet Union. Loss of governance is inevitable if the breakup of the old system overtakes the creation of a new system of governance. This to a large degree is exactly what is happening in the former Soviet Union. The old system of planning is largely dismantled but practically no market or other economic mechanism has started to function. Hyperinflation is playing havoc with most economic processes. Corruption and irresponsibility are undermining the capacity of public administration to function. The disintegration of the USSR has eroded the legal basis of society and the possibilities of law-enforcement.

The problems of transition in the former USSR are grossly aggravated by serious mistakes in the posing of the objectives of transformation. On many issues a totally ideological approach was adopted. A major feature of such an approach (largely inherited from the recent past) is a constant use of highly abstract notions. The objectives of Perestrojka were formulated in very general terms – the achievement of “democracy”, a shift from central planning to a “free market” economy. In reality numerous stages have to be overcome before responsible democratic development can commence. Economically the main task was not to

abolish all state intervention and move to a totally free market (which does not exist anywhere in the world). The task should have been formulated as one of moving from excessive regulation on an anti-market basis to more sophisticated and competent forms of regulation in a market-oriented economy. Such a formulation of the goals would have made it possible to diminish the destruction of the system of governance and to focus more attention on the creation of a new mechanism of regulation. This was not however the case.

The problem was further aggravated by purely political considerations. All those in opposition to the old regime (and first of all the democratic forces) needed an easily understandable economic alternative to the planning concept: this presupposed a total negation of existing practices. The only such concept in their range of view was the idea of the free market, which (if the necessary conditions were created) was supposed to solve most issues automatically – or at least so it was alleged in the textbooks. Moreover such an approach was and is actively supported by Western consultants and international institutions.

The idea of regulated markets was actively criticized and practically discarded by both the democrats and the old-guard nomenclatura. The first said this was a ploy to return to the old system; the second considered this a maneuver to introduce unacceptable changes. For many years very little was done (either conceptually or practically) to introduce any serious forms of market regulation as an alternative mechanism to the existing forms of state involvement through planning. As a result of an exaggerated faith and reliance on free market forces there was a much greater loss of governance in the economy and in society than need have been.

On the whole the splitting up of the Soviet Union was primarily a political phenomenon but in had very broad roots and wide repercussions on the whole process of economic reform and market formation.

*
* *

The falling apart of the Soviet Union created in many respects a totally new situation. It cleared away some of the barriers that had been hindering economic reform. At the same time it gave rise to numerous new and difficult problems. More important still it fostered profound new development trends that will influence the course of events for many years to come.

Any balance-sheet of the events since the break-up of the former USSR will inevitably be highly controversial. This is best reflected in the shifts, that took place in attitudes, strategies and policies during 1992. If at the beginning of the year the prevailing trend was one of euphoria, and an all-out dash towards unlimited national sovereignty and independence (even autarcy), the latter part of the year was characterized by a growing understanding of the numerous costs and problems that the disintegration of the Soviet Union (and a disdain for existing interdependencies) was inflicting on all the areas of the former USSR.

The dissolution of the Soviet Union was in many respects a decisive step towards the breakup of the old system, but it also increased the danger of overall loss of governance. It accelerated the changes begun after the failure of the August 1991 putsch. As a result of the disintegration of the USSR many previously allpowerfull institutions of the command – administrative system were dismantled (All-union ministries, committees, etc.). Some of them, however were immediately reproduced at the level of the newly independent states. The size (and corresponding inefficiency and incompetence) of the bureaucracy did not diminish; in fact it increased tangibly since now the needs of fifteen state apparatuses had to be satisfied.

But independence meant also that each newly created state was much freer to choose and pursue its own strategies and policies. No previous rules having been set, the hope was that common sense and inertia would keep the newly independent states from demolishing too much of their existing mutually-useful relations and economic ties. But (especially during the first months) the desire of each political and economic unit to broaden its area of competence took undisputed precedence over other considerations. This lead in many cases to confrontational attitudes prevailing. Numerous arguments over the distribution of functions, rights and responsibilities, over subordination of armed forces, property rights, etc. ensued. The dispersal of arms control and military potentials (sometimes into irresponsible hands) greatly increased the danger of new areas and forms of conflict, armed banditry and civil wars. The main achievement of this period was probably that territorial, ethnical and other strife did not as a rule degenerate into shooting wars; non the less they kept smouldering through numerous political discussions, decisions and compromises.

Political ambitions and desires for selfassertion often lead to extreme forms of intolerance and uncompromising nationalism, destabilized the situation, created endless new areas of tension, undermined the foundations of governance. Long overdue decisions were not taken thus increasing the already existing question marks about the legitimacy of the whole newly-created decision-making process and institutions. Old rules were increasingly invalidated, established laws and agreements-challenged.

This was very evident *inter alia* in the field of economics. The prevailing mood (especially in the beginning of 1992) was that political sovereignty had to be complemented by economic independence. Accordingly in most new states there was a scramble for most of the attributes of economic independence – customs, finance, currency. During 1992 all the newly independent states of the CIS elaborated their own socio-economic strategies and policies of development with a very strong autarcic element and a scarcely veiled orientation towards economic relations with countries outside of the former Soviet Union (preferably with hard currency areas). Existing interdependencies with other former USSR regions was largely or totally ignored, if not replaced by economic warfare, protectionism, priority rules of catering to local needs, attempts to introduce local financial and trade systems, credit and/or currency controls, reglementation of trade flows, pricing, etc.

This could not but have disastrous effects. For many decades the economy of the Soviet Union had developed as a single entity and the degree of interdependence in the economy of the newly independent states of the CIS is now comparable to that of the member states of the EEC. For decades special attention was given to a more equitable territorial distribution of industrial capacities. This was one of the explicit aims of central planning and – at the same time – a very important means of fostering meaningful forms of dependence. As a result very few large or medium size plants in Russia (and even more so in other CIS countries) do not depend on deliveries from at least four-five other member states of the CIS. On a macro scale this translates itself (in the latter years of the USSR) into a very high share of national produce channeled into intra-regional exchange. For the Baltic republics and for most of Central Asia this share exceeded one-half, for Kazakhstan about 40%, for the Ukraine, 35% and only for Russia it was a bit below 20%.

This situation is aggravated by a very high level of concentration and monopoly control of production. For a very long time high levels of specialization were considered a major source of efficiency. Moreover the military complex fostered overspecialization as an element of secrecy

and of additional control over large areas of production. The recently created committee against monopoly practices has established a list of enterprises which have a monopoly position in the production of individual products. According to its findings 90% of metallurgical products are the produce of monopoly enterprises. The high level of monopolization means that there are practically very little possibilities of substituting one source of supply by enterprises are now widely dispersed among newly independent states, the rupture of economic ties between the newly independent states is having devastating consequences – shortfalls in deliveries, delayed payments, embargos, etc. – and as a final result the plummeting of production.

According to calculations made in the Russian Institute of Economic Forecasting a 10% drop in Russia's trade with other former USSR regions causes a 3-3,5% drop in national income. For the Ukraine the equivalent drop in the national income is over 4% and for most other former republics – over 7%. According to existing statistics Russia's trade with other CIS countries fell in 1992 by at least 15%. Most experts agree that the disruption of economic ties between the countries of the CIS is the biggest single factor, determining the tremendous fall last year in their production.

On a longer-term, macro-economic basis Russia and other CIS countries are faced with difficult strategic and policy choices. Disruption of trade and economic relations is creating a dangerous instability which makes economic reform very difficult to carry out. The disintegration of the Soviet Union has already changed many important parameters of economic development and market formation in Russia and the other CIS countries. It has externalized many issues, which previously did not constitute serious problems. Structural imbalances have grown tangibly. Potentially the issue of an outward orientation of the economy is gaining importance for each country. Each one has to decide whether it agrees to be dependent on unpredictable sources of supply in the CIS or whether it intends to become more self-sufficient and/or increase ties outside of the former USSR. Increasingly these countries are faced also with a choice of import substitution in one field or export promotion in another. Many other more specific issues also arise. For instance Russia to a very large degree has become overnight a semi-land-locked country with totally inadequate access to sea transport. This is becoming an additional factor in its relations with some of the CIS countries.

And all these questions have to be decided upon when market parameters are non-existent, when normal cost-price calculations and market analysis are impossible. Private enterprise cannot actively

participate in the solution of the accumulating problems since it is impossible to foresee how far political ambitions and nationalistic feelings will prevail over economic considerations. On the other hand with the degradation of the economic situation, autarcic trends are gaining force and policies aimed at priority satisfaction of local basic needs is acquiring wide support.

The desirability and even need to end existing overspecialization is apparent. But the practical possibilities of doing so are limited both by insufficient investment possibilities and by the mutual social disruption that such attempts will create. It is increasingly evident that the establishment of new enterprises as alternative sources of supply for local industries (say in Russia) will mean bankruptcies and increased unemployment in other republics, which previously were the main source of supply. Moreover a new structural strategy needs large investments which are not available. Hence the growing need to strike a viable balance between autarcie tendencies and the creation of workable conditions for the continued functioning of existing interdependencies.

Another significant factor defining a revision of previous approaches is the rapid crumbling of unjustified expectations according to which severed economic ties within the former Soviet Union could easily be replaced by trade and investments from the West. Very early in 1992 the limits of such expectations became apparent. And this was a very sobering factor for policy-makers in the newly independent countries. The need to seriously consider the possibilities of reviving old economic ties is rapidly becoming apparent.

On the other hand very unexpected mixtures of economic, social, ethnical and other problems have arisen. First and foremost among them is the question of refugees. Up till recently this has never been a serious issue. Now with growing civil strife and armed banditry swelling streams of refugees are moving in different directions across newly created frontiers. Infringement of human rights (often with an ethnical connotation) threatens to tangibly increase this flow. Added to the repatriation of armed forces from outside the borders of the former Soviet Union, the stream of refugees promises in the near future to create a very difficult social and economic problem.

In less flagrant forms numerous other social-cum-economic-cum-ethnical issues constantly arise in a decomposing ex-USSR. For instance different interpretations of citizens rights threaten the future of millions of people. What, for instance, should be the social security status (and who should pay the pensions) of workers of one nationality who have

been employed (say on contract) on the territory of another republic – now a newly independent state? All such questions were previously non-existent and now create a complex interdependence between states, which can be resolved only on the basis of good will and mutual respect.

Against a background of rapid economic deterioration in CIS countries all the abovementioned issues are forcing policy-makers in the newly independent states to take a new look at their economic problems, to acknowledge the existing degree of interdependence and to seek mutually acceptable solutions. The nationalistic euphoria and extremism of the first months of independence is giving way to more sober calculations and long-term thinking.

If a first the newly independent states were firmly opposed to any coordinating machinery and/or institutions within the CIS (viewing them as a ploy or attempt to revive old imperial practices), the last months have seen the creation of many such institutions – at parliamentary and NGO level, in many branches of industry, the acknowledgment of the rouble as an exchange currency, the establishment of a common clearing-account bank, etc. Simultaneously bilateral and multilateral agreements are being signed between different regions of the former USSR at an accelerating rate.

1992 was a period of violent and not very competent trial and error. If anything the first year of independence has shown the limits of that independence; it has taught most member countries of the CIS that they cannot individually try to solve their problems totally disregarding the problems and conditions prevailing in the other regions of the ex-USSR. The heritage of the former Soviet Union (as most historic inertia) has been shown to have a lot of strength and vitality.

But nationalistic and separatist trends continue to play a very important role in most areas. Fear of subjugation by Russia persists. Only time can tell how strong will the emerging acknowledgment of the existing interdependencies become. At present the balance between disintegrating and reintegrating trends is very fragile. Many important issues remain unresolved.

Among them is the possibility of internal disintegration in Russia itself. It is rather doubtful that the process which broke up the Soviet Union will spend themselves at the borders of Russia. Up to now the central Russian government is restraining separatism by granting various economic concessions to local authorities. This might easily undermine the future income base of the central government and/or lead to a loss of control over vital economic functions. At present wild-cat privatization is reportedly on the increase in many regions. A scenario envisaging *de*

facto power shift to local authorities (especially on economic issues) is very likely. Events in this field deserve constant monitoring. The more so that inside Russia no serious concept of regional policy or strategy or stage-by-stage decentralization has been elaborated as yet exists.

*
* *

The processes evolving in the former Soviet Union, and first of all the transition to market relations together with the disintegration of the USSR, are going to have an ever increasing impact on the world economy. Both these processes constantly interact and reinforce each other. The first stages of market formation in the USSR gave a strong additional impetus to nationalistic and separatist movements in many parts of the country. On the other hand, the disintegration of the Soviet Union is to a very large extent fashioning the type of markets that are emerging in the member states of the CIS. Whatever form and magnitude these two processes acquire, they are in the final analysis going to have a very serious impact on the world economy-disrupting some sectors, reshaping others.

This stems from the very size of the former USSR economy and its historic heritage. The former Soviet Union was comparable in territory, natural endowment and population with a whole continent, like Latin America or Africa. When analyzing economic and social processes in Russia (and even more so in the CIS) we have to think in terms of continents and not small countries – say hyperinflation or privatization not in just Bolivia or Argentina, but in the whole of South America.

In contrast to most developing countries the former Soviet Union was a highly industrialized state, with intricate interdependencies, long production chains and a high level of division of labor. The main problem facing the country previously, and its constituent parts now is not just structural or financial adjustment, but a profound socio-economic transformation in a situation when the life-support system of society already depends predominantly on industries, whose distorted features were defined irrespective of market forces and which today moreover has become technologically totally obsolescent. Previous decades of centrally planned economic growth have led to a highly autarchic, inward-concentrated type of development. To-day both market formation and the disintegration of the Soviet Union (to say nothing of world market forces) are giving a tremendous boost to more outward-oriented development.

Market formation, whatever course it takes, will by itself inevitably stimulate outward-orientated development. This stems from the need for foreign markets, technology, know-how, international standards and all the other attributes of competitiveness, which are so lacking in the Russian economy. This also stems from the numerous difficulties of market formation – limited capital accumulation, growing unemployment (and ensuing emigration), etc. – which most often have a tendency of spilling over into new demands on the world economy.

The disintegration of the Soviet Union will have analogous effects. It has narrowed down (and is continuing to narrow down) the strategy options open to the newly independent countries. At the same time it has (together with other factors) seriously shifted the balance between inward and outward oriented strategies in favor of the latter. Previously each component part of the USSR was closely interwoven with other economic units in the country; external economic ties were limited to a minimum.

Now the situation is rapidly (shock therapy) and drastically changing. Internal economic ties, production chains and interdependencies, established during decades of forced industrialization are being mutilated or torn to pieces. The overspecialization of previous years is being brutally distorted by economic policies, which give priority to local needs and attempt (in a situation of increasing shortages) to satisfy these needs by annulling previous commitments and directing available resources into channels of direct barter trade. Internal production is falling. Local purchasing power is shrinking. At the same time a distortedly high premium on hard currency makes most economic units (to the extent that they have become sufficiently independent) extremely interested in exporting abroad instead of selling for roubles.

Some idea of the size of the possible disruptions this can create can be gained from the following facts. During the last years out of all the exportable surplus of oil and gas production in Russia, about half was shipped to other republics of the Soviet Union and only the other half was exported outside of the USSR. A fall in deliveries of oil-drilling equipment from Azerbaidjan is at present leading to a fall in oil production and exports. But a big potential is still there. With the independence of the Ukraine and Kasachstan, Russia has been cut off from two vital sources of its grain deliveries (in a situation when Russia was already a large importer from world markets). The disruptive spillover into the world economy from such processes could be enormous. It could (and in many cases should) be mitigated by

an acknowledgment of existing interdependencies and the need to take into account the whole spectre of mutual problems within the CIS (as discussed above). But this (as we have seen) is not an easy matter and to what extent progress in this area can be expected only time will tell.

As things now stand Russia and some of the other members of the CIS are selling large quantities of gold, platinum, uranium, aluminium and other primary commodities as well as military equipment for hard currency anywhere they can with disruptive consequences for the respective markets. The grain imports and the needs for humanitarian aid of both Russia and the other CIS states are already too well known to warrant special elaboration.

All in all the impact, that the former USSR is going to have in the not too distant future on international economic relations is going to increase tremendously. Previously the former Soviet Union had a small but more or less stable share in world trade and financial operations. At present this share has fallen due to the overall crisis the region is going through. But in the not so distant future there is every reason to expect a rapid expansion. The longer and deeper the present breakdown of the life support system, the greater the poverty and shrinkage of internal purchasing power, the more attractive will the world markets become to the more enterprising entrepreneurs in the CIS. The previously capsulated economic potential of this area is bound to burst forward onto international markets. This potential (as it is easy to understand) consists not only of natural resources but also of the scientific and technological fall-out of disarmament.

In overall terms such an opening up of the former Soviet Union will have both positive and negative effects for the Western economy. The capital-investment demands of the CIS countries can only aggravate an already difficult problem. The market formation process (if it takes the appropriate forms) could create new outlets for many branches of Western industry, that are today battling with saturated markets. The prospects of increased raw material and energy exports from CIS countries – depending on their magnitude and terms – could either be conducive or dangerous for the Western economy.

The concrete impact that the increasing outward orientation of the former Soviet Union will have on developments in the West will to a very large extent depend on two factors: firstly, the concrete form and scale of economic reform in the constituent parts of the former Soviet Union and, secondly, on the degree to which their integration into the world economy will be left to spontaneous market forces or will be

regulated by a preconceived, negotiated strategy of changing East-West economic relations.

In existing circumstances both the former countries of the Soviet Union and the countries of the West would seem to be most interested in a slow and minimal disintegration of existing economic contacts within the CIS countries and in the adoption by them (as a whole) of more inward – looking development strategies. Only such an approach could cushion the negative consequences of a rapid disintegration of the Soviet Union on many sectors of the world economy. The short-term profits that could be reaped by some groups fostering wild-cat activities in Russia are incomparable with the likely medium and long-term losses if an active outward oriented development were to unfold rapidly in Russia (as well as in other member-states of the CIS).

Unfortunately, however, the policies pursued (if not the strategies declared) both by the ruling circles in the CIS countries and in the Western democracies will have quite opposite results.

During all of 1992 the government in Russia conducted a policy of liberalizing prices, and of imposing high taxes to overcome a growing budget deficit. In the present conditions of a highly monopolized economy this inevitably lead to a further distortion of productive and financial structures, to a tremendous shrinkage of the real purchasing power of the population and of the internal market. This seriously undermined many internal consumer markets and lead to a marginalization of numerous key industries. Monetary emission has been for a long time totally out of control. In addition a noninterventionist attitude and a *de facto* lifting of some foreign exchange controls has lead to an undisputed undervaluation of the rouble, which in its turn is very actively fostering an outward orientation of economic activities. Anyone who has anything to sell abroad does so (at the cost of established internal ties and commitments). Profits from external economic activities are such that entrepreneurs will not stop at anything to export. A severely undervalued rate of exchange coupled with internal instability are fostering a tremendous flight of capital. All this is leading to various malpractices, corruption, mafia-type operations.

During a very long period of time the strategies and policies pursued were virtually turning the economy inside out. To the extent that market relationships emerged they tend to orient all economic activities towards the external world market. New investment and production was (and is) thwarted, by exorbitant taxes, levies and numerous bureaucratic restrictions.

The situation has not seriously changed now. But the new government has proclaimed its intention to counter some of these trends. However its room for maneuver is very restricted and it is facing opposition from different vested interests.

The major Western powers and international economic organizations are supporting outward-oriented development with large doses of neo-liberal advice, demands for a more rapid introduction of the market, of convertibility as well as for a punctual servicing of the external debt. Such an approach is embedded in the whole conceptual framework of post Second World War international economic relations and foreign aid. But they do not take into account neither the specifics of the present problems facing ex-centrally planned societies, nor the possible international consequences of mistakes in handling these problems.

*
* * *

An increasingly outward-oriented pattern of development in the whole area of the former Soviet Union is going to have serious consequences for the West.

Depending on the overall strategy, and even more on the priorities and emphasis adopted in the ex-USSR countries (*e.g.* sequence of economic change, emphasis on structural change or financial adjustment, optimal ratio between economic efficiency or socially oriented development, etc.) Russia will become:

- a larger or a smaller import market for the world economy;
- a more or a less disruptive force on world primary commodity and financial markets;
- a bigger or a smaller source of migration to the West;
- a more or a less dangerous polluter on a global scale.

The tremendous changes taking place in the former Soviet Union cannot but have a very serious effect on the world economy. The character and scope of this effect will to a very large degree depend on the concrete direction and magnitude of the changes that will take place in the near future. One of the most important factors – will the trends towards national independence, separatism and economic disintegration prevail inside the former Soviet Union or will they increasingly be moderated by new forms of interaction and a revival rehabilitation of part of the previous ties?

Given the continuing differentiation of development strategies and policies in various parts of the former Soviet Union no linear, uniform

impact can or should be expected. Events in one region may lead to increased disruptions and instability, cancellations of deliveries or disruptive sallies into world markets, infringement of foreign investments or legal unpredictability. Other areas will very likely be more far-sighted and amenable; they will be more open to mutually advantageous interaction with foreign investors and traders.

On the whole, however, the concrete forms that the economic transformations in Russia are now taking are probably among the more disruptive and this will in the very near future have dangerous repercussions for the economy of Western nations. Absence of any comprehensive planning or even concise long-term strategy is leading to very inconsistent and myopic policy decisions.

Moreover, as has been mentioned the changes, taking place in the former USSR are spilling out upon the world economy a very large natural resource, economic and intellectual potential, which previously had remained predominantly capsulated. The participation of Western states in this process is very uneven. Hence major shifts in competitive positions in world markets can be expected.

Special mention in this context should be made of the possible impact of the large scientific potential of Russia (and some other CIS countries), as well as of the high educational level of their population. At present there is an increasing brain drain from Russia which promises to have a tangible impact on the relative competitiveness of individual Western countries, depending on the migration and other policies they pursue.

But there is also another dimension to this problem. The population in the former Soviet Union is far from illiterate and has a significant level of education. Educational standards are not the best in the world but they create a totally different set of issues and possibilities than those that exist in traditional developing regions. The population as a whole can articulate its desires and demands much better; en masse it has a deeper grasp of the problems of society and a more independent opinion in relation to different policy proposals. Most people understand the need for fundamental changes and are ready to endure a certain amount of hardships as long as they have a realistic hope of bettering the situation. But disenchantment can bring about a more sophisticated and dangerous explosion than could be expected in typically developing countries. At the same time the human factor is a much smaller impediment to socio-economic development than it is in most developing countries; on the contrary it is on the whole a major asset.

One more aspect of the human factor problematique deserves serious attention. Russian scientific research is disintegrating. Funding has been slashed several times and many promising young researchers are forced to seek employment in business and increasingly in shady professions, where they have to jettison a large part of their moral restrictions and values. Some are being drawn into mafiosi-type activities. The potential political and social significance of this process – nationally and internationally – is hard to assess; but it is viewed in many circles with deep apprehension.

Accordingly it is imperative to seriously analyze the interaction between the external and internal aspects of economic reform in Russia (and the other CIS countries), taking into account the interests of both East and West and exploring the possibilities of mutually advantageous, negotiated solutions.

Unless the concrete forms and methods of integrating the former Soviet Union into the world community are elaborated and negotiated, this process might easily take on a very destructive character. Production and internal markets for most consumer goods in the former Soviet Union are shrinking drastically and the prospects of their expanding in the near or medium-term future are not very bright.

As a result a large part of the entrepreneurship, that is so actively being pampered will take the form of undesired and uncontrolled wild-cat export initiatives. The military-industrial complex is giving some idea of the possible implications this can have with its widely discussed arms sales abroad.

If local producers are squeezed in between contracting internal markets and protected foreign markets a serious backlash against market formation as such can be expected. Russia by historical and cultural heritage is very wary of private property and market relations. Growing economic difficulties are creating popular support for a nationalist and/or communist political backlash (with all the international security implications this can entail).

The scrapping of social standards in the former USSR is not only a factor of social instability in that region. It is also a permanent source of international tension. Cheap labor in a vast region will inevitably translate itself into low-cost exports, lead to trade frictions, accusations of unfair trade practices, dumping, etc. Moreover a large reservoir of cheap labor in Eastern Europe will inevitably depress living standards in Western Europe, accentuate problems of migration and ethnical relations.

In the worst scenario, a desperate, disintegrated, loosely governed former Soviet Union with a very big natural, human, industrial and scientific potential (with very limited aid from the West) and no clear sense of purpose could selectively become prey to organized crime and various mafia-type constellations, operating in the world economy. Such a turn of events could easily lead to very unexpected and undesired consequences for the world.

A new conceptual framework has to be worked out to help the countries of the former USSR find their place in the world economy. This should be a two-sided process. On the one hand these countries should elaborate a strategy, defining the role of external economic relations in their overall development effort, taking into account their priorities and possibilities on the other hand the Western countries must define the concrete scope, the forms and methods of integrating these countries into the world economy, which they consider acceptable and which they are ready to foster.

An unplanned, unregulated, unnegotiated, free market solution of this problem will be extremely costly and disruptive for both sides. The level of wages in Russia and the other former USSR countries is such that in a competitive struggle they can seriously impinge on the existing social framework in the West (even leaving aside the use of semi-legal institutions and practices). In the final analysis an all-out unregulated, unnegotiated, competitive struggle of these countries for a place in the world economy will in no way be conducive either to world security or to a rational and durable international division of labor.

Given the embryonic state of market relations in the former Soviet Union (and the new problems created by the disintegration of the USSR), it is more than doubtful that the traditional sequence of economic integration is applicable and even desirable here. With the existing industrial structures, price discrepancies and distorted foreign exchange rates any lowering of external trade barriers will inevitably have very little positive effect. Accordingly another sequence of integrating the former USSR into the world economy seems more plausible. This would lay more stress (in the primary stages) on coordination and harmonization in the environmental and social fields. Specific agreements in this area could be viewed by both East and West as a commitment to further integration in the future. At the same time such agreements could stem the threat that the former USSR countries will attempt to achieve international competitiveness not so much through the modernization of their productive apparatus as by measures depressing wage costs and the environmental quality of life. The already emerging trend in CIS

countries towards lowering social and environmental standards could constitute a real danger to cohesion and modernized development in the West – through increased migration pressures, reorientation of foreign investment inflows, unfair competition, to say nothing of unpredictable socio-political upheavals. Another fruitful area of cooperation and future integration are the possibilities of joint scientific research and active Western participation in conversion programs.

The changes taking place in Russia and the other CIS countries are creating a totally new international situation. Much will now depend on the concrete development strategies and the policies, that these countries adopt. It is in the interests of both East and West to elaborate a viable concept of transition to regulated market relations in this area. Such a concept is imperative if a peaceful transformation of these societies and their negotiated integration into the world economic community is to be achieved.

Chaotic targeting on the market clearing price *

Amit Bhaduri

Indian Institute of Management, Calcutta
and, Jawaharlal Nehru University, New Delhi

Assuming incomplete information about demand and a steady flow of supply, the paper shows how chaotic price adjustment may result from an attempt to equate demand with supply. In a discrete time model, the percentage price revision of each period is supposed to be governed by the unplanned inventory change as a proportion of total supply in out-of-equilibrium trade.

Supposant une information incomplète sur la demande et un flux constant de l'offre, l'article montre comment un ajustement chaotique par les prix peut résulter d'un effort d'égaliser la demande avec l'offre. Dans un modèle en temps discret la révision du prix en pourcentage de chaque période est supposée être déterminée par la variation non planifiée des stocks en proportion de l'offre totale dans un état d'échange hors de l'équilibre.

Imagine a former centrally planned economy attempting to make its transition towards the free market mechanism. For some agricultural or natural product (e.g. oil) which has a more or less constant flow of

* Research support from the Norwegian Research Council on Science and Humanities and, from the Centre for Environment and Development, University of Trondheim is gratefully acknowledged.

supply per period, the attempt on the part of the relevant authority is to let the forces of demand and (fixed) supply determine the market clearing price perhaps because it wants to do away with a complex regime of quantity control, subsidy and import duty. However, like in most realistic situations, information is incomplete. The authorities neither know the “demand curve” for the product; nor do they have – unlike the Walrasian auctioneer – any viable mechanism to prevent trading at out-of-equilibrium prices. Thus, prices are set at the beginning of every period and, trading takes place at that price during the period. The total amount of demand *at that price* is met through appropriate adjustment of inventories (or imports). However, such inventory changes are not allowed to influence a constant flow of supply to which, demand is expected to adapt over time through suitable price adjustments. The total expenditure on the commodity at the existing price is noted at the end of each period to provide guidance for setting price for the next period. This is done according to an apparently plausible, “common sense” rule that, total expenditure divided by total (fixed) supply would more or less approximate the market-clearing price – a procedure which appears reasonable so long as total expenditure does not vary sharply with price (*i.e.* price elasticity is around 1).

Thus, if P_t = price set at period t

X_t^d = demand during period t

then, total expenditure, on the commodity during period t is.

$$E_t = P_t^* X_t^d \quad (1)$$

We postulate that the demand curve, unknown to the price-setting authority, has the simple linear form,

$$X_t^d = a - bP_t, \quad a, b > 0 \quad (2)$$

while, the supply level is kept constant at,

$$X_t^s = \bar{X} \quad (3)$$

Thus, the targeting on the market clearing price for the next period is governed by the rule,

$$P_{t+1} = E_t / X_t^s \quad (4)$$

Using equations (1) to (3) in (4), we obtain on manipulation,

$$(P_{t+1} / P_{\max}) = (a / \bar{X}) \cdot (P_t / P_{\max}) \cdot [1 - (P_t / P_{\max})] \quad (5)$$

where the maximum price, P_{\max} is reached at zero demand, *i.e.* from (3)

$$X_d^t = a - bP_{\max} = 0, \text{ implying } P_{\max} = (a/b) \quad (6)$$

Thus, equation (5) may be rewritten in the form of perhaps the most familiar chaotic equation (Devaney, 1986; May, 1976),

$$x_{t+1} = Ax_t(1 - x_t), \quad 1 \geq x \geq 0 \quad (7)$$

where, $x_t = (P_t/P_{\max})$ and, $A = (a/\bar{X})$ (8)

Moreover, if a market clearing equilibrium price P^* exists, then from (2) and (3),

$$P^* = (a - \bar{X})/b \quad (9)$$

Thus, the positivity of the equilibrium price P^* entails, $(a - \bar{X}) > 0$, which also implies $A > 1$ in (7). Alternatively, parameter A may be interpreted as the ratio of maximum demand to fixed supply. Maximum demand is reached when price is zero *i.e.* the commodity becomes a "free good". From (2), this implies, $X_{\max}^d = a$, at $P = 0$ so that, A in (8) becomes (X_{\max}^d/\bar{X}) , a ratio which is likely to be far above unity for most commodities.

It only remains to note that, for relatively low values of A , $3 > A > 1$, the convergence of x_t to its non-trivial equilibrium value in (7) is assured, so that P_t also approaches its equilibrium value P^* in (9). However, as A passes through 3, approaching 4 or above, the dynamics of x_t becomes highly complex (*e.g.* Devany, 1986; especially p. 31-38). As a result, even though a market clearing equilibrium price exists in (9) and, the price setting authorities are constantly trying to target on it according to adjustment equation (4), the prices set may behave in an apparently chaotic manner over time.

Two comments are in order. First, our argument depends crucially, not only on a linear (approximation to the) demand curve, but also on supply being exogenously fixed. The qualitative dynamics are likely to be different, if supply also depends (positively) on price⁽¹⁾. Second, it appears somewhat counter-intuitive that even when the demand for the commodity is relatively insensitive to large variations in price (*i.e.* the

(1) Since changes in inventory (or import) depends on the prices set, this also means that the extent of inventory change is not allowed to influence the fixed level of supply. The rationale is that, the authorities believe inventory changes to be temporary until the equilibrium price is reached, while long-term considerations determine the fixed level of supply (*e.g.* a fixed rate of depletion of some natural resource).

coefficient “ b ” is small), the “chaos” in price setting is not avoided by following rule (4). Since (4) may be rewritten as,

$$\Delta P_t/P_t = (X_t^d - X_t^s)/X_t^s \quad (10)$$

it follows that the percentage revision in price is governed by the discrepancy between demand and supply as a percent of constant supply. Therefore, to the extent a lower value of the coefficient b implies smaller variation in demand in response to price change, the percentage variations in prices over time may be restricted to a narrower range in general for a smaller value of b . However, the qualitative dynamics of a chaotic pattern of price setting over time would not be avoided. The prices set by rule (4) may wander “aimlessly” around the equilibrium value for an indefinitely long time.

REFERENCES

- Devaney, R.L., *An Introduction to Chaotic Dynamical Systems*. Mento Park: Benjamin/Cummings, 1986.
- May R.M., «Simple mathematical models with very complicated dynamics», *Nature*, 261, 10 June 1976.

économie appliquée

Archives de l'Institut de Sciences Mathématiques et Economiques Appliquées
Revue fondée en 1946 par François Perroux

Directeur de la Publication : Gérard Destanne de Bernis, Président de l'ISMEA

Comité de rédaction : Christian Barrère (Université de Reims), Maurice Baslé (Université de Rennes), Claude Berthomieu (Université de Nice), Rolande Borrelly, éditeur (Université Grenoble 2), André Cartapanis (Université Aix-Marseille 2), Robert Delorme (Université Paris 13), Roland Lantner (Université Paris 1), Jacques Léonard (Université de Poitiers), Manh Lan Ngo (ISMEA), Sylvain Wickham (ISMEA), Yves Younès (CNRS).

Comité international d'orientation : Irma Adelman (Berkeley), Amit Bhaduri (Calcutta), Augusto Graziani (Rome), Joseph Halevi (Sydney), Vladimir Kollontai (Moscou), Marc Lavoie (Ottawa), Jun Nichikawa (Tokyo), Luigi Spaventa (Rome), Adrian Wood (Brighton).

Secrétariat de la revue : François Denoël, ISMEA, 14 rue Corvisart, 75013 Paris,
Tél : 33 (16 - 1) 44 08 51 42 – Télécopie : 33 (16 - 1) 44 08 51 34

Administration – Abonnement – Diffusion :

Presses Universitaires de Grenoble

BP 47 – 38040 Grenoble cedex 9,

Tél : 33 76 82 56 51-52 – Télex : Unisog 9809 10 – Télécopie : 33 76 82 56 54

Tome XLVII, N° 1 – 1994

Prix au numéro : 150 FF (TTC)

Abonnement 1994 (4 numéros) : - *France* : Franco, 660 FF (TTC)

- *Etranger* : 690 FF (port avion 100 FF en sus)

Pour les numéros publiés avant 1968, s'adresser à Kraus Reprint :

Millwood - NEW YORK 10546 USA.

Copyright 1994 - Reproduction interdite sans l'accord de l'administration de la revue
ISBN : 2 7061 0582 8 – ISSN : 0013 0494

François Perroux a fondé *Economie Appliquée* pour contribuer au développement d'une «économie d'intention scientifique».

Ce «savoir économique scientifiquement contrôlé» repose sur la confrontation des observations (enquêtes quantitatives et qualitatives, travaux économétriques) et des interprétations correspondantes (visions, concepts, lois, théorèmes). La mise en lumière des «dynamiques structurelles» à l'œuvre en chaque période de l'histoire exige en effet ce travail permanent d'élaboration et de rénovation théoriques.

L'orientation générale de la revue répond à deux préoccupations : faire vivre le pluralisme, soutenir le développement d'une théorie de la dynamique économique.

Une préférence est désormais affichée pour les recherches qui contribuent à une meilleure intelligibilité de la dynamique socio-économique. La revue entend ainsi s'adapter à sa façon au développement rapide que la discipline économique a connu au cours des dernières décennies. Cette extension de la recherche économique – plutôt que l'accroissement du savoir – oblige pratiquement à une spécialisation. Privilégier les travaux qui relèvent d'une approche dynamique est une manière d'y procéder qui conserve à la revue son caractère généraliste.

La pratique des numéros spéciaux est depuis longtemps une caractéristique de la revue. Les propositions de «surveys» et de «review articles» sont examinées par la revue qui accorde une attention particulière à ceux qui s'efforceraient de présenter un thème ou un concept dans une perspective comparative, prenant en considération les développements auxquels il a donné lieu, aussi bien dans que hors de la discipline économique.

Instructions aux auteurs

La revue publie des articles originaux en langue française, anglaise, espagnole. Les articles doivent être adressés au secrétariat de la revue en triple exemplaire.

Leur longueur totale ne doit pas excéder 30 pages dactylographiées au recto du feuillet exclusivement et en double interligne. Ils doivent être accompagnés d'un résumé (en français et en anglais) qui fasse clairement apparaître l'intérêt et la valeur ajoutée de la contribution.

Le recours au langage mathématique doit être réduit à ce qui est nécessaire au raisonnement et à sa clarté ; il est recommandé de reporter en annexe la présentation des démonstrations. Les auteurs des articles économétriques s'engagent à fournir l'intégralité des données aux lecteurs qui leur en feront la demande.

Les originaux des graphiques doivent être joints sur feuilles séparées, sous la forme dans laquelle ils seront directement reproductibles en subissant un taux de réduction qui n'altérera pas leur lisibilité. Les notes de bas de page doivent être aussi peu nombreuses que possible et présentées sur une feuille séparée. Les références bibliographiques doivent être fournies par ordre alphabétique et placées à la fin du texte.

Les textes reçus par le Comité de rédaction seront soumis à l'évaluation de deux lecteurs extérieurs. Le Comité fera connaître son avis dans un délai approximatif de quatre mois. Lorsque l'article est retenu, le Comité de rédaction peut demander aux auteurs de procéder aux corrections suggérées par les lecteurs.