Social Costs and Public Action in Modern Capitalism

Essays inspired by Karl William Kapp's theory of social costs

Edited by Wolfram Elsner, Pietro Frigato and Paolo Ramazzotti

Routledge
Taylor & Francis Group

Social Costs and Public Action in Modern Capitalism

Essays inspired by Karl William Kapp's theory of social costs

Edited by Wolfram Elsner, Pietro Frigato and Paolo Ramazzotti

2006

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

Social Costs and Public Action in Modern Capitalism

The Social Costs approach to the globalised capitalist market economy has gained new relevance in recent years. The present situation is one of widespread and increasing deterioration of the social, cultural, democratic, and environmental frameworks of advanced capitalist market societies. This deterioration is indicated by the threats of unemployment, precarious working conditions and increasing income/status inequality, uneven geographical developments, and the exploitation and undermining of the institutional fabric of the society. It is aggravated by the rapid extension – at local, national, regional and global scales – of ecological disruption. So the global capitalist market economy is characterised by a great deal of instability and so-called *true uncertainty,* which largely undermine its coordinating and welfare-enhancing capacities.

The view and heuristics suggested by Karl William Kapp's seminal *evolutionary open-systems approach* is that these processes and the resulting social costs are the outcome of a widening gap between private individualistic economic and societal values or, to use Karl Polanyi's term, of the increasing disembeddedness of the economy from society.

Kapp's analysis of social costs contends that, from the perspective of the community, assigning a money value to social costs is inappropriate because prices cannot adequately reflect the full magnitude and significance of the real losses involved. In addition, innovation continuously changes them in order to *shift costs* from business to sections of society. By focusing on the market as the only possible economy, formal theory implicitly favours those economic and social interests that have most to gain from a disembedded market. Even those interests, however, are likely to suffer the long term consequences of disembeddedness. Indeed, while conflicts of interest among different sections of society are crucial to understand the origin of social costs, the notion of social cost transcends sectional interests and involves the overall well-being of society.

Kapp's approach to *public policy* suggests that the economy must be functional to society, not the other way round. *Societal goals* should be a priority for the economy, and the economy should be the means to accomplish them rather than a constraint. The performance of the economy should be valuated in terms of the societal goals that it can actually achieve.

Current analyses of the global capitalist market economy are overdue to be undertaken making use of Kapp's powerful analytic frame. *Social Costs and Public Action in Modern Capitalism* examines and applies this approach from theoretical, conceptual, empirical, policy and case study levels.

Wolfram Elsner is Full Professor of Economics at the University of Bremen, **Pietro Frigato** is PhD, Sociology of Development, University of Pisa and **Paolo Ramazzotti** is Associate Professor of Public Policy at the University of Macerata.

Contents

Contributors

Michele Cangiani
Dipartimento di Studi Storici
University of Venezia
Italy

Wolfram Elsner
University of Bremen
Department of Economics
iiso - Institute for Institutional and Social-Economics
Germany

Maurizio Franzini
Dipartimento di Economia Pubblica
University of Rome 'La Sapienza'
Italy

Pietro Frigato
Via Santissimo Rosario 7
39100 Bolzano
Italy

Regine Heidenreich
Department of General Practice - Family Medicine
University of Goettingen
Germany

Lorenz Jarass
University of Applied Sciences
Wiesbaden
Germany

Gustav M. Obermair
Department of Physics
University of Regensburg
Germany

Richard Peter
Department of Epidemiology
University of Ulm
Germany

Paolo Ramazzotti
Dipartimento di Istituzioni Economiche e Finanziarie
University of Macerata
Italy

Marco Rangone
Dept. of Sociology
University of Padua
Italy

Johannes Siegrist
Department of Medical Sociology
University of Duesseldorf
Germany

James Swaney
Department of Economics
Wright State University
U.S.A.

Introduction

Wolfram Elsner, Pietro Frigato and Paolo Ramazzotti

The renewed relevance of social costs

An old joke from mainstream economics is that economists are divided into two groups: the optimists, who believe we live in the best of all worlds, and the pessimists, who fear that the optimists may be right. The conclusion obviously is that, independently of how you view reality, you should not change the status quo.

Possibly because not all of the authors of this book are economists, the papers in this volume are based on a different premise. Based on a long-standing tradition of institutional thought, the point of departure for the studies presented here is that economics is unavoidably a normative science and that it is worth investigating precisely to achieve a better world. The book is optimistic in a different way. Its aim is to investigate how we can put to use the approach and insights of an evolutionary and institutionalist economist, Karl William Kapp, who dealt with concrete social problems by means of a very sophisticated theoretical approach, which drew on both the European and American institutionalist traditions. The concrete problems he dealt with are clearly pointed out in the title of his most famous book, *The Social Costs of Private Enterprise* (Kapp 1950). The theoretical approach is *open-systems* oriented and holistic: it is institutional and evolutionary, systematically influenced by new developments in the *natural and social sciences*, and strongly policy-oriented (Kapp 1976b). The key tenet of the papers in this book is that Kapp's insights are still extremely useful to understand economic change today and to identify appropriate policies to deal with its – ever more ubiquituous – social costs.

The social costs approach to the globalized capitalist market economy has gained new relevance in recent years, and increasingly so. The present historical situation is one of widespread and increasing *deterioration of the social framework* of advanced capitalist market societies. This deterioration is indicated by the threat of *unemployment*, precarious *working conditions* and increasing income/status *inequality*, uneven geographical developments, and the exploitation and undermining of the institutional fabric of society. It is aggravated by the rapid extension – at local, national, regional and global scales – of *ecological* disruption. What is often presented as a substantially efficient competitive process is not only unable to solve these problems. It is itself characterised by a great deal of *instability*

and so-called true *uncertainty*, which undermine its coordinating, stabilising, and welfare-enhancing capacity.

While most people will acknowledge this, a more controversial question is what originates these phenomena. The prevailing view is that they are the consequences of *'imperfections'* in an otherwise effective *'coordination mechanism'*, i.e. the 'market economy'. The view suggested by Kapp's approach is that they result from intrinsic features of capitalist market economies. They are the outcome of a widening gap between private individualist *economic* and *societal values* or, to use Karl Polanyi's terms, of the ever increasing *disembeddedness* of the economy from society and of the 'embedding' of society within the economy. The key actor in this process is business or, more specifically, it is the increasingly dominant, globalised, deregulated and disembedded hierarchical and power system of business enterprise.

It is somewhat surprising that, under these circumstances, K.W. Kapp's classical works on social costs (Kapp 1950, 1963) have been largely neglected by the economic mainstream orthodoxy. The term 'social cost' is, in fact, fairly commonplace in economic discourse since the development of the *'New Institutional Economics'*. In Kapp's view, however, social costs have little to do with the absence (or inadequate definition) and enforcement of *property rights*, and dealing with them requires a definitely more active *public policy action*. Social costs are, according to Kapp, losses suffered by sections of the community (including business itself), or the community as a whole, as a consequence of unrestrained business activities. They involve the living conditions of people as they are affected by working conditions and by all sorts of environmental disruption and downgrading of institutions of social coordination and cooperation.

Social costs are all but the exception, in a capitalist market economy: they generally arise when business firms *shift* their *private costs* on to other parties, be they workers, consumers, neighbourhood communities, the public sector, the environment, or society as a whole. They do not involve a merely distributive issue, however, whereby the cost to one party is a gain for the other. Quite to the contrary, in so far as costs are shifted not only on third parties but also on common resources, they involve a cost for the community as a whole, including – in some instances – those who have generated them.

Although there are instances when these costs can be monetized, it is not possible to avoid them by relying on relative prices alone. The reason for this is not that the market lacks the background institutions (i.e. individual property rights) that would allow it to be 'efficient', thereby either determining 'wrong' prices or no prices at all as the neoclassical 'new institutional' transaction cost and property rights school suggests. A different perspective is required.

In order to evaluate the performance of markets – thus also social costs – it is essential to understand that they are *not isolated systems* (Kapp 1976a, b). They interact with the social and natural environment they are a part of. *Relative prices* depend on the specific features of this interaction, quite independently of relative scarcity and rational need satisfaction. How the interaction is defined – thus whether the economy is functional to society's needs or the other way round – is part

and parcel of public policy action. In this perspective, social costs may be conceived of as the indicators of the gap between the economy's values and society's values. If they could all be monetized, society would be a sub-system of the market economy rather than the other way round.

This is what links these costs to the above mentioned deterioration of the human, social and ecological frameworks of today's societies. Disregard of social costs or attempts to *reduce them to monetary values* do not remove them: they subsume societal values to economic ones; they subsume society to the economy.

Let us list a few major developments that occured since Kapp published his works.

Features of the globalized 'new' economy

Six major dynamics of change deserve attention.

1. *Technology and organization of production*: Technological change, especially in communication, has determined a widespread transformation in production, distribution and consumption. The development of *net-technologies* as the technological base of the socio-economy has added momentum to the increasing *fragmentation of value-added chains*. This implies that decisions are much more *interdependent* than they were in the past. Owing to complementarities in production, arm's-length (market) transactions alone are inadequate because firms need to cooperate in order to produce complex products and to innovate. Hierarchical (firm) transactions alone are also inappropriate, because they undermine the flexibility that technology allows and that innovation requires. The diffusion of alternative organizational setups, e.g. networks, provides some potential solution beyond those two extreme cases of the 'market'. However, the uncertain and turbulent environment has reduced the control that single actors have over their specific activity. As a result single actors attempt to reinstate and exert their *power* over the dispersed production process. This has two implications. First, trust and collective learning are hindered, thereby determining a further increase in strong, or true, uncertainty. Second, firms resort to internal flexibility – intra-firm flexibility – in order to make up for the lack of control that external flexibility – inter-firm and customer-firm flexibility – entails. This leads to a lack of security in working relations: workers cannot be sure that they will keep their jobs; in order to avoid unemployment, they are likely to accept a deterioration of their wage and of their working conditions. Although internal flexibility is usually justified by the need to adapt to new technology, the opposite is likely to occur: the organization of production is made ever more flexible in order to cope with uncertainty, in a way that unevenly allocates the costs and benefits among the capitalists and the employees.
2. *Globalization*: National economies are increasingly interdependent. International trade flows have substantially grown since the end of World War II. The internationalization of production has also grown. It originally consisted in investment to supply foreign markets or to gain access to resources

that were not available in the home country. It has subsequently become the international dimension of the above-mentioned fragmentation of value-added chains. These processes were enhanced by the progressive deregulation of trade and financial movements, under the influence of international agencies such as the WTO, the IMF, the OECD as well as of nation-state institutions. As a consequence, constraints on the balance of payments – and thus on competitiveness – have become more stringent, thereby forcing nation-state efforts to comply with the emerging rules of the new world economic order. In most instances, when technological capacity was lacking, this has reinforced attempts by firms to achieve competitiveness through cuts in production costs, i.e. through the mere redistribution of value added.

3. *Decline of the welfare state*: The social compromise of the Keynesian era – the thirty years following the end of World War II – was ended by a range of circumstances. The above-mentioned features of technological development require greater flexibility. Workers have to adapt to the uncertainty involved in the new organization of production, which means that they must be willing to change their jobs more frequently than in the past. They are also expected to adapt to the new learning requirements that innovation involves, which also means that when innovation is fairly radical, they risk either permanent unemployment and/or a downgrading of their status. In this perspective social security 'from cradle to grave' is viewed as an impediment to adaptation and to the efficient working of the labour market. Furthermore, high capital mobility forces countries to keep their interest rates in line with those of international markets. Since this determines financial constraints on domestic expenditure, private business puts pressure on governments to keep welfare expenditure low so that they will not have to bear high interest rates when they seek the finance for their real or financial investment.
4. *Ideological shift*: The theoretical ambiguities underlying the Keynesian and welfarist consensus of the post-World War II period precluded a thorough explanation of the processes here depicted. As a result there has been a shift, in academic circles as well as among political parties and public administrations, towards a '*neo-liberal'* view of the economy. An idealized view of the market was used not only to interpret economic relations among private agents but as a benchmark to assess what policy makers do, the underlying assumption being that policy makers act in accordance with the 'political market'. The claim that the latter was opaque suggested that the less political intervention the better. Thus, not only was welfare claimed to preclude the efficiency of the labour market. It was also claimed not to meet the real needs of people in an efficient way. The prospected solution was that individuals should seek welfare provision on the private market, freely choosing what best fit their needs. Similarly, since private markets were claimed to be less opaque than political ones, public provision of goods was deemed inefficient just as the pretence to publicly coordinate markets.
5. *The collapse of the State-socialist system*: Following the events of 1989, Eastern European countries sought to integrate their economies in the system

of capitalist market relations. This implied a transition process which involved the substitution of market-related institutions for the ones that prevailed in the former regimes. In most instances, a naive view of the market made the governments of those countries confident that market institutions would arise as soon as the former regimes were dismantled. By opening up to foreign trade and foreign direct investment Central European countries were, albeit to different degrees, integrated in the globalization process and in the transformation of production processes which was occurring in Western industrial countries. This integration, however, was accompanied by the rising commodification of labour, unemployment and social insecurity. The absence of social buffers to absorb the impact of this transformation implied transitions to markedly disembedded economies.

6. *Environmental disruption*: Economic growth involves the depletion of existing non-renewable resources as well as the saturation of the earth's absorbing capacity to assimilate wastes. The effect of the former is a growing world instability, associated to the pursuit of control over those resources that are crucial for economic activity. The effect of the latter is particularly evident in terms of climate change. The instability of economic relations outlined above and a rather unexplained technological optimism led to postpone attention to these issues. Technological optimism also underlies lack of attention towards another aspect of environmental change. Technological progress in a range of fields may be beneficial but it may also lead to dramatic effects. Nuclear power, even when it is not used for armaments, may have long-lasting consequences on the environment. Chances are that genetically modified organisms may also cause hitherto unpredictable effects. Research and development in these, as well as in other, fields of inquiry is basically carried out according to 'market' criteria and its results are commodified by assigning property rights that transform public property and knowledge into private property and knowledge. This reduces the possibility to control how these technologies are used and whether different uses or altogether different technologies would be socially efficient.

As we mentioned above, the changes outlined here have led to a number of problems. They include precarious working conditions – involving the threat of unemployment, low wages and a weaker enforcement of basic rights in the working place – and precarious living conditions, since the above situation is complemented by less public action to support people's welfare, rights and capabilities. They also include environmental disruption. Finally, owing to the rise in uncertainty and the decrease of social commitment and social institutions already outlined, they also include low growth rates as recent international surveys on the levels of trust and their impacts on economic performance demonstrate. The causes for these complexes of social costs are the *incapability* of markets and of market-centred policies of national governments and international agencies to *coordinate* economic activity effectively and to support joint learning investments and the building of trust and commitment.

The issue, therefore, is to provide an interpretation of the above sketched events that allows us to identify the means to prevent them and to overcome their negative consequences.

Two interpretative frameworks

The conventional way to look at these changes is to identify a range of *'exogenous' shocks* on the economy. Thus, according to this view, technological change forces firms to adapt to new efficiency requirements. Adaptation to technological change depends on how efficient the 'market' is as a coordinating mechanism. Market failures may preclude efficiency. Consequently, governments – and international agencies – who are concerned with efficiency must ensure competition and determine institutional arrangements to prevent or overcome other forms of market failure. *Externalities*, i.e. effects on third parties that the market does not register, are a typical case of such a failure. Under these circumstances, public action is suggested that consists either in assigning property rights to stakeholders, so that market transactions will eventually allow compensation, or in acting directly – e.g. through a Pigou tax – to make up for the market failure.

This emphasis on efficiency is not to deny that policy makers may be socially concerned. Despite ideological scepticism concerning the role of governments, public action may actually focus on issues such as equity or social cohesion. It may therefore act in favour of some social actors, possibly the 'less fortunate' ones, but any such action must take into account an equity/efficiency trade-off: social goals generally constrain economic activity.

Kapp's alternative view is that the negative consequences of recent economic evolution are not side effects. They are the social costs that individuals and communities bear as a consequence of the unrestrained – or improperly embedded – activity of private business. Thus, they are the point of departure for any policy agenda that is independent of those vested interests. It is appropriate to elaborate on this issue.

Consider flexibility in the labour market. It is fairly intuitive that it allows business to shift the costs of the fragmentation of value-added chains, of resulting uncertainty and coordination failure, on workers and on the community. It is far less intuitive that the latter should bear those costs instead of firms. Truly, the case may be made that business should be relieved of the costs it has to support in order to be competitive, and that it is appropriate from the community's point of view to transform private business costs into social costs. To do so, however, requires an appropriate *accounting* of those costs. This is a fairly straightforward task if we look at it from the point of view of business. Costs are quantifiable: the more they get shifted – thereby turning from private into social costs – the higher money profit (the lower money loss) is.

On the other hand, from the perspective of the community, assigning a money value to social costs is more complicated. The redistribution associated with cost-shifting may well affect household (money) income but it may produce other sorts of effects as well, e.g. on the *health* of workers. Truly, providing compensatory

health care may lead to a quantification of these costs. It is also possible to place a price tag on the psychological and social consequences of suffering and impairment. One should be aware, however, that such a procedure is highly discretionary. The 'market efficiency' claim, for instance, assumes that prices provide information in terms both of vendibility and serviceability, which, at the very least, is open to dispute. Even if we were to accept this assumption, however, prices would still be unreliable: relative prices depend on the *distributive effect* of cost-shifting, so it is not clear whether the unit of account to assess social costs should be the relative prices that precede or those that follow cost shifting.

The above considerations are restricted to a static context. In an *evolutionary* one business continuously attempts to devise technological and institutional innovations that will increase its profitability, and devising ways to make the labour market more flexible is just one way to innovate. Consequently, relative prices constantly change and, along with them, the above-mentioned distributive effects. This raises some questions as to what public policy should focus on. Should it be ex post compensation or the determination of a legal-economic nexus that anticipates the consequences of innovation? A rather different approach would be to focus less on which social section is to bear the cost of change than on the direction of change itself. The underlying premise is that, contrary to the conventional approach, whereby the market must adapt to exogenous shocks, economic circumstances interact with a range of elements: political, social, environmental, cultural and technological. Although, for simplicity's sake, it is sometimes convenient to isolate them and discuss each one of them separately, they are part of an overall process and it would be misleading to consider them independently of each other. There is more to the unreliability of markets than the unreliability of prices.

This same interdependence – which reflects the *systemic openness* of the economy – shows a further drawback of market-centred approaches. A typical Coasian claim is that it is sufficient to assign property rights to stakeholders, so that whoever is prejudicial to someone's property may be held responsible for his/her action. The fundamental problem of neoclassical Coasian – New Institutional – Economics is that the most basic and ubiquitous circumstances in the real world are not the result of a specific agent's conduct but of a range of interdependent conducts. Some of these conducts might not even be harmful were it not for the uncoordinated and concurring actions of other actors. Supposing it were possible to assign property rights on polar bears, who should be held responsible if they were to starve because the Arctic Sea ice shrinks?

Systemic openness also suggests that we do not focus on direct causal links alone. Change may be originated – or influenced – by purposive action but it often has a dynamic of its own, based on the interaction of factors that lie *beyond the control of individual actors*. In an interdependent evolving system, it is therefore difficult to discern whether something could be different from how it actually is, thus whether its negative connotation actually represents a cost suffered by someone to the advantage of someone else. Consider, for instance, how firms in the 'New' economy try to shift the costs of strong uncertainty, non-coordination and

lacking collective action capacity towards each other (may it even be a mutual blockage of action, or the incapability of innovative action, including technological lock-in) with the result that they are all at a loss because of the overall rise in uncertainty and decline in trust. This is why a policy may be required that establishes greater confidence, coordinated expectations, social commitment and, thus, stability in the future, eventually reducing those costs in what appears to be a positive-sum game.

This issue is important because it points to another important feature of Kapp's thought. By focusing on the market as the only possible economy, formal theory implicitly favours those economic and social interests that have most to gain from a disembedded market. Kapp suggests, however, that those same interests are likely to suffer the long-term consequences of disembeddedness. Thus, while conflicts of interest among different sections of society are crucial to understand the origin of social costs, the notion of social cost transcends sectional interests and involves the overall well-being of society.

This *substantive view of the economy*, as opposed to the formal one of conventional theory implies that the economy cannot be considered independently of the society it is a part of. More specifically, the economy must be functional to society, not the other way round. *Societal goals* should be a priority for the economy – and the economy should be an enhancement of societal opportunities – rather than a constraint. Thus, the performance of the economy should be valuated in terms of the societal goals that it can achieve, and little matters, in the first instance, whether the economy fails to meet social requirements because the market does not function properly, because institutional truces among social sectors prevent change or because of some other reason.

The identification of societal goals – and, symmetrically, of social costs – leads to a *cognitive dimension* in Kapp's approach that should not be disregarded. Cognition is not always associated with positive-sum games, at least not in terms of monetary payoffs. It took some time before firms were forced to use mechanical presses with a two-hand switch so that workers would not get their hands caught in the machines. Until then, the cost of speed in production was 'paid' by workers with accidents on the shop floor. Since then, the social cost of such injuries has been shifted back to firms. The cognitive relevance of this case is that: (a) perception that production could be carried out in a different way was conditional to the identification of a special type of social costs and to their substantial reduction if not elimination; (b) the value assigned to the health of the workers was deemed superior to the value of productivity; (c) this new arrangement was deemed appropriate through a (cognitive) shift from money accounting to social cost accounting. The latter approach framed the problem in terms of a positive-sum game whereas the former was more likely to frame it as a zero-sum (merely distributive) game or even a negative-sum game, in so far as the new arrangement negatively affected productivity and growth.

These issues suggest that a policy centred on the notion of a trade-off between equity and efficiency is misleading for a range of reasons. First, if societal goals are the priority, other indicators of economic performance – such as those pertaining

to efficiency – may point to the existence of constraints but they cannot be considered on the same standing as the goals themselves. In so far as these constraints exist, they should be viewed as impediments that policy ought to remove or bypass, not as features of a status quo that society is subject to. This statement may appear less awkward to economists with a conventional approach if they acknowledge that a range of features of the economy – e.g. technology, preferences, institutions – can be assumed to be exogenous only in order to simplify their models, not because this is actually the case.

Second, precisely because the above features are not exogenous, the economy cannot be reduced to a device that allocates given resources according to given preferences, technology and institutions. Consequently, equity – in the restrictive sense of distribution – and efficiency are not the only possible goals. Whatever has to do with the material reproduction of society and with how this relates to societal values – what K. Polanyi would refer to as embeddedness – has to do with the goals assigned to the economy. Societal goals involve not only the choice of what goods should be produced and the techniques used to produce them but also what is generally referred to as the quality of life, along with the freedom required to pursue it. There is more to societal goals than equity.

Third, the variety of social goals and the endogeneity of preferences make it more difficult to identify a social welfare function than is the case in the conventional approach. The issue is not to somehow take account of given individual preferences – possibly as they are expressed by political votes – but to allow actors to identify and formulate them, to allow the community to deal with inconsistencies among individual preferences as well as between preferences and social values. This is symmetrical to identifying social costs. In so far as this process occurs both within the economy and within other domains of society, it cannot be assumed away from economic inquiry. The characteristics of democracy are also an economic issue.

Who knows and who chooses turns out to be a key issue, at the borderline between economics and political science. It is also a link between Kapp's lifelong work, which ended in 1976, and Amartya Sen's more recent depiction of development in terms of *capabilities*, i.e. the positive freedom to choose how to conduct one's life. Despite different emphasis on specific issues, there is a range of elements which suggest that the two authors, albeit independently of each other, have been pursuing a common research programme.

Kappian perspectives on today's globalized capitalism – the papers of this volume

In the above, we have provided a short outline of the problems that most capitalist economies are confronted with and we have suggested that Kapp's theory of social costs provides a broad, comprehensive and highly relevant approach to the negative consequences they lead to. In many instances we have already drawn on issues that the individual papers of this volume address. We will provide a brief survey of their main arguments now.

The first part of the book is concerned with theoretical foundations of open systems, social costs and entropy generation, and capabilities of coordinated and deliberate action.

Kapp's approach to social costs can be appreciated only as a part of his general theoretical framework which is clearly an institutionalist and evolutionary one. This issue is at the core of *Michele Cangiani*'s paper. Cangiani stresses the *open-system approach* followed by Kapp and the dehumanisation of economics that follows when conventional scholars depict the economy as a closed system. In this perspective, openness involves that collective values co-exist with individual ones and that social values co-exist with individualistic ones. A societal and economic coordination which relies on the 'market' favours business-centred values at the expense of society. The Kappian alternative to market coordination is *societal planning*, as Cangiani stresses, which is meant explicitly to set out the social and collective priorities that must underlie economic activity.

While Cangiani's paper situates Kapp's theoretical contribution in the context of a more general discussion of social institutions and planning, *Regine Heidenreich*'s paper provides a focus on Kapp's *cognitive premises*. She points to the key role that cognition plays in the understanding of economic action, especially technology, ideas, markets and preferences. Subsequently, she stresses its relation with the identification of social costs, which requires their social recognition by means of a twofold process: scientific dialogue and *collective decision-making*. Here, Heidenreich identifies what remains an open issue in Kapp's approach, namely which social or political agent is supposed to identify – or to arbitrate on disputes over – social costs: given that divergent interests co-exist, what justifies the fact that some are going to prevail over others? Although Heidenreich does not provide a solution to this 'anti-liberal fallacy', she suggests that Kapp's analysis of political processes of recognition and correction of social waste and the general direction of his policy prescriptions may have to be further developed today.

Drawing on a discussion between Beckerman and Kapp, *Maurizio Franzini* points out that the conventional and Kappian notions of *social costs* have little in common. While conventional views are strictly related to the efficiency of the market, Kapp's emphasis, according to Franzini, is on the *violation of social rights*. The elimination of this type of social costs is not incompatible with the market: just as in Coase's efficiency proposition, the issue might be one of individual property rights. Problems arise, however, in so far as the 'market' relies exclusively on mere 'willingness to pay'. Extra-market remedies may therefore be required, but this raises the issue of who decides and what the decisions should consist in. Franzini's suggestion is that insights on this issue may come from recently revitalised research on *deliberative democracy*.

The second part of the book is concerned with policy. *Wolfram Elsner* focuses on the scope for public action following the failure of the market as a coordinating instance under modern net-technological, value-added chain, cluster and network conditions. The paper provides an account of how changes in ICT increase the degree of complexity and interdependence among actors, thereby leading to a rise

in uncertainty in the economy. Under these circumstances neither traditional – arm's-length, or hierarchical, or power-led – nor simple network forms of coordination can cope with such complexity in a self-sustaining evolutionary process. Strong uncertainty, mutual blockage of basic innovative action and technological lock-ins, therefore, are fundamental social costs of the contemporary de-regulated and globalized economy. They indicate widespread Pareto sub-optimal outcomes. Elsner does not draw the conclusion, however, that strong public intervention in substitution of jointly learned coordinated action of private agents is required. Quite to the contrary, a *'lean' interactive and 'institutional' policy* is outlined, based on pecuniary and non-pecuniary incentives, and on *futurity* as well, i.e. on the establishment of cooperation-enhancing processes and the learning of institutions to deal with social values and society's structure of 'merit' goods. This leads to a new hybrid governance approach of 'good' network (self-) governance and an institutional policy that is interactive with the networked interaction processes of the private agents.

James Swaney focuses on *environmental policy*. After a brief outline of the differences between market efficiency and community standards, he discusses the relevance of two non-market environmental thresholds: the limit to the supply of natural resources and the limit to the assimilation of wastes. He argues that technology may postpone but not stop the earth's trend towards these thresholds, i.e. towards what Kenneth Boulding has referred to as the 'spaceship economy'. Swaney then moves on to policy issues. He criticises Coase's approach in the light of Coddington's asymmetry between commodities and discommodities, making the point that while there usually are economic incentives to create the former, there are very few ones to avoid or eliminate the latter. This is the reason why, despite the shortcomings of price-centred policies, Swaney argues in favour of Pigouvian green taxes.

Gustav Obermair and *Lorenz Jarass* discuss the scope for *environmental taxes* in the German case. They begin their argument in favour of public policy by stressing that, given insufficient information and, consequently, the existence of a variety of local optima, achieving the 'greener valley' may require too much hardship by all social and economic actors. Following a discussion of the complementarity between normative and price instruments, they argue in favour of taxes that favour ecological and labour-intensive policies. Such measures were originally proposed by former EC-President Jacques Delors in terms of a tax shift: higher taxes on fuel and lower taxes on labour. Obermair and Jarass examine how this suggestion was eventually put into practice in Germany. Their general conclusion seems to be that the German case was successful, despite some rather grievous shortcomings.

Paolo Ramazzotti and *Marco Rangone* discuss 'voluntary' *unemployment* as a special case of a social cost. Workers who choose not to accept jobs that do not meet their expectations generally pursue goals that are socially valued, e.g. social mobility: by trying to improve their social status they indirectly – and, generally, in an unintentional way – tend to ameliorate the economic conditions of their society. As single individuals they may be mistaken about the market, when the

jobs they want are not available, but precisely because of the social values that underlie labour supply it is misleading to claim that the cost they have to bear – unemployment – is a private rather than a social one. The authors subsequently examine the demand side of the labour market. They argue not only that the technology underlying labour demand should not be assumed to be exogenous, but that its creation should not be claimed merely to depend on consumer requirements. Both private and public technological research tends to interiorise business goals on the grounds that they are beneficial to society as well. Thus, when claims arise that may prejudge profit, they tend to be perceived as constraints on economic activity. Technological change is sought to remove them and to restore profitability, regardless of the other interests at stake. Here again, there is a social dimension to unemployment that begs for appropriate public action.

Pietro Frigato shifts our attention towards *health economics* by focusing on two cases of what Kapp termed 'hybrid fields of investigation', i.e. those disciplines that arise in order to deal with issues that lie beyond the boundaries of conventional disciplines. The problem with these hybrid fields of investigation is that, while they do acknowledge the need to deal with new issues, they also exacerbate scientific segmentation, in stark contrast with the need for inquiries based on an open-systems perspective. Frigato stresses how '*socio-economic epidemiology*' provides a great deal of clues to the understanding of recent issues. At the same time, he argues, its difficulty in trespassing its own disciplinary boundaries precludes it from appreciating all the connections that Kapp's work would suggest. Similarly, '*occupational safety and health economics*' provides interesting insights which are apparently in no relation with the results of research in socio-epidemiology. The implication, from a policy perspective, is that single measures are proposed that seem to miss the general picture.

Last but not least, *Richard Peter* and *Johannes Siegrist* examine the *social health costs* of recent historical changes in the labour market, with special reference for the European transition countries. In a typically Kappian attempt to link policy suggestions to an interdisciplinary scientific discourse, they examine the relation between precarious work and health. After a survey of the literature, they lay emphasis on the potential gap between work effort and expected reward that incomplete contracts lead to. This gap, and the stress it builds up, activates direct and indirect mechanisms which eventually cause damage to the health of the workers. Their policy suggestions, which are grounded also on the data concerning East European countries, parallel their interdisciplinary inquiry, ranging from measures on the quality of the working life or on income distribution to preventive medicine.

Kappian themes and public policy

The policy objectives in the essays vary in scope, testifying to the richness of the approach fostered by Kapp. In some instances they consist in very specific goals, to be pursued by standard measures, e.g. Pigouvian taxes. In other instances they include action which is generally not taken into account by conventional theory.

Thus, while they acknowledge the need for increasing investment in research and development and in education, they also stress that the direction of – the goals pursued by – R&D or education must be *functional to social, rather than mere business*, concerns. Similarly, a range of measures outlined transcend conventional distinctions between the economy and other dimensions of society. Owing to near-decomposability, policy-makers may occasionally identify and circumscribe a sub-system and focus on measures that are appropriate to its evolution within and as a part of societal evolution. Thus, in some instances they may identify measures which are in many ways like those that a conventional economist would deem appropriate. They must acknowledge, however, that those measures are appropriate only under the specific circumstances of a given moment and of a specific economy. Policy measures that are appropriate today may prove to be inadequate tomorrow.

In general, both private actors and policy makers seek solutions to problems. In so doing, they learn. As their knowledge changes, they tend to perceive problems differently. Thus, choice contexts – those of single actors and those of policy-makers – change over time as a result of both real and cognitive evolution. Under these circumstances, policy-makers need to acknowledge that whatever they do affects what and how people learn. Consequently, it will eventually affect how all actors – including themselves – perceive reality.

In this perspective, while the features of the globalized 'new' economy involve problems that specific policy measures can and must address, they also raise dramatic questions concerning the kind of life people want to live and the priorities that must underlie purposeful action. They question not only what conventional economics usually takes for granted but also internalized beliefs concerning the overall relation between economic and societal values. Precisely because economic policy requires a reassessment of these issues, it is strictly connected to the issue of *participatory democracy*. The social costs approach does not provide a ready solution to all of the above issues. It does provide a framework, however, that puts these issues in the context of a comprehensive analytical and value perspective.

In sum, it is widely acknowledged today that the recent development of the globalized and de-regulated 'new' economy has increased interdependence and complexity and, thus, strong uncertainty, and has deteriorated trust, commitment, coordination, institutions, stable expectations, and social capital. At the end of complex causal chains, therefore, a drastic increase of social costs in human, social, institutional and ecological terms is generated. Current analyses of the global capitalist market economy are overdue to be undertaken making use of the powerful analytic frame of Karl William Kapp's open systems economics, including the conception of social costs. Most surprising, though, the analytical devices that Kapp, with his evolutionary institutional political economy, has developed for this task, i.e. *open systems*, *social costs*, *deliberative democracy* and *social planning* theories, have been largely undervalued for a long time. So time is due now to make a new start with social costs analysis – applied to contemporary ecology, economy, society and policies. It is towards this task, and the analytical and policy opportunities it contains, that the present volume wants to start off, on

theoretical, conceptual, empirical, policy and case-study levels. A starting point, we do hope, for more powerful analyses to come, that bring back to the core what certainly is one of the most powerful analytical economic devices – and what displays a most prominent dynamics in today's real-world economy – i.e. social costs.

Bremen, Bolzano, Macerata
Wolfram Elsner, Pietro Frigato, Paolo Ramazzotti

References

Kapp, K. W. (1950) *The Social Costs of Private Enterprise*, Cambridge, Mass.: Harvard University Press.

——. (1963) *Social Costs of Business Enterprise*, Bombay: Asia Publishing House.

——. (1976a) 'The Open-System Character of the Economy and its Implications', in K. Dopfer (ed.) *Economics in the Future: Towards a New Paradigm*, London: Macmillan.

——. (1976b) 'The Nature and Significance of Institutional Economics', *Kyklos*, Vol. 29, n.2: 209-232.

1 Freedom to plan: on Kapp's institutional outlook

Michele Cangiani

Karl William Kapp's theoretical achievements are both consistent all along his life and linked to the changing actuality of capitalist development. A continuity is clearly noticeable in his critique of the method of mainstream economics, as well as in the issue he correlatively raises, of the efficiency of the market system, and more in general 'of existing institutions', as he says in his posthumous article about 'The Nature and Significance of Institutional Economics' (Kapp 1976: 216). An institutional method, explicitly opposed to neo-classical theories, characterizes, in fact, also Kapp's early works, culminating with *The Social Costs of Private Enterprise* (1950), and even his Geneva thesis (Kapp, 1936).

The interrelation between methodological concern and political commitment, which is evident in Kapp's approach, is typical of what is commonly denoted as the 'substantive' point of view, and can be traced back to the 'classical' institutional critique (by 'Veblen and his followers', as Kapp says) of the market-capitalist system as well as of neoclassical economics. In this sense, Kapp's 1976 article coherently develops the basic questions he afforded in his first publication (1936); and the theory of 'social costs' appears as an important milestone along that run, precisely because it corresponds to the deep and consistent inspiration of a lifelong work.

There is, on the other hand, a development in Kapp's reflections, which can be referred also to insurgent problems and institutional changes along the history of contemporary society: from the economic and political crisis and the new features of the market and business organization in the inter-war period, to the global crisis of the post-war period of economic growth in the middle of the 1970s.

In the present historical phase, which started with this latter crisis, at least two among the tendencies Kapp highlighted have become even more important and manifest. Firstly, capitalist development is paralleled by a growing burden of social costs, social uneasiness and inequalities, waste and environmental disruption. There is then the impairment of democratic institutions and individual freedom, as a general tendency, which can remain in abeyance or be contrasted in the boom time, but is stirred up by economic depression and the need for controlling social discontent. These two tendencies strengthen each other, as is logical, and nowadays easily verifiable.

The institutional-substantive paradigm

Karl Polanyi, at the end of his lifetime, annotated and conserved among his papers two writings by Kapp (*Karl Polanyi Archive Catalogue* 26/6). The first one is the article 'Social Costs and Social Benefits – A Contribution to Normative Economics' (1963). Polanyi notes 'substantive concepts' in the title page; in the first page, he underlines Kapp's proposal of discussing 'the normative implications of social costs and social benefits for the formulation of economic policies and economic planning'. His attention seems again attracted by Kapp's analysis of the distinction between the 'formal' and the 'substantive' concepts of 'economy' (where Polanyi's 'The economy as Instituted Process' (1957a) is mentioned), and its bearing on the problems of social costs and benefits and of 'the starving of the public sector in modern affluent societies' (Kapp 1963: 190).

The second of the two writings Polanyi read and underlined is a typed research project ('American Institutionalism: The System of Economic Analysis of Veblen and His Followers') Kapp presented to the Brooklyn College, probably at the beginning of the 1960s. Kapp's purpose – 'to systematize institutionalist analysis'– is grounded on the idea that institutional thought developed 'an essentially common view of the economic process', from Veblen and the American economists, who were influenced by Peirce's and Dewey's pragmatism, to Gunnar Myrdal and François Perroux in Europe (Kapp n.d.: Chs 1 and 2). As Kapp makes clear in his account, the institutional paradigm – contrary to a diffused preconception (see e.g. Yonay 1998; older examples are mentioned in Gruchy 1947: 605) – is not reducible to an 'impressionistic description of an ever changing economic environment' (Kapp n.d.: Ch. 1), or to a demand for empirical data, but represents 'an alternative analytical framework' (Kapp 1976: 229), a theoretically fruitful blueprint, endowed as it is by a consistent methodological and conceptual core.

According to Kapp, the functioning of the economic order is considered as a whole by institutional economics ('holistic economics': Gruchy 1947: 594 ff.), and 'the social or collective nature of all human action' is stressed. This wide theoretical scope of institutional thinking doesn't hamper empirical accounts. On the contrary, it is this way, by analyzing the 'institutional components' of economic relations, Kapp observes, and not by keeping to 'a study of price and market phenomena', that it is possible to understand the economic reality as a process going through 'structural changes' – changes concerning the 'institutional structure', Polanyi would say. On this basis factual data can be meaningfully collected and interpreted, and such aspects of capitalist development as economic instability, corporate structure and corporate finance, collective bargaining, 'mixed economy' and 'administered prices' can be explained (Kapp n.d.: Chs 2 and 3). We will see in the next paragraph that Kapp adopts this line of thought as early as 1936, at the very beginning of his career; he joins, thus, the lasting tendency of institutional, critical economic thought, which he continues and develops all along his life. (For a comprehensive reconstruction of Kapp's thought in relation to institutional economists, see Heidenreich 1994.)

In that same year Adolf Löwe characterizes as follows the institutional method: 'pure economics' consists in the 'theory of choice', but the problem of the data of choice – of 'their occurrence, structural order and evolutionary tendency' – cannot be avoided. The need for a cooperation with other sciences is thus asserted by Löwe; yet, he considers a preliminary understanding of 'the social process as a whole' to be of the greatest importance (Löwe 1936: 21; see also Löwe 1935). The attention is thus shifted from isolated economic actions, from the logical form of *ceteris paribus* individual choices, to the conditions of individual valuations, and therefore, primarily, to 'the structure of the whole order' (Löwe 1936: 24).

In his article of 1976 Kapp resumes and refines his methodological reflections, significantly beginning with a reference to Robert Lynd: social sciences must be able to analyse the social whole, and, on this basis, to raise the question of 'what functionally more useful kinds of order can be created' (Lynd 1939: 126). Kapp, in the conclusion of his article, alludes to 'a new normative approach to social analysis and social action' (Kapp 1976: 228).

In fact, the link between the institutional method and political commitment of social scientists is the leitmotiv of this article in particular, and of Kapp's work in general. On the one hand, 'institutional economics replaces the concept of the economic man by something which we may call the "institutional man"' (Kapp 1968a: 93). This means that the institutional and technological environment and its transformations are no longer considered as given, known and constant, indeed external and extraneous to the subject matter of economics (Kapp 1968b: 71-72). Furthermore, the economic process and its transformations, the data of individual choices and choices themselves are intelligible only by referring them to the characteristics and dynamics of the market-capitalist institutions. On the other hand, as the institutional outlook implies a social accounting whose items depend on historical socio-economic institutions, it implies as well the pursuit of alternative, more efficient institutional arrangements (see Heidenreich 2000: 98 and 120).

This attitude implies a theoretical paradigm, which is opposed to that of neoclassical economics. The latter, Kapp observes (Kapp 1976: 210), has adopted 'methodological individualism' and has developed 'into a self-contained body of knowledge'; besides, it has privileged the concept of equilibrium 'by analogy to mechanics', and has conceived the economic activity as 'a particular type or "form" of human conduct', and therefore its own task as the study of human behaviour as a relationship between ends and scarce means which have alternative uses (according to the famous formulation by Lionel Robbins). As the economic system tends to appear as closed and self-sustaining, so conventional economics claims autonomy and turns the way 'of a formal, self-contained, closed mechanical analytical system' (Kapp 1976: 211). Kapp stresses the inadequacy of this kind of 'habits of thought', by referring to some aspects of the crisis, which was then (middle Seventies) becoming apparent and is still enduring. Presently, Kapp says, 'as always in times of economic and social crisis',

> 'normal' economic theory (...) is criticized for its inability to provide an analytical framework for the diagnosis of the problems and the formulation

> of more adequate criteria, policies and remedies designed to cope with increasing internal and international disorganization, environmental disruption, stagnation and inflation as well as unemployment, conflicts over terms of trade, etc.
>
> (Kapp 1976: 212)

A different standpoint is therefore indispensable: that of the interaction of the economy 'with a more comprehensive social and political as well as physical system from which economic processes receive important organising (and disorganising) impulses and upon which they exert their own negative and positive influences'. By relying on this principle, 'institutional economics has always aimed at a coherent representation of economic processes within and as a part of a complex social system and their interaction' (Kapp 1976: 213).

Robbins himself, in *An 'Essay on the Nature & Significance of Economic Science'* (1932), maintains that conceptually distinguishing the realm of the economy from that of other aspects of individual and social life doesn't prevent us from putting them into relation; on the contrary, moral, social and political values would be appreciated as such, and as such would determine choices. Max Weber's distinction between 'formal' and 'material' economic rationality is interpreted by Robbins as a support to his thesis. However, from the institutional standpoint that is shared by Kapp and Polanyi, and corresponds to the sense of Weber's theoretical and methodological reflection, two objections to Robbins' position can be made.

First, Robbins' 'formal' definition of 'economy', characterized as it is by the postulates of the economizing attitude and scarcity, is biased by its adherence to a historically given economic organization, which he fails to define and analyse as such. Polanyi sees in that definition the core of the 'economistic fallacy', because, by conceiving the economy in terms of 'scarcity plus economizing' (Polanyi 1977: 20), it mistakes a given institutional arrangement of the economy for the economy in general, an element of the set for the set. The market system, the 'system of price-making markets' to which the organization of the economic activity is committed, is in fact an 'integrative exchange pattern unique to our times' (Polanyi 1971: 19).

In that specific historical situation – in which traditional norms disappear, needs cease to be culturally predetermined, and livelihood is no longer warranted to individuals within a network of personal links and reciprocal rights and duties – economic behaviour is institutionalized as 'economizing'. 'The principle of gain and profit' spreads 'as the organizing force in society', and masses of individuals must sell their work-force as the only means to avoid 'the penalty of starvation' (Polanyi 1957b: 170 and 164). Money becomes the universal medium, scarce by definition and requiring a choice among different uses. As a consequence, scarcity acquires a specific meaning as a factor of a given institutional system, the modern economic organization; neoclassical economics entangles that meaning of scarcity with the more general one, concerning the condition of human beings on earth. It is thus on the real, *institutional basis* of the market-capitalist society that mainstream economics builds its 'economistic' generalizations, from the 'formal'

definition of the economy, to the economically rational individual choice, methodically taken not only as a basic element of the market system, but as the starting and central point of the analysis of economic activity in any society.

Second, the want of an institutional analysis of the 'mode of production' prevents the formal-neoclassical approach from raising the problem of the actual differentiation and autonomy of economic activity in the market-capitalist society, and from understanding the causes, changing modalities and consequences of the peculiar 'place' the economy therefore occupies.

In this kind of society the economy follows rules of its own; it is – to quote Polanyi's famous concept – no longer 'embedded' within social institutions of diverse sorts: it is, so to say, 'economically' instituted. Economic behaviour stands out as such; everybody must 'economize'. Economizing is carried out for its own sake as profit, and as such it becomes the dominant and typical form (structure, Gestalt) of economic activity. No wonder economics has developed, on this basis, the tendency to consider the economic system as a closed system.

That tendency is one of the main objects of Kapp's criticism. In fact, not only does the economic system tend to be self-reflexive, that is, to select and elaborate information coming from its environment only for the sake of the reproduction of its own organization; it is precisely thanks to its autonomy that the economy exerts its constraints and influences upon the functioning and development of modern society. Opportunities and freedom, welfare and health, life and death of all individuals are thus affected. The 'form' of the economy – in the sense of its market-capitalist organization – heavily conditions the 'substance' (ends and modalities) of economic processes and social institutions, and therefore individual and social choices, as well as society's relationships with its human and natural environment. As Polanyi says, the autonomy of the economy not only represents a 'complete reversal' (Polanyi 1957b: 68) of the pre-modern 'embeddedness' of the economy in society, but the *market economy* – with its congenital capitalistic relations of productions – 'gave rise to yet another, even more extreme development, namely a whole society embedded in the mechanism of its own economy – *a market society*' (Polanyi 1977: 9).

The economic system cannot be considered as closed, all the more that this mystification comes out of the very reality of the market-capitalist organization, from its autonomy. Not only can we not take it for granted that the system of market prices responds in the most efficient way to individual and social needs, but we must acknowledge that its functioning changes its environment, individual needs included. The problem is therefore to oppose and to reverse this tendency, by *politically* recovering and organizing – hopefully in a democratic way – that freedom to know and to choose, which is taken for granted by Robbins and his 'formal' definition of the economy and 'the nature and significance of economic science', but is actually jeopardized in a market society. Nothing less than man's survival in a technologically developed society is at stake; institutional conditions directed towards transforming man from object into subject of the economy are required for that minimal achievement, besides becoming an end in themselves (see Kapp 1968b). Kapp's concern with planning, which spans his entire intellectual

production, meets that problem, that exigency. A few hints will be made below to the fact that the same concern characterizes in general institutional and 'critical' economic thinking, in the face of capitalist development and, in particular, of its crises and turning points.

Kapp opposes to the 'formal' concept of 'economy' the 'substantive' one, which he shares with Polanyi (1957a) and finds also in an earlier work by A. G. Gruchy (1947): the object of economics is, for the institutional approach, 'not a particular *form* of behaviour', but rather a 'set of interconnected dynamic problems which arise in the satisfaction of individual needs and public objectives' (Kapp 1976: 214). From the 'formal' point of view, 'the subject-matter of economics' are the relationships between ends and 'the technical and social environment' (Robbins 1962: 38). 'Ends as such' – Robbins continues – 'do not form part of this subject-matter. Nor does the technical and social environment'. Kapp's opinion is the opposite. First of all, the analysis of the diverse '*forms*' (shapes, structures) of social-economic organization is an issue of primary importance. In this sense he observes that our attention is to be focussed 'on the evolution of social systems and social processes' (Kapp 1976: 214). And the first step of such inevitably comparative analysis is the question: what are the main factors of that evolution in our epoch? Kapp quotes, in this regard, the first page of Veblen's *Theory of Business Enterprise*: the 'Capitalistic System' is characterized by 'the industrial system' as its 'material framework', and by 'investment for a profit', as pursued in particular by 'business enterprise', as its 'directing force' ('the principle of profit' – according to Polanyi's statement quoted above – 'as the organizing force in society'). As a consequence of this institutional theoretical basis, the ends, limited and conditioned as they are by the historically specific characteristics of the economic system, cannot but be included in the scope of the economic theory, as well as the consequences of the functioning of the economic system on its environment. This is a logical premise to the concept of 'social costs'.

It is on this ground – on the theory of the specific institutional characters of capitalism as a social form of economic organization – that, according to Kapp, institutional economists have been able to analyse not only economic development and fluctuations, but also such factors of the complex social-economic process as the 'legal foundations of capitalism' and 'vested interests', 'collective and political bargaining' and 'public utility regulations', 'the industrial-military complex' and the 'power elite'. Analyses of this kind support the 'critique of classical preconceptions' by institutional economists, and in particular their scepticism about market prices ('in terms of which business enterprise tends to measure its performance and efficiency') 'as criteria of optimal decision-making and as indicators of economic rationality in general'. It is not surprising – Kapp adds – that institutional economists have called attention to 'the social costs of production' long before the present evidence of 'serious threats to social reproduction and the quality of individual and social life' (Kapp 1976: 215).

Conventional economics cannot account for those problems; also the idea of 'planning for national development', as well as the 'domination effect' (François Perroux) or the 'center-periphery problem' cannot find any place within its 'paradigm'.

Kapp here makes (Kapp 1976: 217) reference to Thomas Kuhn's theory: generally speaking, a new paradigm, a new 'theoretical framework' is necessary, if the analysis of social-economic processes is to be consistent with 'the principle of interlocking circular interdependencies within a process of cumulative causation' (Myrdal 1957: 23; quoted by Kapp). Kapp himself recalls that this principle is central in Veblen's methodological reflection, in particular in the article in which he pleads for a shifting of paradigm that would make economics an 'evolutionary science' (Veblen 1898).

A 'new theoretical framework' (Kapp 1976: 223) of this type is clearly working, according to Kapp, in Myrdal's analysis of the process of underdevelopment, but has a general meaning, which can be illustrated in the following, interrelated points.

1) Elements and conditions of the process, as well as its effects, are not conceivable in strictly economic terms. In particular, if institutional, cultural and natural elements are to be considered, a 'trans-disciplinary' approach is required.
2) There exists 'a mutual responsiveness, i.e. a capacity of the different conditions to react upon changes of one or several elements'; the result is a 'circular and cumulative interaction which shapes the dynamics of the system' (Kapp 1976: 220). Any sort of balance between forces and countervailing changes, and therefore the mechanistic notion of 'stable equilibrium as a paradigm for problem solving in the social sciences' (Kapp 1976: 223), is thus refuted. This is even obvious if we want to understand, for instance, such 'self reinforcing movements' as inequalities of income, 'disparities and polarization', the differentiation 'between "center" and "periphery"', and 'emerging scarcities of non-renewable resources' (Kapp 1976: 223 and 221).
3) Economic processes are to be represented 'as radically open systems which exchange energy and matter with the environment'; in the course of this exchange, 'qualitative changes take place both with respect to the environment and the process itself'. Besides, this question is to be envisaged with reference to an appropriately long run (Kapp 1976: 222 and 223).
4) The method thus sketched enables us 'to transform problematical situations and unsolved open problems (...) into "puzzles" which can be solved', through 'the specification of possible measures and priorities' (Kapp 1976: 220 and 221).

It is clear here that the 'market' is to be considered as an institution, that is, as a social, historical form of economic organization; therefore, it is to be analysed in its concreteness as a social-economic process, on the basis of that general concept and the dynamics of capitalist accumulation. The market's ongoing reality therefore appears as the result of processes of 'cumulative causation', which are not pre-determined, but are biased by the constraint of 'the principle of profit', and are complex and changeable, though largely irreversible. As a consequence, for the institutional outlook, as opposed to the conventional one, *data* of economic choices cannot be taken tautologically in their given, instantaneous existence,

but must be genetically explained as the result of those processes. Choices, then, are always *free within constraints*. The ongoing economic system, while imposing its own constraints, tends to ignore and conceal constraints concerning its human and natural environment, as well as many of the consequences of its functioning on its environment: choices are correspondingly limited and biased. In order to understand both the evolution of the economic system in terms of a social process, and the mutual, cumulative consequences of its interchange with its environment, in order to cope with problems of irreversibility and 'communication', the point of view of thermodynamics and cybernetics is to be adopted, instead of that of mechanics. The theory of social costs is meaningful within Kapp's purpose to plead the awareness and augment the knowledge about these problems, as a means to improve the freedom of choice, and therefore welfare. Indeed, not only the welfare of human beings, but even their survival are at risk, if we consider the destructive potentialities of technical advancements in our time.

The obvious starting point is that the optimality of resources allocation is not granted by the price system. This can appear as an understatement, but leads us to raise an important, general question: in what sense and to what extent the market-capitalist economy is economic, how does it perform the economic function? This is a typical 'substantive' question, which presupposes the distinction between a general definition of 'economy' and 'economic', and 'economizing' as the 'pecuniary' form of the economic activity (as well as between 'means of production' and 'capital' etc.). Institutional thinkers are thus led to question the efficiency of capitalism, and to assert the need for an informed planned intervention and for an inquiry about its possible purposes and ways.

The institutional and 'normative' approach Kapp expounds in the article of 1976 reflects his lifelong work. Besides, this article is linked to his field experience in economic policies for development as well as to the crisis of the post-war model of growth and of the related hopes about a stable and universal development. In this situation, the question both Veblen and Weber raise, concerning social orientation of the economic system and its interaction with its environment, becomes manifest and urgent. Also systems and communication theories, and cybernetics, have supplied a way of conceiving that question and facing it, theoretically at least. Some of the above quoted statements by Kapp reveal that he was not unaware of such new methodological opportunities, suitable as they are for a clear and consistent critical analysis of the disadaptive, harmful and potentially catastrophic effects of the economic process, when it pretends to be, and is conceived as, an autonomous system. Indeed, Kapp does not fail to declare that his 'open-system approach' to the complexity of economic problems is perfectly consistent with 'systems analysis' (Kapp 1985: 150-51).

The problem of planning can thus be formulated as the need for a suitable organizational intervention, in order to provide the economic system with the required capacity of perceiving and processing information coming from its environment. This way, the tendency to disorder, to entropy, from the point of view of human society as a whole and of the larger eco-system to which it belongs, could be contrasted. This requires the contribution of diverse social and natural sciences

as well as the provision of appropriate organizational support. While the analysis of social costs intends precisely to supply indispensable information, the condition of a suitable organizational alternative to the market is of primary importance, and most difficult, since it involves politics and, as Veblen would say, vested interests and habits of thought.

As for the relevance of systems analysis, which the circumstances of the crisis of the post-war phase of development made more evident in the field of social sciences, it is possible here only to allude to a few examples. The very concept of system, Alessandro Pizzorno observes, implies the problem of defining its purpose and controlling possible gaps between its achievements and that purpose; systems, then, are always to be conceived as inserted within systems of higher order. 'Normal' economics, on the contrary, conceives the economic system as self-referential, in the sense that its purpose is reduced to its internal equilibrium, to a very elementary regulation, to a self-reflexive 'economizing' (Pizzorno 1973: 180; see also: Georgescu-Roegen 1971; Wilden 1972, 1973 and 1977; Morin 1980; Luhmann 1986).

Further developments of a systemic, communicational and ecological reflection about the economy, which can be considered as close to Kapp's approach, continue up to the present. Within an autonomous, self-reflexive economic system, the problem is that of a systemic 'economizing', which implies an endeavour to reduce costs accounted for in private enterprises' outlays, also by *systematically* shifting them – this is the point of Kapp's concept of 'social costs' – onto both human and natural environments. This is the reason for the inadequacy of any attempt by conventional economic thought to face the problem of social costs (reductively called 'externalities') through *ad hoc* measures, such as taxation, protection of 'property rights' of the victims, negotiations on 'pollution permits' etc. Confidence in such measures presupposes, once more, an economistic and mechanical model, within which equilibrium can be restored by corrective interventions concerning *random* deviations. Kapp maintains that social costs are instead 'normal and typical phenomena' within the market-capitalist system, and therefore the neglect of them cannot be understood, as 'some analysts' do, as 'the result of a lack of information' or of 'an inadequate system of adjudication of damages caused and sustained' (Kapp 1969: 334-35). More generally, Kapp observes, the expedients those analysts suggest amount to 'a return to the conventional wisdom of relying on property rights and market costs and returns instead of preserving the principle of maintaining communal rights to nature and treating them as social use-values serving fundamental human requirements' (Kapp 1985: 158).

Systems theory has been important in supporting and diffusing a critical analysis of that kind of, so to say, neoclassical-neoinstitutional amendments. Olivier Godard, for example, looks for a possible overcoming of 'economic reductionism', and for institutions capable of defining the 'collective frame' of the economy. He points out the inadequacy of neo-classical economics for the analysis of the autonomy of the economic system and the very concept of environment; he criticizes, in particular, the attempts at internalizing environmental externalities into the market order (Godard 1984; see also Martinez-Alier 1987; for a recent discussion about

diverse ways of 'reinstituting the economic process' and diverse kinds of 'environmental economics', see Adaman *et al*. 2003). Finally, an explicit reference to Polanyi and an implicit one to Kapp can be found in James O'Connor's theory of the 'second contradiction of capitalism'. Economic growth has been possible, according to O'Connor, at the condition of depleting and damaging human and natural environment, without taking these costs into account. This, in the long run, provokes a rise of costs within the economic system itself and therefore the risk of a runaway process (O'Connor 1991).

Market efficiency

Both the origin of the debate about planning and the reasons that have made it endless are to be found in the class clash that marked the crisis of liberal capitalism, culminated in the period of World War I and continued over the greater part of the 20th century. As soon as the war was over, in the face of general control of the economic system by the state, what Luigi Einaudi (1933) called 'war collectivism', liberal thinkers such as Einaudi himself or Ludwig Mises, urged a return to the 'old liberal creed' and the rejection of the socialist alternative as a new form of the 'statist' tendency, of 'an authoritarian idea of the state' (Mises 1983: 5–6). Noble-minded, old-style liberals of that kind were still confident in the *institutional feasibility* of a true free-market system. They acknowledged the new reality of big business, of the interlacing of industrial and financial interests and of their ability to direct state policies to their own advantage, but considered all that as a vestige of pre-modern times or a degeneration to be amended. It is interesting to note that an influence of this point of view is traceable in Joseph Schumpeter's 'essay on imperialism' (1919).

At the opposite side of the barricade we find such scholars as John Hobson and Otto Neurath, Veblen and the American institutional economists, as well as socialists of diverse sorts. Max Weber, too, questions acutely and radically the market system as such, beside analyzing the historical making of its institutions. Roughly simplifying, the following points were opposed to liberal positions.

1) The market system tends unavoidably and irreversibly to transform its institutions toward growing capitalist concentration, conscious organization and political interventionism.
2) That system – not only on account of its departing from the perfect competition model, but also as such – does not guarantee either freedom and social justice or efficiency.
3) Other forms of economic organization are conceivable, in which state intervention and even the abolition or at least a social control of private property of the means of production would not undermine, but, on the contrary, would improve both economic efficiency and democracy, thus freedom.

Furthermore, there existed a diffuse confidence in the self-governing ability of men and women, if only they were free of disposing of their own destiny. On this

base different projects of an 'industrial democracy' were built, as the presupposition for the 'realization' of democracy, i.e. of an informed and responsible participation of the individuals to decisions concerning their life. Democracy should not be limited to 'formal' political rights, but should pervade everyday life and should be extended to every aspect of the social process, first of all to the economy. Social relations would thus become transparent, and freedom promised by the bourgeois revolution would be achieved and extended to everybody.

Polanyi was deeply influenced by this kind of political and intellectual atmosphere in the 'Red Vienna', where, after World War I, it was particularly lively. At that time it seemed possible to extend the recently conquered political democracy to the economy: the overcoming of the crisis of liberal capitalism, and in particular of what Polanyi later called the 'utopia' of a self-regulating market, would thus lead to a real, not illusory freedom (Polanyi 1924, 1925).

In Vienna, Polanyi participated in the debate on the possible forms of socialism. In his 1922 essay about this subject – in which he outlines the traits of a non-centralized socialist organization, in opposition to Mises' refutation of its very feasibility – the problem of the objectives pursued by the economic system is central. Economic production and exchange should be 'framed' by a social, conscious and democratic process of choice concerning goals and constraints.

The starting point of Polanyi's essay is that in a market economy – that is, a capitalist economy more or less characterized by big business – the economic system tends to be self-referential. Since choices, and the very selection of information, are determined or biased by 'the principle of gain and profit', 'economic' efficiency cannot immediately and in general be considered as coinciding with efficiency from the point of view of society. The capitalist economy, Polanyi writes, 'by nature' cannot be guided towards 'social utility'; indeed it lacks 'the sense organ' in order to perceive social needs and evaluations. Not only that, but the 'retroactive effect' production has on the community and on individual lives and needs cannot be systematically taken into account and even understood as a problem (Polanyi 1922).

It would be easy to show how much the idea of the inefficiency of the market system was widespread in those times. We could, for this purpose, refer to Veblen and the American institutional economics, or to the 'economic heretic' (as he called himself) John Hobson. The latter, as well as Veblen, pays attention to the institutional transformations, in particular to the control that 'strong business organizations' have acquired over government, with a view to turning internal and external policies to their own advantage, thus compromising the interest – and, of course, the freedom – of the greatest part of the citizens and of the nation as a whole. The overcoming of this class supremacy appears to Hobson as the condition for establishing an effectively democratic government, founded on the development of 'intelligent co-operation' with a view to 'clearly defined ends' (Hobson 1919: 75, 87 and 143).

In Veblen's works, as is well known, but also in Weber's *Economy and Society*, the inefficiency, indeed the 'irrationality' of the market system is stressed. As Veblen analyses the opposition between 'industrial and pecuniary employments', so Weber

maintains that 'rational' capital investment, oriented as it is to 'profitability', biases the choices about objects, technology and organization of production in such a way, that, in general, consumers cease to be 'sovereign'. Furthermore, not only the choices, but even the needs of the latter are purposefully influenced (Weber 1980: 88-90). Kapp would say, as we will see better below, that the market-capitalist co-ordination of different plans gives producers the opportunity of planning also on account of consumers, whose freedom is thus jeopardized.

In an early article, in which different meanings and forms of planning and 'economic regulation' are illustrated, Kapp explains that the 'free market', considered in its reality and not as an abstract model, 'at best is only inadequately equipped to measure social losses and to supply the means for the satisfaction of social ends and objectives'; indeed, it leads to a utilization of resources that can be 'wasteful' and even 'disastrous' (Kapp 1939: 772-73). This article marks the continuation of Kapp's concern about planning, which has a central relevance in the development of his thought.

That concern constitutes a starting point, as Kapp's dissertation of 1936 shows. He cites there the collection of essays on planning edited by Friedrich Hayek the year before (Hayek 1935), which summarizes and continues the above-mentioned Vienna debate. The great crisis and the institutional transformation of capitalism is a second fundamental issue in Kapp's reflection; he significantly quotes, in this regard, the book by Adolf Berle and G. C. Means of 1933, *The Modern Corporation and Private Property.* Present economic systems, Kapp observes (1936: 15), are generally very far from the model of liberal capitalism ('*liberal-kapitalistisch Idealtypus*'). Already before the slump, self-regulation of the economy through a free price formation was nothing more than a fiction. The concentration of capital and production, industrial cartels, financial links among companies (through their stockholders), entrepreneurial and workers' associations, and collective bargaining, reveal and enhance the tendency toward a restriction of competition in the supply-side of the economy. Big business holds a monopolistic power on the market; small firms depend on the price policies of larger ones (Kapp 1936: 20).

Nobody could thus deny that 'the market has lost its importance as regulator of the capitalist economy' (Kapp 1936: 21). The question is, Kapp continues, if this depends on the internal development of the free-market capitalist economy, or on external conditions. The entrepreneurial associations could grow up, in fact, thanks to state intervention, and the control of prices would be difficult without customs barriers. But these policies have not been decided and implemented by a 'neutral' state. Economic power has always meant political power too. Entrepreneurial associations themselves have urged governments to take protectionist measures. The cause of the above-described developments of the economy, Kapp concludes, is internal rather than external.

Some liberal economists – Einaudi, for example – also shared the view that capitalists were the people chiefly responsible for the decay of the free-market system; however, they didn't give up their confidence in the possibility of recovering it. Kapp's position seems instead closer to Polanyi's: the free market has been more a utopia than a historically feasible institution; anyway, in the course of capitalist

development, it has become irreversibly obsolete, even as a normative model. Kapp, too, rejects the thesis of an 'anti-liberal conspiracy', as Polanyi calls it: namely, the idea that a perfectly competitive and self-regulated market could have worked, and could still be restored, if only private and corporate interests would cease to jeopardize it. Kapp's outlook is fully institutional also in this regard, in the sense that his explanation of the process of development and transformation of the capitalist economy (and society) is grounded on the general features of that kind of economic and social organization. Only a perfectly competitive market would be effective in regulating the economy in sight of the optimum satisfaction of needs: but capitalists systematically endeavor to avoid it, because it would be 'ruinous' for them. Kapp mentions in this connection the problem of power and several economists who dealt with it. His opinion is similar to Weber's, that the market is *constituted* by struggles for power and asymmetrical relations. This is a good instance of the institutional vision of economic reality as a complex and cumulative process, rather than in terms of deviations from a model, from 'equilibrium'.

Kapp considers the crisis as having been fuelled by the above-described developments. The fact that a return to a free-market economy appears as utopian, suggests that the capitalist development itself has determined the conditions for a new planned organization of the economy (Kapp 1936: 24). This not only alludes to a possible 'structural order' (Löwe) which would be different from that of capitalism: it also means, according to Kapp, that even 'the maintaining of the discretional power on the means of production in private hands' may require 'a stronger restriction of the free enterprise and the autonomy of individual entrepreneurs' (Kapp 1936: 22). Anyway, Kapp continues, state intervention in different branches of the economy will never go against the interests of the entrepreneurs concerned, or at least of the greater and more powerful part of them. Polanyi's opinion can be recalled in this connection, that a firm hold of power by the ruling class was a precondition for the institutional transformations, which in the Thirties were aimed at facing the crisis of market society and the *impasse* it had reached.

The analysis of the developments that the market-capitalist system, pushed by its internal drives, has undergone, and the crisis of the market as 'regulator' of the economy, suggest the question Kapp doesn't fail to raise, concerning economic organization in general and of its diverse forms. What is, indeed, 'the essence of the *economy*'? In order to understand it, Kapp continues, we have to oppose it, as Weber does at the beginning of his *Economy and Society*, to 'the essence of the *technique*', which is to achieve a given result with the minimum cost. The *economic principle* concerns, instead, a choice among many possible uses of available means, in order to produce the most valuable products. Whether a given production is 'economic' depends, then, on 'factors going beyond the ambit of individual productive processes' (Kapp 1936: 27 and 28). The key questions concern *evaluation:* how is it performed, through which institutions, and how is it implemented? Kapp as well as 'classical' institutional thinkers criticize market-capitalist institutions as to the form (organization and purpose) and the results of the evaluation; they are thus led to an inquiry about more efficient alternative institutions and the very concept of 'social value'.

Löwe's above-quoted article begins with similar considerations about the 'economic principle', which consists, he says, in the arrangement of data according to their 'subjective importance' (Löwe 1936: 19). Economic theory, as a formal theory of choice, can then proceed deductively. Löwe however raises the problem of data itself, of its 'structural order and evolutionary tendencies'; the problem of the 'concrete order of ends and means' as a result of 'the social process as a whole' (Löwe 1936: 19 and 21). In what conditions do valuations and choices take place, and how are they determined?

In a planned economy, according to Kapp, there would be a conscious coordination of productive processes. Within the market system, instead, the economic principle is implemented by the market, which is supposed to allow entrepreneurs' rational choices and to connect them with consumers' evaluations, through the mechanism of prices (Kapp 1936: 30 and 34).

At this point, however, the real social structure and the evolution of the market-capitalist system are to be considered. Since the purpose of any capitalist entrepreneur, Kapp observes, is to minimize costs in comparison with prices of the goods he makes, only the actual existence of both perfect competition and absolutely sovereign consumers would make the productive choices by the entrepreneurs coincide with an optimum allocation of resources. But, if the market is to be considered in its reality, and not as an abstract utopia, the lack of free competition, and therefore of a free price formation, is manifest; besides, consumers' choices are limited and conditioned (Kapp 1936: 35).

These considerations lead Kapp to refute the 'preconception' of the efficacy of the would-be free market; a critical analysis is necessary about the results and efficiency of the way in which disposable resources are employed and the economic principle is actually implemented. Not only is the market far from being perfect, but its changing conditions have to become the object of an 'evolutionary' analysis. Kapp shares Veblen's method when he asserts the need to take into account 'social and political opposing forces' (Kapp 1936: 36) and processes of 'cumulative causation'. The resulting historical process is neither determined nor casual, but constrained within the general features of a historically given form of socioeconomic organization, and meaningful in relation to them. The 'evaluations' and the 'data' – that is to say, the elements of economic choices, the conditions to which the 'economic principle' applies – are correspondingly biased.

On the basis of the preceding considerations, Kapp's criticism of market-equilibrium theories can be both theoretically motivated and empirically grounded. We must acknowledge, he observes, that in the existing economic conditions and social relations the process of price formation is distorted, because individuals' needs can have an influence (if any) only to the extent that they are supported by purchasing power. This distortion is systematic, since differences of income are again and again reproduced by the 'economic organization' (Kapp 1936: 37). As a result, in Carl Menger's words (1923: 49), 'the lightest desire' might be satisfied, while 'the most urgent need' would not be taken into account and even acknowledged; besides, the perception of needs can be imaginary, not developed,

and morbid. Kapp adds in his turn that the knowledge of the utility of different goods is often not only limited, but also influenced from outside.

It is interesting that Kapp, in dealing with this aspect of what he calls a 'critique of the economic calculation through market prices', makes reference both to Menger's posthumous book and to Sidney and Beatrice Webb's study on *Soviet Communism* (1935) – two works that, at first glance, it seems difficult to compare. The latter had just been published, and obviously was of topical interest in the times of the great crisis and the 'great transformation', when both the fascist assault and the communist challenge were growing harder and harder, and the definitive decline of the institutions of liberal capitalism was demanding a search for a new economic and social order. As for Menger's book, it was, according to Polanyi (1971), the result of a half century-long attempt to distinguish the price theory, and the very concept of 'economy', that Menger himself had so successfully proposed in the first edition (1871) of his *Grundsätze*, from a more comprehensive definition of economic activity. Polanyi interprets Menger's distinction between two 'basic directions' of the economy – 'the economizing direction stemming from the insufficiency of means' and 'the "techno-economic"' (Polanyi 1971: 18) – as premonitory of his own distinction between the 'formal' and 'substantive' definitions of the economy (see Polanyi 1957a). As we have seen above, only the latter (the economy as the social – 'instituted' – activity of supplying material means for human needs) is actually general.

We can understand Polanyi's tenacity in tracing the distinction between the two 'definitions' or 'directions' of the economy in one of the founders of neoclassical economics, only if we realize that the opposite paradigm, the institutional one, is grounded on that distinction. An institutional and critical approach cannot but stem from 'the "distinctive determination" of modes of production' Menger was looking for; and this, in its turn, logically presupposes 'a distinction between the economy as the sphere of man's livelihood', and the different forms in which the economy has been 'integrated' (organized institutionalized) (Polanyi 1971: 22). This method of institutional or – as Kapp also says, and the reference to Veblen is not casual – evolutionary economics supports the analysis of the actual functioning and evolution of the market-capitalist economy. Furthermore, it makes it possible to raise the problem of the *efficiency* of this kind of economic system from a societal point of view. This problem, which is typical of Kapp's approach, presupposes not only the above-mentioned analysis, but also, more in general, a *substantive* definition of the economy, since its general form is the following: how does a given institutional asset perform the economic function? How and to what extent are individual evaluations actually possible, and transformed into relevant information, within a given organization of the economic process? What are the conditions to which the 'economic principle' applies and how is it realized?

The market apparently, *formally*, performs the allocation of the means of production according to individual evaluations of producers and consumers. But, as we have seen, demand is taken into account only to the extent that it is endowed

with purchasing power (that it is 'effectual', as Adam Smith says). Furthermore, Kapp observes, the economic system tends to overlook the relative importance of the goods for satisfying needs, as well as the tentative of achieving the highest welfare, and even the possibility that productive processes or consumption might be harmful. Capitalists try rather to influence consumers' evaluations of the relative utility of goods by all kinds of means. In the end, from a market- and business-like point of view, a production is 'economic' when it gives rise in whatever way to a (solvent) 'need'. Therefore – this is Kapp's conclusion – the economic calculation through individual evaluations 'realizes the economic principle only *formally*' (Kapp 1936: 38–40).

Furthermore, the decisions private entrepreneurs take on the basis of the market system must be 'economic' in the sense of profitable for them; but this does not imply the absence of 'inconveniences and damages' for society. Kapp supports this early formulation of the issue of 'social costs' with a reference to *The Economics of Welfare* by A. C. Pigou. Beside such problems as professional diseases, the effects on public health of air pollution, and the increase of criminality as a consequence of social uneasiness, Kapp mentions the costs and sufferings, which affect the whole of society, but come from diplomatic and military support to the interest of a few foreign investors. A reference to John Hobson's *Imperialism* about the latter question seems unavoidable, though Kapp doesn't make it but includes that question among those dealt with by Pigou (Kapp 1936: 41).

Other costs falling on society come from restrictions to the diffusion of inventions and technical advancements, and from advertising. Furthermore, Kapp asks, in what sense and to what extent is an unlimited private exploitation of natural resources 'economic from the point of view of society'? It is impossible to answer this question on the ground of the price formation in the market. This is even more true as to 'social needs and interests', which simply cannot be accounted for, to the extent that they do not find any expression through individual demand and supply (Kapp 1936: 42).

Furthermore, in a market economy, Kapp finally observes, the state tries to face social problems through economic policies. Legislative interventions, social policies and charitable measures try to face the inadequacy, to say the less, of private evaluations and choices in a market situation. Only this kind of *ex-post* correction is admissible within a capitalist economic organization. Economic planning would allow, on the contrary, to take *a priori* and systematically into account any possible disadvantage that the economic process could cause to society (Kapp 1936: 43).

Kapp goes back to the post-war debate on planning and '*Sozialisierung*'. A planned economy is characterized by the fact that a conscious coordination of productive processes is supposed to take the place of the market; but is planning able to make the economic principle work? Mises gives a negative answer to this question; in particular, according to him, a comparative evaluation of disposable resources would be impossible, owing to the want of a unitary yardstick (Kapp 1936: 29 and 30).

Different solutions have been proposed, in order to refute Mises' criticism: for instance, Emil Lederer relies on the evaluation of means of production by

their cost, other scholars on the persistence of a modified form of market also in a planned economy or on a 'guild-socialist' organization. In fact, the problem of economic accountancy in Mises' sense, Kapp observes, would not exist if we were to cling to the classical 'objective' concept of value, to time of labour as the measure of value (Kapp 1936: 33).

Kapp ends his survey on planning with a few general considerations on man and society. If man is to be considered as a social being – and not as a single individual, as is generally the case in economics – he has to protect society's existence (from epidemics, external attacks etc.). Furthermore, as a member of society, he is interested in those institutions that favor public health and education, and general welfare. If people conceive of themselves as social beings, they can objectively recognize social needs as needs of all individuals. Exchange values, in a market-capitalist system, appear to be unsuited to this task, since they express producers' and consumers' evaluations of their interest as individuals, rather than social needs (Kapp 1936: 44). In a planned economy the allocation of resources would take into account social rather than individual evaluations: 'the economic principle would be considered from society's point of view' (Kapp 1936: 45).

It would be easy to compare this kind of arguments, made by Kapp, to those of 'classical' institutional economists. In *The Social Costs of Private Enterprise*, Kapp recalls the fundamental contributions by 'Veblen and his followers' not only on social costs, but also in providing a 'critique of the basic preconceptions of economic science' (Kapp 1950a: 41). Suffice it to refer, in this regard, to Veblen's distinction between 'industrial and pecuniary employments', and to quote a statement from his article on that subject. Gains, Veblen maintains, do derive from vendibility, but 'vendibility need not, even approximately, coincide with serviceability, except serviceability be construed in terms of marginal utility or some related conceptions, in which case the outcome is a tautology' (Veblen 1901: 309). Both Veblen and Kapp want to go beyond a formal-tautological concept of the 'economic principle', and to question from a 'substantive' point of view the economic system's 'serviceability', that is its capacity of realizing both the welfare of all the individuals and a successful long-term adaptation of human society to its natural environment. Both allude to the resulting necessity to lay down an economic calculation, which would not be constrained within the market mechanism and the 'business' outlook. As Kapp will say many years later, if we critically analyse 'what the conventional theory treats as given' – from asymmetrical power and information, to 'the institutional *status quo*' which constrains social and individual opportunities – 'market costs and prices' appear as 'unreliable measures of economic rationality', and social costs as 'normal and typical phenomena', not as exceptions to the rule (Kapp 1969: 335 and 338).

An analysis of the actual efficiency and effects of economic activity, and an open discussion about its ends have always been the purpose of the institutional approach – marked by its, so to say, 'substantive' paradigm – since the times of the parallel crises of liberal (19th century, free-market, Victorian etc.) capitalism and of classical economic thought. Institutional transformations stressed the utopian character of the utilitarian shortcut. Utilitarianism tried to ground on the

natural order an immediate and perfect correspondence between individual 'economizing' behaviour on the one hand, and actual individual and social welfare on the other, between the 'pecuniary' form of the economic activity and its 'serviceability', between market values and social values.

From Veblen onwards, institutional economists (and multifarious social and socialist thinkers) have tried to draw economic choice out of the 'economizing' attitude and calculation, and to base it on knowledge about socioeconomic institutions, conscious plans and 'social engineering', democratic diffusion of information and responsibility. Also Weber's distinction between 'formal' and 'material' (or 'substantive') economic rationality alludes to the impossibility of keeping economic questions outside the realm of social processes, cultural values, and – last but not least – politics. In the just-quoted article of 1969 Kapp observes that

> the concept of social rationality is an elusive one which cannot be formulated simply in terms of market prices because many of the relevant social costs and social benefits have no market value and, even if they did or could be assigned such value for purposes of 'objective' cost-benefit calculations, we would still be faced with the problem of their *political evaluation*.
>
> (Kapp 1969: 346)

Hence the need for considering 'social choices, social evaluation and social value' as a 'central problem of economic theory' (Kapp 1950a: 255–56). Kapp quotes on this subject J. M. Clark, who maintains that the concept of economic value should be developed 'with reference to society as a whole, independent of market valuations and capable of scientific application to concrete cases'; such 'intellectual instrument' would 'pierce the insulation' of economic theory 'and establish a connection with the ideas that are making things happen' (Clark 1936: 53).

The issue of social value further clarifies the opposition between the two paradigms (neoclassical and institutional). Kapp observes that the principle of marginal valuation can apply also 'to social valuation and collective choice' (Kapp 1950a: 259); concepts and methods of neoclassical economics can be employed also for that purpose. What makes the difference between the two paradigms is in fact the very definition of the economy and, accordingly, of the scope of the economic science. The 'traditional approach' is 'based upon a "methodological subjectivism"' (Kapp 1985: 153) and the economic system appears as closed; 'the resulting perspective and the criteria of evaluation are those in terms of which the system of business enterprise tends to evaluate its performance' (Kapp 1936: 145–6). As we have seen above, the institutional approach considers the traditional one as an uncritical generalization of the market-capitalist habits of thought, and the alleged efficiency of the market as correspondingly biased.

The article from which the preceding quotations are drawn was published in 1976; it represents, as well as the article examined above (par. 1), Kapp's final reflection on 'economics as a normative science'. There is a development of his thought, but also a fundamental continuity. His first publication has been considered from this perspective in the present paragraph; it would be easy to do the

same for his 1950 book, which closes with the wish that 'by including social costs, social returns and social values within the range of its analysis, economic science would become "political economy" in a deeper and broader sense than even the classical economists conceived of the term'. (Kapp 1950a: 261–62)

In reality, the 'conventional' tendency, as opposed to the institutional one, historically prevailed, and is still the norm. That tendency attempts to make economic science not only more abstract and formal and therefore independent from institutional forms and transformations of the economy, but also not interested in them and in 'substantive' issues, as if they were external to its scope, extraneous to its subject-matter (see Cangiani 2000). According to Schumpeter (1954), 'economic analysis' was finally clearing its way and distinguishing itself from 'political economy'; according to Thomas Balogh (1982) economics was condemning itself to 'irrelevance'. Irrelevance, and tautology, we could add, as to the necessity of questioning the actual 'serviceability' of the economic activity: but not as to supporting the reproduction of the ongoing institutional order.

Planning and freedom

The institutional perspective is immediately, closely and meaningfully linked by Kapp to the problem of planning. That link is traceable, as we have seen, in his thesis of 1936: it is even more apparent in an article of 1950, 'Economic Planning and Freedom'. The market system is here represented as an 'institutional arrangement' – to employ Karl Polanyi's concept – that is, as a social, historically specific way of organizing ('integrating') individual economic activities. The market, Kapp writes, 'under conditions of private enterprise', can be understood as a mechanism providing for the 'coordination' of the economic plans of private individuals (owners of labour, consumers, producers) with each other and with 'the planning of public bodies' (Kapp 1950b: 34).

At a less abstract level than that of the general features of the market as a historical organization, its developments have to be analysed; then, different forms of market coordination (different 'institutional structures', Polanyi would say) appear to be characterized by different combinations of types and methods of planning (Kapp 1950b: 33). For instance, anti-depression policies, such as those adopted by the New Deal and those recommended by Keynesian 'new economics', do represent an important aspect of the new, post-liberal institutional structure; however, they 'do not suspend the market mechanism' as such, and have in this sense an 'essentially conservative character' (Kapp 1950b: 37).

Not only do the transformations of the market system call for the acknowledgement of the growing planned organization of the economy, but economic activity appears essentially as a planning activity. According to Kapp, planning, which implies 'valuation and choice', is to be considered in general as 'a prerequisite and an essential element of rational conduct', and indeed a typical feature of 'human action' (Kapp 1950b: 30). Besides, at a less abstract level of analysis, we must acknowledge that government planning, and therefore 'social ends', are always present, though in different ways and measure. As critics of the neo-classical

point of view had long since asserted, the problem therefore is not *if*, but *how* to plan. Both 'in a system of private enterprise' and 'in a socialist economy' there is 'the same need for a coordination of the planning of the individuals and government' (Kapp 1950b: 37). Capitalism and socialism, and the different forms in which they can be instituted, differ from each other because they adopt different kinds of planning, that is different ways of defining and coordinating the plans of individuals and private and public organizations.

Kapp appears to be here, once more, radically opposing neo-classical economics, in spite of the latter's stress on subjective choice: he considers institutions instead of individual behavior, and plans instead of price mechanisms. Furthermore, it becomes clear that both institutions and purposeful actions should be considered as to their efficacy for the attainment of individual *and social* ends. Industry, profit and economic growth must cease to be (and to be considered as) ends in themselves, and the economy must not be considered a closed system.

It is on the ground of the above sketched considerations, and of the institutional method they imply, that the problem of freedom is to be raised and discussed. This is what Kapp does.

A non-centralized socialist system, in which the plans of individual consumers and producers would be taken into account, could be feasible. Kapp makes reference to Oskar Lange's analyses on that subject, and dwells upon it in the last pages of his article. In a socialist system of this kind, the volume and the nature of investment would be the object of a political decision, instead of being determined by private investors, as in the capitalist system. According to Kapp, anyway, in both types of economic organization the problem is that of coordinating different plans with one another, therefore of *planning:* only the 'kind of planning' is different (Kapp 1950b: 39).

To reject any sort of planning, the very idea of planning, is a prejudice, just like the claim that 'the free and unregulated market economy is the only effective guarantee of personal freedom' (Kapp 1950b: 41). In fact, on the one hand, 'neither the capitalist market economy nor national economic planning offer an absolute guarantee against totalitarianism and the subordination of the individual under arbitrary and absolute power' (Kapp 1950b: 51). On the other hand, 'there are no elements in national economic planning or in political democracy that make the two incompatible with one another' (Kapp 1950b: 51). According to Kapp, the 'real question', that is 'which forms of social organization are destructive of personal freedom and which reforms are compatible with free institutions' (Kapp 1950b: 40), is to be dealt with by analyzing historical reality and existing institutions.

'Market economy', generally speaking, 'has been accompanied by the growth of freedom (whenever it gave rise to higher standards of living) and by a return to totalitarian controls (when it produced a state of affairs which large masses of people considered intolerable)'. Furthermore, Kapp continues, 'the market economy itself, despite or more correctly because of the absence of central planning, produces conditions which may put us on the road to serfdom' (Kapp 1950b: 40). This thesis is similar to what Polanyi maintains in *The Great Transformation*, and

is opposed to that Hayek maintains in his book of the same year – *The Road to Serfdom* (1944) – to which a reference is clearly made in Kapp's sentence.

We have seen that Kapp since 1936 contests, in keeping with the tradition of institutional economics, the very ability of the market to generate knowledge and to diffuse it through the price mechanism. As Bruce Caldwell (1997) has highlighted, Hayek stresses precisely that ability in his critique of socialism following on from that of Mises. In his article of 1950 Kapp refutes also Hayek's political argument, that market (capitalist) freedom constitutes the true ground and only guarantee of liberty and democracy. Indeed, Kapp turns upside down Hayek's argument, by questioning the basic institutions of capitalism. 'What are' – he asks (Kapp 1950b: 41–42) – 'the effects on the individual if the pursuit of monetary gain becomes the supreme and only goal of life?' The freedom and even the 'mental health' of the individual can be impaired by his 'isolation from his fellow men or from values which transcend his monetary goal', and by 'economic insecurity'.

These general characters of the market organization are related to the autonomy of the economic system, to its differentiation, which is, according to Max Weber, the leading aspect of the process of modern 'rationalization'.

The problem of the autonomy of the economy is a central one, if the contradictory existence of freedom in modern society is to be understood. The process by which the economy differentiates from other aspects and functions of the cultural whole and begins to follow rules of its own, is a fundamental aspect of the process of rationalization. As a consequence of the latter, individual activity is no longer culturally pre-determined, and societies are no longer compelled to keep to tradition. In a 'disenchanted' world the new degree of freedom so attained implies a conscious planning of the interplay between man and nature, and of social institutions themselves. Planning, as Kapp maintains, is in general typical of 'human action'; as Marx put it, it makes the difference between an architect and a bee. This becomes, so to say, more true, true at a higher level, in the modern epoch. Modern rationalization, which is accompanied by the disruption of the cultural totality of traditional societies, enhances to a higher degree the freedom of human beings, who become aware of their quality of subjects of their own history and of human history in general. The individuals conceive themselves as such, and society as problematic: not as a given (natural, i.e. divine) reality to be interpreted, but as a human and historical reality to be analysed and changed.

The autonomy of the economy, on the other hand, comes into existence and develops together with its market-capitalist organization: thus it limits and even undermines freedom by imposing the 'economic rule' over society. 'Economic' norms – specifically characterized as they are by that kind of social organization – function as constraints with regard to the relationships of the social system with its human and natural environment, *and therefore to the planning activity as an expression of human freedom*. As early as the beginning of the 20th century, in his essay on *the Protestant Ethic and the Spirit of Capitalism*, Max Weber pointed out a central aspect of the modern paradox of freedom: rational economic behavior is characterized by a congenital irrationality, to the extent that it becomes an end in itself instead of being aimed at individual happiness and social welfare.

Utilitarianism has been a tentative to cope with the novelty of the modern situation. The utilitarian solution – which entrusted the commonwealth to individual choices, the common good to the pursuit of individual interests – relied on natural propensities and the self-regulating market mechanism. Planning as the typical character of the economic activity was kept in the background, until the crisis of 19th century liberal capitalism revealed the utopian nature of utilitarianism and of its related confidence in 'general equilibrium'. Economic activity could thus appear, as Kapp writes, in its general sense of 'decision-making', that is to say essentially as planning; this connotation affects choices by individual consumers and private firms to begin with, but is inevitably to be extended to social, *political* processes of choice. Planning appears at this point as the manifestation and implementation of human freedom, and even more so in modern society.

Now, if we want to face the above-mentioned 'real question' – that of the actual relation between given 'forms of social organization' and freedom – the structure, functioning and transformations of the capitalist society are to be considered. What are the subjects and the aims of planning activities, and how are they coordinated? Kapp, for instance, invites us to wonder about 'the effects of monopoly and big business on effective liberty'. Consumers, far from being 'sovereign', are 'victimized' by big business; they have a 'lack of awareness', and therefore cannot but be 'inarticulate' (Kapp 1950b: 42). Also the freedom of small- and medium-sized producers, Kapp continues, is heavily restricted. And what about inequality and penury, that persist in spite of the growth of productivity and production? We could say, from a substantive point of view, that the economic system, which is supposed to solve the problem of scarcity, creates scarcity as well. If 'free markets and free competition' bring about 'such extreme inequalities in the distribution of income' as to prevent some people from maintaining 'a social minimum of nutrition, health and housing', while some other people dispose of 'the least essential luxuries', then not only 'individual freedom', but also 'all moral prerequisites of social life' are violated (Kapp 1950b: 43).

Besides, if we look at the capitalist organization of work, individual liberty tends to vanish and 'undemocratic attitudes are fostered'. On the whole, according to Kapp

> if we understand by democracy not merely a form of government but a way of life which is characterized by human relationships based upon voluntary cooperation and participation of all in the formulation of purposes and values, it is indeed to inquire how far these conditions are actually realized in business and the field of industrial relations.
>
> (Kapp 1950b: 42)

Carl Menger's general statement can here be recalled, that wage-earning workers cannot be considered as economic subjects, as for their productive activity at least, since they do not carry out economic choices. Max Weber comments that this exclusion constitutes a 'specific element of *substantive* irrationality' of the social system (Weber 1980: 138).

Karl Mannheim published in different editions before, during and after World War II a well-known book, which is fundamentally dedicated to the need of a (in a wide sense) planned 'reconstruction' of society. One of the starting points of his analysis recalls the above-quoted remark by Weber, while extending it to the general condition of 'man and society'. The individual freedom in a complex industrial society – Mannheim writes (1954: I, Ch. 2) – would demand a dramatic increase and diffusion of 'substantive rationality', defined as 'the capacity of acting intelligently in a given situation'; but this, besides being difficult in itself, does not match with the concentration in a few hands of the means of production, and therefore of knowledge and the power of making choices.

Dependence on 'the amoral forces of supply and demand', 'economic insecurity', 'social injustice', 'the increasing intolerable influence and oppression of vested interests', 'social and commercial manipulation of choice', 'subordination and indifference' in workers' relations to their employers, 'social conventions, patterns and authority', are, according to Kapp, all aspects of those 'forces which are the direct outgrowth of the market system', and by which 'individual freedom and human personality are menaced' (Kapp 1950b: 44–45). As a result, the real situation appears to be the opposite of that which is represented by classical liberal ideals and normally presupposed by economic theories: in fact, Kapp maintains, 'the individual is in danger of losing his capacity of acting according to his own desires and preferences and with it his self-esteem and self-evaluation' (Kapp 1950b: 45).

A similar nonconformist analysis is carried out by Löwe in the above-cited article. He illustrates the advantages of 'a planned market system with public ownership' in comparison with capitalism (Löwe 1936: 28). According to Löwe, the market-capitalist system, particularly 'under the technical conditions of large-scale capitalism', systematically imposes constraints, rigidities and cumulative distortions (e.g. the social results of the exchange process are predetermined by the initial market positions of the partners; the sunk costs of past private investment hamper innovation and oppose the interests of investors to those of consumers, besides involving cut-throat competition; the overhead costs of labour never appear in the entrepreneurs' calculations) (Löwe 1936: 32 and 34). Löwe explains that, instead, under public ownership, investment can be planned efficiently and based on rational accounting of social costs and social returns; moreover, 'most of the social problems which the autonomous process of capitalism solves incidentally are assigned to planned action' (Löwe 1936: 34). As paradoxical as this could appear, Löwe observes, planning as a form of economic organization warrants that fullness of freedom, of 'deliberate action' in economic matters, which is impossible within a capitalist system. Subjective decision and personal responsibility take the place of an objective mechanism; their scope is no longer limited by 'economic' (in reality 'structural') constraints. Only planning allows, in fact, the use of means 'to any political end whatsoever'; it is 'a merely technical instrument', and the question is – can be – entirely that of 'the political organization' (Löwe 1936: 33).

In short, Löwe accepts the issue raised by Robbins, but he turns the latter's solution upside down. Only a public control over the means of production, as a premise for conscious planning, would realize the absence of structural constraints

concerning the economy, and therefore a 'social neutrality of the economic mechanism'; thus 'the crucial tasks' could be transferred 'to the political and psychological sphere' (Löwe 1936: 35). Instead, the hypothetical individual rational choice in a perfect-competition market situation is not a plausible solution: choice is in fact bounded and biased, because it cannot but take place within a business-driven cumulative evolution of the economic process and of the market's setup.

Planning and public control over the means of production would thus allow – speaking in the abstract – a higher degree of freedom; while – in real terms – everything depends on how the 'political aims pursued' are expressed and implemented. Social processes of choice can in fact be performed by 'a democratic representation or by the dictatorial decision of a ruling minority' (Löwe 1936: 33).

Löwe's article seems to be meaningful for the range of problems and the political and intellectual atmosphere that are important in the development of Kapp's thought: roughly simplifying, the definitive crisis of the 19th century liberal system, the criticism of the utilitarian solution to the modern problem of the economy, the changeable fortune of the institutional method, and the debate about alternative forms of economic and social organization. During World War II that debate had a revival, to which, in particular, Kapp's article of 1950 is related.

References

Adaman, F., Devine, P. and Ozkaynak, B. (2003) 'Reinstituting the Economic Process: (Re)embedding the Economy in Society and Nature', *International Review of Sociology–Revue Internationale de Sociologie* 13, 2: 357–74.

Balogh, T. (1982) *The Irrelevance of Conventional Economics*, New York: Liveright Publishing Corporation.

Caldwell, B. (1997) 'Hayek and Socialism', *Journal of Economic Literature* 35: 1856–1890.

Cangiani, M. (2000) 'La scienza economica alla prova della Grande guerra', in V. Calì, G. Corni and G. Ferrandi (eds) *Gli intellettuali e la Grande guerra*, Bologna: Il Mulino.

Clark, J. M. (1936) 'Toward a Concept of Social Value', in J. M. Clark (ed.) *Preface to Social Economics*, New York: Rinehart & Co.

Einaudi, L. (1933) *La condotta economica e gli effetti sociali della guerra italiana*, Bari: Laterza, and New Haven: Yale University Press.

Georgescu-Roegen, N. (1971) *The Entropy Law and the Economic Process*, Cambridge, Massachusetts: Harvard University Press.

Godard, O. (1984) 'Autonomie socio-économique et externalization de l'environnement: la théorie néo-classique mise en perspective', *Economie appliquée* 37, 2: 315–345.

Gruchy, A. G. (1947) *Modern Economic Thought*, New York: Prentice Hall.

Hayek, F. (ed.) (1935) *Collectivist Economic Planning*, London: George Routledge & Sons.

——. (1944) *The Road to Serfdom*, London: Routledge.

Heidenreich, R. (1994) *Ökonomie und Institutionen. Eine Rekonstruktion des wirtschafts- und sozialwissenschaftlichen Werk von K. W. Kapp*, Frankfurt a. M.: Peter Lang.

——. (2000) 'L'economia come scienza della cultura. La fondazione costruttivistica dell'economia sociale nell'opera di Karl William Kapp', in P. Frigato, M. Giovagnoli (eds) *La Continuità della vita umana, 'Harmattan Italia, Torino,* 95–124.

Hobson, J. A. (1919) *Democracy after the War*, London: George Allen & Unwin; New York: Macmillan; a previous version was published in 1917.

Kapp, K. W. (1936) *Planwirtschaft und Aussenhandel*, Liège: H. Vaillant-carmanne, Imp. de l'Académie.

——. (1939) 'Economic Regulation and Economic Planning', *The American Economic Review* 29, 4: 760–773.

——. (1950a) *The Social Costs of Private Enterprise*, Cambridge, Mass.: Harvard University Press.

——. (1950b) 'Economic Planning and Freedom', *Weltwirtschaftliches Archiv* 64: 29–54.

——. (1963) 'Social Costs and Social Benefits – A Contribution to Normative Economics', in E. von Beckerath and H. Giersch (eds) *Probleme der normativen Ökonomik und der wirtschaftspolitischen Beratung*, Berlin: Duncker & Humblot.

——. (1968a) 'In Defense of Institutional Economics', *Swedish Journal of Economics* 70, 1: 1–18. Reprinted in W.G. Samuels (ed.) (1988) *Institutional Economics*, vol. I, Aldershot: Elgar, 1988.

——. (1968b) 'Nationalökonomie und rationaler Humanismus', *Kyklos* 21, 1: 1–25.

——. (1969) 'On the Nature and Significance of Social Costs', *Kyklos* 22, 2: 334–47.

——. (1976) 'The Nature and Significance of Institutional Economics', *Kyklos* 29, 2: 209–232.

——. (1985) 'The Open System Character of the Economy and its Implications', in J. E. Ullmann (ed.) *The Humanization of the Social Sciences*, Lanham: University Press of America.

——. (*n. d.*) 'American Institutionalism: The System of Economic Analysis of Veblen and His Followers', ms., Brooklyn College (*K. Polanyi Archive*, 26–6).

Karl Polanyi Archive Catalogue, 2001, Montréal: Karl Polanyi Institute of Political Economy.

Löwe, A. (1935) *Economics and Sociology*, London: George Allen & Unwin.

——. (1936) 'Economic Analysis and Social Structure', *Manchester School of Economics and Social Studies* 7: 18–37.

Luhmann, N. (1986) *Ökologische Kommunikation*, Opladen: Westdeutscher Verlag.

Lynd, R. S. (1967) *Knowledge for What? The Place of Science in American Culture*, Princeton, N. J.: Princeton University Press; first published in 1939.

Mannheim, K. (1954) *Man and Society in an Age of Reconstruction*, London: Routledge & Kegan Paul.

Martinez-Alier, J. (1987) *Ecological Economics. Energy, Environment and Society*, Oxford: Basil Blackwell.

Menger, C. (1923) *Grundsätze der Volkswirtschaftslehre*, 2nd ed., Wien-Leipzig: Hölder & Co.

Mises, L. (1983) *Nation, State, and Economy*, New York and London: New York University Press, (*Nation, Staat und Wirtschaft*, Wien, 1919).

Morin, E. (1980) 'Per una teoria della crisi', in M. D'Eramo (ed.) *La crisi del concetto di crisi*, Roma: Lerici.

Myrdal, G. (1957) *Economic Theory and Underdeveloped Regions*, London: Gerald Duckworth & Co.

O'Connor, J. (1991) 'The Second Contradiction of Capitalism: Causes and Consequences', *Capitalism Nature Socialism* 8.

Pizzorno, A. (1973) 'L'incompletezza dei sistemi', in F. Rositi (ed.) *Razionalità sociale e teorie dell'informazione*, Milano: Edizioni di Comunità.

Polanyi, K. (1922) 'Sozialistische Rechnungslegung', *Archiv für Sozialwissenschaft und Sozialpolitik* 49: 377–420.

——. (1924) 'Die funktionelle Theorie der Gesellschaft und das Problem der sozialistischen Rechnungslegung', *Archiv für Sozialwissenschaft und Sozialpolitik* 52: 218–28.
——. (1925) 'Neue Erwägungen zu unserer Theorie und Praxis', *Der Kampf* 18: 18–24.
——. (1957a) 'The Economy as Instituted Process', In K. Polanyi, C. M. Arensberg and H. W. Pearson (eds) *Trade and Market in the Early Empires. Economies in History and Theory*, Glencoe, 3: The Free Press.
——. (1957b) *The Great Transformation*, Boston: Beacon Press; first published in 1944.
——. (1971) 'Carl Menger's Two Meanings of 'Economic'', in G. Dalton (ed.) *Studies in Economic Anthropology*, American Anthropological Association.
——. (1977) *The Livelihood of Man*, in H. W. Pearson (ed.) New York: Academic Press.
Robbins, L. (1962) *An Essay on the Nature & Significance of Economic Science*, London: Macmillan; first published in 1932.
Schumpeter, J. A. (1919) 'Zur Soziologie der Imperialismen', *Archiv für Sozialwissenschaft und Sozialpolitik*: 1–39 and 275–310; English translation in *Imperialism and Social Classes* (1951) edited with an Introduction by P. M. Sweezy, New York: A. M. Kelley.
——. (1954) *History of Economic Analysis*, New York: Oxford University Press.
Veblen, T. (1898) 'Why is economics not an evolutionary science?', reprinted in T. Veblen (1994) *The Place of Science in Modern Civilization*, London: Routledge / Thoemmes Press, first published by Russell & Russell, New York, 1961.
——. (1901) 'Industrial and pecuniary employments', in T. Veblen (1994).
——. (1904) *The Theory of Business Enterprise*, New York: Charles Scribner's Sons.
Webb, B. and S. (1935) *Soviet Communism: A new civilization?*, London & New York.
Weber, M. (1980) *Economia e società*, vol. I, Milano: Edizioni di Comunità.
Wilden, A. (1972) *System and Structure. Essays in Communication and Exchange*, London: Tavistock Publications.
——. (1973) 'Ecosystems and Economic Systems', in M. Maruyama and A. M. Harkins (eds) *Cultures od the Future: Ninth Congress of Anthropological and Ethnological Sciences*, The Hague: Mouton World Anthropology Series.
——. (1977) 'Comunicazione', in *Enciclopedia Einaudi*, vol.VII, Torino: Einaudi.
Yonay, Y. P. (1998) *The Struggle over the Soul of Economics*, Princeton, N. J.: Princeton University Press.

2 Political democracy and social costs: reading K.W. Kapp's 'political economy' today

Regine Heidenreich

The notion of social costs reflects such basic concerns as worker safety, environmental protection and consumer protection, but also equity and social justice in society. Talking about social costs today means talking about the costs of labour or costs shifted towards future generations, i.e. the storage of atomic waste. Karl William Kapp, indeed, explored these social costs, and many more, in the USA of the 1950s. But we tend to forget that the idea of social costs has its genuine epistemological background. This chapter reviews the theoretical core of Kapp's concept of social costs and explores its interdisciplinary nature. This translates into a view of public policy providing insights into institutional reform in the German health sector.

Institutions and economic action

Kapp's notion of social costs is entrenched in his institutional economics. To avoid the fallacy of merely reproducing an isolated definition of social costs, this chapter develops the basic terms of his theory. In his unpublished papers, Kapp (1969) develops the following definition: 'Institutional Economics is not simply concerned with the economic and social impact of institutions and technology, but it must be understood as a form of economic analysis, guided by the perspective of a social structure in transformation'.

Favouring the term 'evolutionary economics' to 'institutional economics,' Kapp (1968a: 1) stresses the economy's built-in tendency towards dynamics and transformation with its implications like unpredictability and indeterminacy.

Institutions, as Kapp defines them (1969), are habitual patterns of thought and action, engendered by social arrangements. Economic evolution unfolds in relation to the institutional structure of society. In the evolutionary process, the adjustment of institutions alternates between ceremonial patterned behaviour and progressive institutional change. This evolutionary concept is a non-teleological scheme since the final form is open.

This point of view relates to time and space and thus joins together social and cultural features. Institutional economics emerges from the intersection of economic theory, cultural anthropology, and modernization theory. Kapp had adopted a realist methodology. He favoured a historical, anthropological approach, and an institutional

and context-interested analysis, over methodological individualism. Historical and institutional analysis is vital to cross-national policy research and can illuminate contemporary policy issues. It allows us to sort out short-term from long-term factors and raises sensitivity to how concepts are bound by particular cultures (Kleinman 1980, Payer 1988).

Biographical notes: dehumanization and economic thought

Kapp, like so many German intellectuals[1], left Germany because of the rise of fascism. In America he did not keep to the inner circle of German émigrés but drew key ideas to institutional analysis from modern American social sciences: modernization theory, modern anthropology, psychology, and gestalt theory.

In both a biographical and theoretical sense, Kapp is an important mediator between American and European institutionalism. A scholarship from the Institute of Social Research – the Frankfurt School of Critical Theory in exile, situated at the New School for Social Research – made it possible for Kapp to establish himself in the United States. Coming via Geneva to New York as a refugee from Nazi Germany, he made himself familiar with the philosophy of science of Alfred North Whitehead, Charles S. Peirce, William James, and their successors. In extending Kapp's European background, American social philosophy provided a distinctive vantage point on economics. 'American and European institutionalism,' says Kapp,

> (...) look at the dehumanization of our economic thinking as the result of isolating so-called pure economic phenomena from their social context – an isolation that stands in contradiction to the epistemological demands, for instance, of contextualism as propounded by John Dewey.
>
> (Kapp 1968b: 1)

The pragmatist idea of contextualism implies that individual actors are embedded in social interaction, in networks and social configurations. Actors perceive and found these interdependencies cognitively. The perception of the social world, of institutions and of manifest and symbolic power forms part of the pragmatic approach. Peirce's theory of signs points out that raw sense data are to be interpreted by signs, by symbolic representations or by patterns of mind (Peirce 1967, 1970; Liebhafsky 1993). Including this level of social representation and cognition into institutional analysis can open the understanding of contemporary policy issues.

Assembling American (Veblen, Commons, Dewey) and European institutionalism (Myrdal, Lowe, Perroux), Kapp developed an understanding of the economy as embedded in cultural practice. Economic action thus forms part of the social context. Economic development unfolds as a cultural process in a dynamic world. In Kapp's view (1954), embeddedness is a core element in cultural anthropology and social economics. He agrees with Polanyi's concept of embedded/disembedded societies (Polanyi 1957a) and mentions the importance of substantive rationality

and cognition in Polanyi's approach (Kapp 1983: 236, footnote 27). In an article on 'Political Economy and Psychology' Kapp (1950b) considers the psychological premises of economic action and emphasizes that all economic action proceeds in the field of interpersonal relations. He quotes from *The Great Transformation* (Polanyi 1957b) to emphasize the idea that the economic system does not persue its own objectives: it is conditioned and motivated by individual and group objectives and collective decisions, depending on culture, the drive for security, or on social interests like the idea to improve one's social standing. (Kapp 1950b: 306).

Kapp's theoretical approach to institutional economics combines American pragmatist social philosophy and European institutionalism. This quest for synthesis is reflected in his theoretical and reform-driven works on social costs and developmental issues.

Substantive rationality and cognitions

Furthermore, Kapp (1983: 236) mentions to go on with the continental European tradition of socioeconomics formed by Paul Tillich, Erich Egner, and Karl Polanyi, for they recognized the importance of substantive rationality and cognitions[2] for the understanding of social action. In these works, cognitive processes are social in a sense that they incorporate language, ideas and practices, which relate to a social culture. In Kapp's words, they depend on 'the culture as a whole (by which we mean the sum total of man-made techniques and tools, the symbolic universe as well as the ideas, attitudes, customs and the corresponding behavior patterns of a society)' (Kapp 1954: 215). This contrasts with the programme of modern neoclassical economics reducing culture to rules and norms created by rational actors when transaction costs are high, but ignoring societal objectives.

Kapp and Polanyi were both committed to analyse the 'economy as instituted process', where economic action between individuals is tied together in a pre-existing context of institutional arrangements (Lie 1991). Institutions and individuals cannot be separated: human actors shape institutions and are themselves shaped by institutions. Institutions form actors' identities and preferences. Preference building and value formation on the individual as well as the group level are not stable against alternative outcomes: they are reflexive processes knitting people together. Institutions, as habitualized patterns of perception and action have an impact on identities, preferences and habits.

Over time, institutions attain relative stability (Kapp 1968a: 2). They mediate between individual action and social structure. Thus, economic action as social action is embedded in the social and cultural context. In the introduction to his manuscript on 'The Foundations of Institutional Economics', Kapp (1969) states that institutional arrangements, derived habits, attitudes, and valuations bear their own logic and dynamics. Institutions are a strategic factor in the analysis of processes of production and distribution as well as with respect to the study of long-term development processes.

This has a considerable effect on Kapp's allocation theory: social preferences form part of stated allocation optima. Social welfare has to be constituted in a

social discourse and not by aggregating individual preferences on welfare. The premises of economic action lie in formulating goals and generating preferences on private and social issues. But social efficiency is not defined in a normative or decisionist way. Production, distribution, and consumption are conceived as a network of social relations. They are embedded in a cultural matrix. Thus, the allocation of resources is laden with social meaning that might be reconstructed in a scientific dialogue.

The concept of social costs

In Kapp's classic book on the *The Social Costs of Private Enterprise*, first published in 1950, the notion of social costs 'refers to all those harmful consequences and damages which third persons or the community sustain as a result of the productive process, and for which private entrepreneurs are not easily held accountable' (Kapp 1950a: 14). Social costs originate from productive processes; they depend on the social organization of production. The commodification of nature is a source of production in market economies. However, as Kapp (1970: 841) has stated, 'economic theory continues to treat allocation, production, exchange and distribution as if they occurred in an essentially closed and autonomous "economic" sphere with only minor effects on man's natural and social environment.' The effects of this mode of thought on the environment and society 'are anything but negligible' (Kapp 1970: 841).

This is an important shift in the semantics of the social costs concept. Not primarily a failure or inefficiency, they emerge in the process of production in market societies. In neoclassical economics, the natural environment is considered a public good until its over-exploitation due to free riding (because of inappropriate property rights) makes it a scarce and thus, economic good. An evaluation in monetary terms (e.g. an individual's willingness to pay) facilitates the internalization and reduces environmental damage to an acceptable level. Kapp argued, that he is not

> denying that it is possible to attribute a monetary value to environmental damages, to human health, human life, or for that matter to aesthetic values (...). In fact in markets such evaluations are made constantly; but I'm questioning and, in fact, denying that monetary values constitute appropriate and responsible criteria for the evaluation of the damages caused by environmental disruption.
>
> (Kapp 1977: 531)

Kapp lists the pollution of air and water, soil erosion and deforestation, i.e. the destruction of renewable resources as a result of uncontrolled competition in the utilization of these free goods, the exhaustion of non-renewable resources and the consequences for future generations, as well as the effects of urban growth and the concentration of industries on the human environment. Furthermore, he mentions effects on the human body as harms to health, occupational diseases

because of long-term exposure to hazardous chemicals, industrial accidents, and effects caused by the economic cycle and technological change: unemployment, the obsolescence of knowledge and skills, and the decline in the value of capital goods. He emphasizes the increase in costs as a result of missing cooperation in the production and distribution of goods, R&D, governments and scientific institutions. He further deals with the negative effects of planned obsolescence in consumer goods and deprecates in a Veblenian sense the effect of fashion and styles on scarce resources, as well as the costs of advertisement and aggressive selling.

Looking at the theory lying behind this enumeration of features, we have to state that this concept is less determined than the concept of externalities in economic theory, where institutions are assumed to be exogenous. A neoclassical economist would understand this in the following way: given the 'institutional framework', the society has to identify a trade-off. Kapp's concept of social costs, however, includes the institutional framework and the society, which has to make choices, for example on the qualitative level of environment in a substantive, not a formal way. In the last chapter of his book on social costs entitled 'Toward a New Science of Political Economy' he points out the implications for his institutional approach to economics:

> Indeed the really important problems of economics are questions of collective decision-making which cannot be dealt with in terms of a calculus deductively derived from a formal concept of individual rationality under hypothetically assumed and transparent conditions.
>
> (Kapp 1978: 288 f.)

Kapp refers to 'collective decision-making' as a process of social evaluation, what might be called a social discovery process.

According to Kapp's dynamic approach to an institutional theory of economics, the analysis of social costs is a central element to understand what might induce institutional and social change. Kapp prefers the notion of social costs to the term 'externalities' which, in neoclassical economics, 'implies that uncompensated side effects are exceptional rather than pervasive, incidental rather than systemic' (Swaney/Evers 1989: 8). Kapp (1977: 529) regards environmental disruption and social costs as

> anything but exceptions or minor side effects of economic processes. Rather, they are pervasive consequences having global and regional effects which alter not only the conditions and the quality of human life, but also may affect and endanger the process of social and economic reproduction.
>
> (Kapp 1977: 529)

Social costs are basically non-market interdependencies. Institutional change actually generates social costs. Thus, changes in the institutional structure are particularly significant for the analysis of social costs. This is the dynamism of

market societies: market systems tend to create new institutions that actually generate social costs and democratic market societies perform an agenda of social choice between institutional arrangements in order to internalize social costs.

Substantive rationality and social choice

Kapp's notion of social costs goes beyond Coase's (1960) understanding of social costs as reciprocal and external to the decision-making unit (Kapp 1977: 531). The final determination of social costs is a matter of social evaluation and social value – hence depending on social objectives. In Kapp's institutional perspective social costs become selected in a social process of evaluation. They are the result of interdependent human action. Social costs become relevant in an ambit of socially integrated actors, their cognitions, and social preferences. Thus, the social recognition and evaluation of externalities as social costs form part of economics as a social science (Kapp 1978: 41). Externalities as non-market phenomena, as ever-present results of interdependent interaction, have to become socially recognized. The societal evaluation of externalities as social costs is related to the social context: to communities, interest groups, technical and social science experts, and results in minimum standards reflecting social priorities (Kapp 1978: 293, 297). Social costs may lead to a call for public policy and may even initiate institutional change. Thus, the perception of social costs has an effect on social welfare. The problem of social costs is a problem of social choice: of collective decisions concerning social welfare:

> If modern economic science is to concern itself with social costs and benefits, it will not be sufficient merely to trace and classify these largely neglected aspects of socio-economic reality in general terms. Nor will it be adequate to present quantitative estimates (in monetary terms) of social costs. No matter how important such estimates may be as a first approach to an appraisal of social costs, the final determination of their relative magnitude and significance is likely to be a matter of social evaluation and social value (in the sense of value to society).
>
> (Kapp 1978: 292 f.)

Interestingly enough, Kapp's methodological approach is quite close to that of Nobel laureate Amartya Sen (whose premises, of course, are much more liberal). With reference to problems of environmental degradation and underdevelopment, Sen writes:

> The threats that we face call for organized international action as well as changes in national policies, particularly for better reflecting social costs in prices and incentives. But they are also dependent on value formation, related to public discussions, both for their influence on individual behaviour and for bringing about policy changes through the political process.
>
> (Sen 1995: 17)

The idea of social value or social evaluation forms an important part of Kapp's understanding of social costs and has further implications. Due to the changing nature of society and economy, the idea of measuring (or even estimating) social costs is one step too short:

> (...) the magnitude and relative importance of (social costs and – RH) social returns is not subject to measurement in terms of exchange values but requires some kind of social estimate, it becomes clear that the concept of social value is likely to be one of the most important categories of any system of economic thought willing to explore the nature and repercussions of governmental action and economic policy making.
>
> (Kapp 1978: 293).

With regard to public expenditure patterns, Kapp calls the idea of measuring social costs and benefits a 'positivistic' approach, 'which can at best reveal only ex post facto social priorities as they existed at a point in time under a given distribution of power and counter power between different groups' (ibid.: 294). Power relations among the major stakeholders, agenda control, and the reinforcement of institutional structure, are critical junctures in historical development.

Social costs are bounded by political cultures and they have to become socially qualified before they can – as externalities – become subject to internalising strategies in the Pigou-Coase tradition. If the first step fails, '(...) such studies can hardly claim any objective validity other than that which may be attributed to the historically existing distribution of power in society' (ibid.). Thus, the validation of externalities is culturally dependent[3] (Mayhew 1987). Kapp's theoretical and empirical analysis of the social costs of market economies, again, exceeds the notion of negative external effects of production and consumption in the neoclassical approach. The value of perceived social costs is a product of social institutions and distributional conflicts. The review of Kapp's works shows that social choice processes consist of perceiving and valuing social costs (external effects and other social interdependencies, such as preferences on social relations) as well as in constructing their relevance in a social discourse taking place within the institutions of a democratic society. He promotes these ideas in the last chapter of his revised book on social costs under the headline of social value and social choice:

> A substantive definition of the nature of the economic problem would have to view it within the context of changing human (i.e. socio-cultural) relations where scarce means are adapted to individual as well as public needs and requirements and where the latter are in turn selected, or adapted to socially available resources. Such a substantive definition of the scope of economics leaves no doubt about the fact that economists are concerned primarily with problems of social action and social choice rather than with individual behavior seen as a formal relationship between (individual) ends and scarce means under condition of scarcity.
>
> (Kapp 1978: 288).

Communication on policy outcomes, substantial values, and minimum standards, is within the realm of social economics. Kapp's (1978: 208) institutional approach aims at the creation of pragmatic criteria for substantial rationality and social choice: social indicators and critical values. They form part of an institutionalized framework of socio-economic and legal relations to provide environmental control.

The aggregation problem in welfare theory (aggregation of individual preferences) is now part of the political decision-making process. Dimensions of social welfare emerge in the social discourse. The search for alternative institutional arrangements to improve allocation or social efficiency is integrated in the institutional approach to economics. In this context, efficiency is a result of the societal objectives and of the combination of institutions. The power of definition, articulation, and evaluation evolves in the social choice process. Evidence may become a resource of power. Because of the interrelationship of power and evidence, it is not possible to analyse policies and politics as an exchange process with given preferences at the outset. Policy processes depend on the definition of situations, norms, and interests by collective actors.

Politics is never merely the exertion of influence by organized political actors. Doing politics is always an act of interpretation, of defining situations and developing identities. This has implications for the definition of social indicators.

Social optima depend on the welfare culture of a given society. Welfare institutions and regimes and issues such as environmental protection have to be accepted as public, qualified by society before they become settled into the process of political decision-making. Perceptions of risk and of environmental goods and environmental damage also form part of the analysis. The search for more efficient institutional arrangements includes a values-oriented redesign of institutions. Choice between institutions occurs in a conflicting social process, a process of political-institutional decision-making.

This is one of the main normative preconceptions of Kapp's institutional approach: he is centring on the formation of substantial values and social issues, not on spontaneous coordination (e.g. the unintended effects of action). The evaluation of the necessity to internalize social costs and the perception of risk depend on the actual institutional setting. Thus, not only transaction costs, but also the perception of individual and social risks by social actors form part of the institution-building process. Kapp's notion of social costs contains cognitive and historical-institutional dimensions.

Open questions

Kapp himself, from my point of view today, failed to fully recognize the problem of a culturally open, underdetermined concept of social costs. Collective decision-making, to Kapp, is mainly happening in the political process, by political actors. But who are these actors – experts, bureaucrats, technicians, and scientists? The environmental movement is the one Kapp puts his hopes on to promote change. Bringing society back into politics maybe only became an issue after his death

in 1976. So he did not focus on democratic decision-making techniques that evolved with the new social movements. Kapp feels the urge to 'inquire whether social values and the process of social evaluation can be placed on an objective basis' (Kapp 1978: 294). Kapp asks: 'Are there any criteria for the determination of social values in a more or less objective fashion and in terms which would demand the consent and agreement of reasonable persons' (ibid.). But he didn't see a solution for the problem of minorities who 'either fail to be convinced or refuse to be convinced by these social evaluations (...)' and finds it 'legitimate and reasonable for organized society to overrule the wishes or interests of these groups in the common interest' (Kapp 1978: 294). I doubt that the call for 'public opinion polls and sample inquiries' (Kapp 1978: 299) will be an exit from this anti-liberal fallacy, which could easily put the legitimacy of social evaluations under stress. Kapp himself doesn't see it that pessimistic, he points out that a democratic society has to deal with the issue of conflict:

> Instead of ignoring or concealing the possibility of conflicts between individual and social values by a purely formal concept of social welfare conceived as the sum of utilities of all individuals or by some vague references to 'the will of the people' or to 'government by and for the people' it is more realistic to recognize the reality of conflicts and accept them, along with the Federalists and Karl Marx, as the typical characteristics of the political process.
>
> (Kapp 1978: 300)

Conclusion

Thus, the idea of internalising social costs by reflecting them in taxes or prices has to face the fact that social costs are socially constructed and depend on critical public debate and processes of (social) value formation. They may initiate institutions and enter the political process. Furthermore, they may even react on individual behaviour just as they may change policy agendas.

Whereas standard welfare economics deals with the maximization of social welfare, institutional theory explores the composition of value criteria: the social construction of public objectives and social consent with regard to types and results of allocation form part of the research design. The comparative analysis of institutional arrangements, in the sense of comparing imperfect worlds, is embedded in a social and political arena where social issues are constructed and the definition of problems, of relevance, occurs. In the political process, targets are never given but always changing. What are the implications of the institutionalist paradigm for the understanding of political reform?

Some implications for the understanding of reforms in the German health care system

Political action develops in institutionalized arrangements. Policy analysis has to deal with the problem of the perception and social valuation of social costs: socially

available resources, which flow into the health system (10 per cent of German GDP) and their distribution. This framework proves to be useful to understand policy changes, for example in the health sector in Germany. The chapter concludes with potential consequences of an institutional analysis for reform in the German health care system.

The German statutory health insurance system has been known as a system that provides all citizens with ready access to comprehensive high-quality medical care at a cost the country considered socially acceptable. However, an increasing concern for rapidly rising health care expenditure led to a number of cost-containment measures since 1977. The aim was to bring the growth of health care expenditure in line with the growth of wages and salaries of the sickness fund members. The recent health care reforms of 1989 and 1993 yielded only short-term reductions of health care expenditure, with increases in the subsequent years[4]. In 1993 and 1999 regulation has imposed supply-side limits by introducing sectored budgets and a reference pricing system with spending caps for pharmaceuticals. Legally set as well as negotiated budgets (translated into target volumes for general and specialized practice) and fixed sums per package for pharmaceuticals became important cost-containment strategies. 'Stability of the contribution rate' is the uppermost political objective of current health care reform initiatives. Future options under discussion include reductions in the benefit package and increases of patients' co-payments. Waivers of co-payments are imperative for low-income groups and the chronically ill. There is a need for better coordination across all types of medical, rehabilitative, nursing, and home care. Networks of providers and individual caregivers that cut across the highly sectorized delivery system promote flexible contracts between individual providers and payers, thus bypassing physician associations. These selective contracts are based on defined quality standards, with higher prices for quality care and compensation of risk structure. Improved quality shall be promoted by a spread of disease management programmes (DMPs) for patients with chronic diseases (diabetes mellitus and breast cancer up to now). DMPs are supposed to promote integrated care and overcome the strict separation between the inpatient and outpatient sector (Wörz/Busse 2002, Altenstetter 2003).

The institutional setting: can we measure quality?

The German health system is a neo-corporatist (or conservative-corporatist) one with strong interest groups: the sickness funds, non-profit self-governing bodies financed by mandated employers' and employees' contributions (payroll tax averaging about 7 per cent of salaries and wages for each), the medical professions and their self-governing bodies, employers, churches, and faith-based and secular social welfare organizations. Allocation can be labelled 'regulated competition'. The interests are highly fragmented, but all conflicts end when the political actors have to face political regulation (or want to exercise their influence on political regulation). Avoiding cost-intensive and inappropriate over-regulation in health care is the broad consensus where the interest groups in the German health

system finally meet. The success of a corporatist system is based on the capacity to self-regulation. The instrument is self-obligation.

One broad issue that somewhat unifies the different interest groups is quality management: quality indicators measure attributes of health care aspects on the level of structures, processes or outcomes to be assessed within the scope of quality improvement activities. New quality indicators have to hold adequate information by taking patients' interests into account and by involving the general public in the debate concerning the criteria for GPs' performance measures. Consensus on quality in practicing medicine is difficult to achieve: 'Different parties and disciplines in health care have different opinions on effectively changing care and propose different, sometimes conflicting approaches to improvement of patient care' (Grol 2000: 298).

Health care providers, professional researchers and the statutory sickness funds are in the task of developing evidence-based guidelines as a basis for collective contracts. The Agency for Quality in Medicine (Ärztliche Zentralstelle Qualitätssicherung, Cologne), founded in 1995 as a joint institution of the medical professions German Medical Association and the National Association of Statutory Health Insurance Physicians, for the assessment and preparation of evidence-based clinical practice guidelines in ambulatory and hospital care, promotes the following definitions:

> Medical care is considered inadequate if the kind or extent of services fail to achieve the medical treatment target or if parts of the population do not have access to necessary health services. Treatment is regarded as incorrect if it causes more harm than benefit. Excessive treatment must be assumed if medically unnecessary services or services are provided that fail to have any impact on the treatment target.
>
> (Agency 2001: 4)

Reforms in the health sector have been decided on, enforced and implemented from the top down, as have other measures such as setting specific health goals and moving towards outcome-orientated evaluation. During the last few years, however, greater bottom-up participation of key target groups became a reform issue. Developing a paradigm for a bottom-up strategy in bridging the gap between research and those who are concerned, especially the GPs, becomes more and more popular. Self-regulation (vs. external control) seems to be the best strategy. So, why isn't the performance of the health care system any better? Professional organizations, health insurances, researchers and policy makers anticipate that a systematic improvement of the quality of health care will provide a solution to the problems of adequate health care supply – under- and over-diagnosing as well as under-, over- and misuse. Quality improvement minimizes social costs in the sense of losses to society and the patient when evidence-based diagnosing and treatment is not provided.

Implementing systems for measuring and improving quality is seen as one of the most important challenges in health care today. This includes activities such as setting and implementing guidelines and quality indicators, collecting and

assessing data on the quality of medical practice, disseminating information and providing training for care providers on quality improvement models and methods (Grol 2000). The BDT interface – the German name of the software interface implemented in every practitioner's software – permits the transmission of performance measure data in primary care. Computer-based strategies allow the development and specification of quality indicators.

Going online and enhancing quality?

The use of electronic patient records (EPR) is now almost universal in practice and hospital, and several different EPR systems are in use. Robust information management is a critical challenge for the medical professions. Gathering, managing, and using clinical information is the basis for collective contracts between statutory sickness funds and general practitioners. GPs who cannot present data on their services in order to lower costs through implementation of evidence-based health care, will be excluded from the new contracting models. Integrating decision-support systems with electronic patient records will be a major step in medical informatics and quality assessment. Data on health care will have a major impact for research, not only on the effects of cost-containment, but on needs-based care (in the sense of needs as distinguished from demand and from the utilization of the health care system).

In order to enhance quality, the implementation of guidelines is considered an important instrument in general practice. In several parts of Germany, the sickness funds and the Physician's Association have implemented the systematic implementation of guidelines for the management of chronic diseases (e.g. asthma and diabetes), by means of operationalized quality indicators for the measurement of the processes and outcomes.

Health matters are important public goods (Kaul *et al*. 1999, Kaul/Faust 2001). Reforms (in the health sector and elsewhere) are decisions on the allocation of resources. Furthermore, they are normative and value-driven, e.g. in Germany patient care is based on pluralism, given the historical mix of public and non-profit, faith-based and secular hospitals, and specialized facilities. Reforms have to produce legitimacy as regards

- Costs: Between 1980 and 2001, health care expenses grew from 8 per cent to more than 10 per cent of gross domestic product. During the same period, payroll decreased and thus, the contribution rate grew from 11.4 per cent to 14.1 per cent.
- Quality (cost-effectiveness): Evidence-based medicine, practice guidelines and quality indicators. At the delivery end, effectively functioning circuits of cooperation and communication from one service sector to another are receiving heightened attention given the urgent needs of an ageing population.
- Acceptance is an outcome of a process, whose involved and interested parties have approved internal logic and decision structure. Decisions that derive

from this process are more likely to be accepted. Consequently, the difficult process of filtering and identifying priority health care topics will not only be more acceptable to the group of individuals both directly and indirectly involved but also to the vast number of interested individuals and/or parties.

- Transparency provides a common basis of understanding (internal and external transparency) and is an important element of ensuring fairness and justice (together with information).
- Chances and justice are notions, whose meaning varies widely across population groups. Devising rules to ensure a fair process of decision-making creates more legitimacy than assessing a policy outcome. A sustainable financing of the health system with respect to needs-based access to care and the exigencies of long-term care is an important concept for ensuring justice between generations, i.e. one generation does not burden the following with unsustainable debts.
- Participation is an important democratic criterion, which has to be met. It produces legitimacy.
- Accountability tells us who is doing what in the health system.
- Efficiency and feasibility: A prioritization process employs methodological definitions that are intended to allow efficient implementation within certain limits. In addition, aspects of feasibility including issues of resource utilization and time frames also determine the prioritization process.
- The European process: Harmonization and competition regulations provide the framework for future reforms.

The perception that reserves still exist in the health care system, which should be put in better use, has led to the introduction of evidence-based medicine, practice guidelines and computerization. Apart from the appreciation of cost-containing strategies, there is also an agreement, that solidarity, subsidiarity and self-governance are important features of the German health care system. On the level of society, a communitarian and inclusive culture surrounding the delivery of care must be emphasized. Health policies are the product of politics and a particular institutional and cultural context (Altenstetter 2003).

Do we need piecemeal reforms? Do we need deep institutional change? Probably both. Do we know how interventions successfully induce institutional change? Not sufficiently. Have needs and benefits been communicated well? I suggest not. Patient empowerment is more rhetoric than reality. Proactive articulation of interests by patients and health professionals to policy makers will be necessary to increase priority. Kapp gave us a frame of reference to understand and deal with imperfect information, uncertainties and interest groups in the pursuit of institutional reforms. Against this background, reform is a social process to adjust technological change, patterned behaviour, and institutional change. It is a challenge for the health community to ensure that effective policies and programmes are put on to the political agenda.

Notes

1 See Krohn (1993).
2 The Merriam-Webster dictionary mentions that the notion 'cognitions' goes back to the 15th century and means the act or process of knowing including both awareness and judgement.
3 Kapp adapted the culturalist framework to the analysis of ecological disruption and other social costs (Kapp 1970), see Heidenreich (1998 and 2000).
4 For surveys of the German health care and health insurance system see Kamke (1998) and Altenstetter (2003).

References

Agency for Quality in Medicine (2001) *Critical appraisal of clinical performance measures for quality management criteria and quality indicators in health care*, consensus paper presented by the German Medical Association, the National Association of Statutory Health Insurance Physicians and Association of the Scientific Medical societies in Germany, Jena: Urban & Fischer Verlag.

Altenstetter, C. (2003) 'Insights from health care in Germany', *American Journal of Public Health* 93: 38–44.

Coase, R. H. (1960) 'The Problem of Social Cost', *Journal of Law and Economics* 3: 1–28.

Grol, R. (2000) 'Between Evidence-based Practice and Total Quality Management: The Implementation of Cost-effective Care', *International Journal for Quality in Health Care* 12, 4: 297–304.

Heidenreich, R. (1998) 'Economics and Institutions: The Socio-Economic Approach of K. William Kapp', *Journal of Economic Issues* 32, 4: 965–984.

——. (2000) 'L'economia come scienza della cultura. La fondazione costruttivistica dell'economia sociale nell'opera di William Kapp', in P. Frigato, M. Giovagnoli (eds) *La continuità de la vita umana. Il contributo di Karl William Kapp alla fondazione di una teoria istitutzionale critica*, L'Harmattan Italia, Torino.

Kamke, K. (1998) 'The German health care system and health care reform', *Health Policy* 43, 2: 171-94.

Kapp, K. W. (1950a) *The Social Costs of Private Enterprise*, New York: Schocken Books.

——. (1950b) 'Political Economy and Psychology: the Survival of Hedonism and the Research Tasks of the Future', *Kyklos* 10, 4: 291–315.

——. (1954) 'Economics and Behavioral Sciences', *Kyklos* 7, 3: 205–225

——. (1968a) 'In Defense of Institutional Economics', *The Swedish Journal of Economics* 70, 1: 1–18.

——. (1968b) 'Economics and Rational Humanism', *Kyklos* 21, 1: 1–24.

——. (1969) *The Foundations of Institutional Economics*, Unpublished manuscript.

——. (1970) 'Environmental Disruption and Social Costs. A Challenge to Economics', *Kyklos* 23, 4: 833–48.

——. (1977) 'Environmental Disruption and Technology: New Frontiers for the Social and Natural Sciences', *Journal of Economic Issues* 11: 527–40.

——. (1978) *The Social Costs of Business Enterprise*, Nottingham: Spokesman.

——. (1983) 'Social Costs, Economic Development and Environmental Disruption', in J. Ullmann (ed.), Lanham: University Press of America.

Kaul, I. I. Grunberg and M.A. Stern. (1999) *Global Public Goods*, New York: Oxford University Press.

Kaul, I. and Faust, M. (2001) 'Global Public Goods and Health: Taking the Agenda Forward', *Bulletin of the World Health Organization* 79, 9: 869–74.

Kleinman, A. (1980) *Patients and Healers in the Context of Cultures*, University of California Press, Berkeley.

Krohn, C. D. (1993) *Intellectuals in Exile. Refugee Scholars and the New School for Social Research*, Amherst: University of Massachusetts.

Lie, J. (1991) 'Embedding Polanyi's Market Society', *Sociological Perspectives* 34: 219–235.

Liebhafsky, E. E. (1993) 'The Influence of Charles Sanders Peirce on Institutional Economics', *Journal of Economic Issues* 27: 741–754.

Mayhew, A. (1987) 'Culture: Core Concept under Attack', *Journal of Economic Issues* 21: 587–603.

Payer, L. (1988) *Medicine and Culture. Varieties of Treatment in the United States, England, West Germany and France*, New York: Holt.

Peirce, C. S. (1967 and 1970) *Schriften I. Zur Entstehung des Pragmatismus, Schriften II. Vom Pragmatismus zum Pragmatizismus,* 2 vols., K. O. Apel (ed.) Frankfurt/M: Suhrkamp.

Polanyi, K. (1957a) *The Great Transformation. The Political and Economic Origins of our Time*, Boston: Beacon Press.

Polanyi, K. (1957b) 'The Economy as Instituted Process', in K. Polanyi, C. M. Arensberg and H. W. Pearson (eds) *Trade and Market in the Early Empires. Economies in History and Theory*, New York and Glencoe: The Free Press.

Sen, A. (1995) 'Rationality and Social Choice', *American Economic Review* 85 (1): 1–23.

Swaney, J. A. and Evers, M. A. (1989) 'The Social Cost Concepts of K. William Kapp and Karl Polanyi', *Journal of Economic Issues* 23: 7–33.

Wörz, M. and Busse, R. (2002) 'Structural reforms for Germany's health care system?', *Euro Observer* 4: 1–3.

3 Social costs, social rights and the limits of free market capitalism: a re-reading of Kapp

Maurizio Franzini

Introduction

One of the most impressive and well-known results in economic theory is that perfectly competitive markets, unaided by other institutions, may ensure a Pareto-efficient situation, a condition that no society can disregard light-heartedly[1].

On the other hand, for many decades now, mainstream economists have been well-aware that the ability of a free competitive market system to ensure Pareto-efficiency may be undermined by externalities, i.e. uncompensated damages suffered as a result of somebody else's action not consented to[2].

Externalities, therefore, may strike a hard blow to traditional economics by seriously weakening its appealing thesis on the efficiency of free markets. One important effect of this thesis has been to induce most economists to believe that a decent society requires little more than competitive markets and effective protection of property rights.

Externalities, however, have been metabolized by mainstream economics and there have been no significant changes either in the body of that theory or in the recommendations flowing from it. Basically, externalities are like exceptions which prove the rule.

Outside mainstream economics, several scholars with more radical inclinations based their harsh criticism of free markets in a capitalist society precisely on uncompensated damages. In their opinion, such damages – which they preferably call 'social costs' – make it necessary to revise the conventional wisdom about the virtues of free markets. This is what economists like William Kapp and Federico Caffè, among others, strongly believed.

A direct comparison between these two competing views took place at the beginning of the 1970s when William Kapp and a distinguished orthodox economist, Wilfred Beckerman, quarrelled specifically over the impact of social costs and externalities on conventional economic theory. In the article that started the debate, Kapp stated that 'the conventional framework and tools of economic theory are ill-adapted and in fact irrelevant for the analysis of the phenomena under discussion' (Kapp 1970: 839). To this Beckerman replied: 'I cannot accept without protest the general theme of Professor Kapp's paper' (Beckerman 1972: 103) and 'Perhaps Professor Kapp has missed something through dismissing contemporary

economics too contemptuously?' (Beckerman 1972: 106). In his rebuttal, Kapp patiently spelled out some of his arguments, which were undoubtedly interesting, but surely did not change Beckerman's mind. The problem is extremely complex. However, it can be approached in many different ways. In this paper I offer my reading of the debate and my opinion on what lies at the root of this controversy. I will not try in any sense to construct a different type of economic theory compatible with the social costs phenomena, as Beckerman repeatedly and polemically asked Kapp to do in the mentioned paper. It is not a matter of building a radically new economic theory but of drawing the right implications from economic theory and from real-world phenomena that can hardly be considered minor exceptions. The dispute revolves mainly around the soundness of economic arguments in support of free markets in a capitalistic system.

As I understand it, Kapp and Caffè, on the one side, Beckerman and lots of orthodox economists, on the other side, are talking about completely different problems, even though the wording might appear similar. Beckerman looks at uncompensated damages from the point of view of economic theory and as a challenge to the Pareto-efficiency of free markets. Kapp and Caffè, on the other hand, were concerned not with the efficiency properties of the market but with the violations of social rights, which are seen as closely tied to social costs. Their primary intent, I believe, was how to make social rights compatible with free-market capitalism. And they found that this was not an easy task at all. The implications of these opposite ways of looking at uncompensated damages are far-reaching not so much for the foundations of economic theory as for the recommendations economists can make on how to build a decent society and the role that free markets should have in it.

This paper is organized as follows. In the first section I present the dominant interpretation of externalities and its almost exclusive concern with Pareto-efficiency. Then I show how different Kapp's social costs are – being the other side of social rights, which are non-existent in traditional economics. In the next two sections I briefly explain why social rights protection may conflict with both capitalistic property rights and free markets, showing the inadequacy of the institutional recipe (free markets plus capitalistic property rights) recommended by traditional economics. Finally I go into the very difficult and unsolved question of how to protect social rights in a non-paternalistic society – a crucial problem to which, I believe, Kapp so timely drew our attention.

Externalities and efficiency

It is widely believed that the notion of externality made its first appearance in Alfred Marshall's *Principles of Economics* in 1890. Since then a huge amount of economic studies has been devoted to externalities but, unsurprisingly, no shared consensus on the exact meaning of the term has emerged.

Several economists, all of them adopting the mainstream approach, have proposed their interpretations, which differ from one another in several important respects. After the founding fathers Marshall, Sidgwick and Pigou we can recall,

among many others, Knight (1924), Meade (1952) Scitovsky (1954), Bator (1958), Coase (1960), Buchanan and Stubblebine (1962), Demsetz (1967), and Arrow (1971)[3].

This is not to deny that a necessary condition for a negative externality is an uncompensated damage inflicted to a non-consenting party. Everybody agrees on that. Still there are many unsettled questions.

Nevertheless, a slow process of convergence seems to have unfolded over time and a dominant approach has taken shape. The first distinguishing feature of such an approach is the lack of concern for the context within which the uncompensated damage originates. It makes no significant difference whether the damage arises in the production or distribution of goods or in the consumption sphere. Nor does it matter whether the damage is caused by a monopolistic firm or an unaware charity.

All that matters – and this is the second distinguishing feature – is that there is a loss of welfare suffered by a non-consenting and non-compensated party. No other dimension of the damage is given. A good example is the natural environment: its degradation, however serious it may be, is not an externality insofar as nobody believes its welfare is impaired. This is a logical consequence of the welfaristic hallmark of modern economic theory.

The third and, in my opinion, most important feature of the dominant approach is that it focuses only on uncompensated damages that do infringe Pareto efficiency. Buchanan and Stubblebine (1962) make this point very clear when they define Pareto-relevant externalities i.e. externalities the elimination of which may benefit all the parties. In their own words

> (...) externalities, external effects, may remain even in full Pareto-equilibrium. That is to say, a position may be classified as Pareto-optimal or efficient despite the fact that, at the margin, the activity of one individual externally affects the utility of another individual (...) This point has significant policy implications for it suggests that the observation of external effects, taken alone, cannot provide a basis for judgement concerning the desirability of some modification in an existing state of affairs. There is not a prima facie case for intervention in all cases where an externality is observed to exist. The internal benefits from carrying out the externality, net of costs, may be greater than the external damage that is imposed on other parties.
>
> (Buchanan Stubblebine 1962: 208–9)

Therefore when economists speak of (negative) externalities they refer only to those actions that cause losses to somebody greater than the advantages they bring to somebody else. Actions whereby losses are less than benefits are disregarded, regardless of any other consideration, because they are consistent with Pareto efficiency – i.e. it is not possible to make both parties better off. From the standpoint of Pareto efficiency, a rich acting in such a way as to increase their welfare more than the welfare of a non-consenting poor decreases would not inflict significant damage, even though such little loss may drive the poor to starvation.

In a free market only those transactions that have a net positive value will be carried out: nobody would spontaneously accept to sell something that costs more than the price the buyer is willing to pay. This helps to clarify the close connection between Pareto-relevant externalities on the one hand, and efficient markets on the other.

Externalities that are not Pareto-relevant are outside the reach of free markets, therefore their persistence cannot be blamed on the bad working of the market. Strictly speaking, the market fails only insofar as Pareto-relevant externalities persist. Therefore the agenda of the orthodox approach is set: focus on that type of externality only and endeavour to demonstrate that its origin lies not in badly working markets but in the lack of some market, in the failure to satisfy all the preconditions for markets to develop. It is a well-known fact that this is what the Coasian 'revolution' has brought about, at least in the dominant interpretation.

This concern for the efficiency property of the markets explains why the dominant approach picks up externalities that are 'anonymous', uncompensated damages, defined as a loss of perceived welfare, the magnitude of which is greater than the benefit accruing to the advantaged party.

It is my contention that to fully understand what marks the difference between orthodox and more radical approaches to externalities one has to bear these features in mind. Otherwise, authors like Kapp or Caffè,[4] who fully belong to the more radical camp, would be incomprehensible. I believe that their first concern was not the efficiency property of the market, even though this is a very important issue.

To them, and to other unorthodox scholars, externalities and social costs were the sign that market capitalism was plagued by a very serious problem: the lack of institutional protection for some basic social rights. Therefore to talk about externalities and social costs is not, primarily, to talk about the Pareto-efficiency of the market but how to reconcile free markets with social rights, and how difficult it is to bridle the capitalistic market so as to eliminate the most unacceptable of those uncompensated damages determined by the violation of social rights. This approach implies that the type of uncompensated damage to look at is quite different from the one which takes centre stage in the more traditional analysis.

Losing sight of these differences makes it impossible to fully appreciate Kapp's ideas and to understand the deepest reasons for his distance from the more traditional approach. The importance of these reasons should not be discounted, as borne out by the rather clumsy debate with Beckerman, a stern defender of orthodoxy, which took place at the beginning of the 1970s[5].

In the next sections I will try to make these points as clear as possible, starting from a brief reconstruction of Kapp's conception of social costs.

Social costs and social rights

'In short, the term social costs refers to all those harmful consequences and damages which other persons or the community sustain as a result of productive processes and for which private entrepreneurs are not held accountable' (Kapp 1963: 13).

This concise and clear definition of social costs highlights some important aspects of Kapp's vision.

First, it is irrelevant whether the action under consideration brings about damages that exceed the gains it yields to the acting party. Secondly, and more importantly, Kapp explicitly and exclusively mentions damages resulting from production carried out by private businesses.[6] He does not care about uncompensated damages of different origin. The anonymous agent of the orthodox approach now takes on a clear identity.

To shed more light on Kapp's vision, let us consider another quotation where he intends to assert that social costs are not the result of environmental disruption alone:

> (...) by stressing the ecological aspect may divert our attention from those social costs which find their expression in such phenomena as work injuries and accidents, rhythms of work inimical to human health, crowded and inadequate housing conditions, damaging levels of noise, enforced and uncompensated adaptations to structural changes, workman compensation systems rendered inadequate by inflation and, last but not least, monopolistic determination of real estate values and rents in congested urban areas, all of which can and do arise in contemporary industrial societies.
>
> (Kapp 1970: 838)

This is a long and apparently messy list. Kapp himself admits that social costs so defined cover a great variety of social losses and therefore lack definiteness and precision. But

> Our concept identifies these characteristics in terms of the empirical consequences of production activities which the entrepreneur does not bear but is able to shift to other persons or to the community at large. In this sense our concept is both concrete and operational.
>
> (Kapp 1963: 20)

It is definitely clear that Kapp believes that the distinctive character of social costs is that they should be borne by entrepreneurs, but are not. It is as if they were dodging their responsibility. If entrepreneurs have a duty to bear these costs then the 'community at large' has the right not to be burdened by them. In other words, entrepreneurs are also violating someone else's rights. In a nutshell, social costs are the monetary side of unprotected social rights. To look at social costs as violated social rights has, in my opinion, far-reaching consequences and puts on a firmer ground the difference between orthodox externalities and more radical social costs.

Indeed, all the listed items in the preceding quotation relate to some type of social rights. Also, because of this it is foolish to assume that Kapp would consider these violations less relevant and would not care about them, were the value of damages in terms of perceived welfare loss smaller than the advantage enjoyed by business firms violating those rights.

I am sure he was not interested at all in Pareto-inefficient situations where some rich people suffer a significant loss as a consequence of an action yielding only a small advantage to a poor. This may be Pareto-relevant but hardly fits into the conception of social costs as infringements of social rights. Social costs might very well arise in a context that is not relevant to Pareto-efficiency, even though this is not a necessary condition. Kapp's perspective is by no means Pareto-efficiency.[7]

He wants to draw our attention to the many types of damage that businesses impose upon non-consenting people in the real world. He looks at the whole set of uncompensated damages caused by private capitalism in its real working. It is not a coincidence that the examples he provides are so different from the neutral and anonymous cases typically considered by the more traditional literature on externalities (the rancher and the cattler, not to mention the bees and the flowers).

Kapp's crucial problem – and also Caffè's – is not, therefore, how uncompensated damages undermine the market ability to ensure Pareto-efficiency. His problem is whether market capitalism can be reconciled with some basic social rights. As I say below, he is very sceptical about this possibility, to put it mildly.

Here lies, in my opinion, the crucial difference between Kapp and the more orthodox approach. He is at pains to make this difference clear in his writings arguing that social costs are different from externalities, making several points to this end.

Kapp argued that social costs are more serious, pervasive and widespread than externalities are believed to be. Therefore they are not a minor disturbance that can be omitted from a more general analysis and mended by the market itself (Kapp 1963: 271). He tried to show that externality is too vague a concept, an empty box, a cover-all concept devoid of practical relevance. He also criticized the static nature of externalities but, to be honest, did not develop his analysis of social costs in a full-fledged dynamic setting. He pointed his forefinger at the closed model used to study externality that underestimates interdependencies and cumulative development.[8] From all this he drew the conclusion that 'the conventional framework and tools of economic theory are ill-adapted and in fact irrelevant for the analysis of the phenomenon under discussion' (Kapp 1970: 839).

All these points are important. However none of them goes right to the root of the problem, as I see it, which is the nature of the uncompensated damages under consideration. Social costs are infringements of social rights perpetrated by market capitalism, be this in contrast with Pareto-efficiency or not. From this perspective, they can support a very radical critique of real markets. To reconcile those rights with market capitalism is not an easy task at all, because most social costs grow out of the lawful exercise of basic capitalistic property rights. Actually, Kapp's problem can be framed as a clash between social rights and capitalistic property rights.

Social costs and property rights

There cannot be free-market transactions without well-defined property rights. Should there be no agreement on the person who owns specific resources, and therefore is to be paid for the right to use such resources to be transferred,

no exchange could take place in the market. Starting from this uncontroversial but widely overlooked observation, Coase (1960) came to the conclusion that the lack of well-defined property rights is the ultimate cause of externalities.

Indeed, when transaction costs – broadly conceived – are nil, well-defined property rights are a necessary and sufficient condition for all the reciprocally advantageous transactions to take place between rational and free agents. In other words all the opportunities for Pareto improvements will be exploited and, therefore, no Pareto-relevant externalities will persist. Clearly, Coase too looked at externalities as a worrying phenomenon only because they endangered the efficiency properties of the market.

His contribution has been interpreted as an attempt to demonstrate that far from being a proof that markets fail, externalities are the consequence of a missing precondition for their working. Given the emphasis, too often ignored, Coase placed on transactions costs as obstacles to the efficient working of the market, this interpretation may be a bit unfair to him. On the other hand, I believe that it was no little merit to remind that markets too need preconditions, i.e. they are not natural, self-sufficient systems. In particular, from Coase on it is quite clear that the market can, at least in principle, support well-diversified systems of property rights and it is a mistake to confuse the market (a system for making decisions, in the end) with capitalism (a system of property rights that has highly significant power and distributive implications). Our concern is not the debate on Coase. However, his approach is useful in several respects. First, it helps to go into the problem of the relation between social rights and property rights – an apparently awkward matching. Secondly, it allows us to identify and keep separate the criticisms that, starting from social costs, can be levelled against capitalistic property rights, on the one hand, and against market transactions, on the other. I believe that in Kapp's writings both capitalism and the market are under the fire of his critical remarks, but for different reasons.

According to Coase, if Pareto-efficiency is the goal, all that is required is a clear allocation of property rights, whatever it may be. If transaction costs are not significant, that sole condition will allow markets to work efficiently. Assume now that your goal is the same as Kapp's, i.e. to protect some social rights, the violation of which gives rise to social costs. Can Coase's proposition be reformulated so that the free market will not interfere with the protection of social rights? In principle it can. Social rights could be interpreted as property rights on some specific actions or resources.

Let us consider an example taken from Kapp's list of social costs: technological change. Kapp states that the loss of income and welfare on the part of the workers displaced by innovations introduced by business firms is to be reckoned as a social cost. More to it, he believes that this driving force of capitalist dynamics is one of the most serious and systematic sources of social costs (Kapp 1963: Ch. 10). It is not my purpose to go into this thorny problem. To highlight the relation between social and property rights, let us assume that workers have property rights on a technology so that entrepreneurs cannot change it without getting their permission, which implies buying the technology from them.

If all the agents are rational only those changes in technology that make the workers better off will take place.[9] Their social right is, therefore, protected by a combination of property rights and free exchange. Moreover, social costs will disappear, because there are no longer benefits accruing to those who, one way or another, are not entitled to them. This is further proof that Kapp's problem can be formulated in terms of property rights.

It seems we have found the solution to make free markets compatible with social rights. But the question is: would that be compatible with capitalism as a system of property rights?[10] The answer is no, at least for a wide range of social rights.

The fact is that the assignment of property rights that would eliminate social costs is hardly acceptable in a capitalistic system. Can we imagine firms paying their workers to introduce technological innovation and still speak of a capitalistic system? Or can we think of firms paying their employees when they want to introduce systems which worsen the quality of work?

To make free markets fully compatible with social rights, anti-capitalistic property rights may be needed. Looking at them from this angle, social rights clash not so much with free markets but with the system of property rights which distinguishes capitalism. This is not yet the whole picture, still it is an important part of it.

Our reasoning demonstrates that Coase's efficiency proposition can be turned upside down: since markets work in a context of predefined property rights, the market may still work in presence of anti-capitalistic property rights. Of course transaction costs and other obstacles to the working of the market, usually referred to as critical arguments against the Coasian solution, call for much greater caution in making such a statement.[11] But the point is that whether social rights conflict with capitalistic property rights and whether they are in contrast with free markets are two separate issues.

Approaching the social costs problem from the side of social rights we reached the conclusion that capitalistic property rights are very often hostile to social rights. I believe that it is not fanciful to assume that an idea like this captures a large part of what lies at the bottom of Kapp's vision.

Social rights and free markets

Free markets and appropriate property rights may not suffice to guarantee social rights protection. The problem can be approached by quoting at length from a recent writing by Kaushik Basu

> we do not have an equivalent moral code against one person becoming richer than another by outwitting the latter (...) one reason we have so few safeguards against people cutting irrational deals is that economists' assumption that all human beings are rational has seeped out into our everyday thinking. If no one is irrational there can be no need to protect the irrational. Deals that look lopsided must merely reflect differences in preferences (...) but some

> people are impoverished because they cut poor deals over and over again, just as some village lenders are rich because they systematically cut "good" deals (...) Outwitting an intellectually weaker person of her wealth (...) has far-reaching consequences, especially in global contexts (...) Hence, the lack of legal or norms-based protection against such possibilities leads to under-contracting and consequent inefficiencies.
>
> (Basu 2003: 896)[12]

One rarely reads sentences like these from an economist. Few are ready to admit that outwitting is a problem because this would seriously undermine the high moral standing of the market as an epitome of freedom.

The problem Basu draws our attention to is very general, even if specific markets present a different exposure to the risk of unfair free exchanges. The problem may be extremely serious when social rights are exchanged, and this is why the issue is so important to us. The market solution to the social rights problem, regardless of the anti-capitalistic bias of the necessary property rights, must be gauged also from this point of view. Due to lack of knowledge and information, liquidity problems or many other reasons, individuals may be too weak to resist the pressure to undersell their social rights. Thus, the goal of protecting social rights through the formula 'anti-capitalistic' property rights plus free market may be missed. The conclusion is that to cope with the social rights problem a more comprehensive institutional framework is needed. To this I shall return in the next section.

In his writings Kapp made several remarks on the working of the market – specifically interpreted as a system of individually conducted exchanges – that fit very well with the problem raised by Basu.

In his reply to Beckerman, Kapp argued that individual Willingness To Pay (WTP) as expressed in the market is not a reliable basis for evaluating things. In particular he was sceptical of the use, suggested by Beckerman and all orthodox economists, of WTP to assign a value to improvements in the natural environment.[13] He believed WTP played down the phenomenon of unpaid social costs 'by making them appear more harmless than they actually are' (Kapp 1963: 317).

Kapp lists three reasons for his scepticism. The most important for our purpose is the second, which is about the 'individual's inability to ascertain the full range of short and long run benefits of environmental improvements or, for that matter, of the full impact of environmental disruption upon his health and his well-being' (Kapp 1963: 314).[14]

Indeed there is a problem of information and knowledge, because some losses remain hidden for a long time so that the injured persons are not immediately aware of them (Kapp 1963: 13). But there is more to it, a sort of inherent irrationality in the allocation process which must concern us (Kapp 1970: 843). This is why Kapp believes 'monetary criteria are not cognitively responsible'.

These critical remarks lead to the conclusion that a richer and more comprehensive institutional framework is necessary, in addition to a clearer departure from free markets. Moreover one can argue that traditional economics cannot be of

much help in such endeavour. This is what Kapp strongly believes. And this is what Beckerman firmly denies: 'I cannot accept without protest the general theme of Professor Kapp's paper in which he argues that "the phenomena of environmental disruption and social cost demonstrate once again that the scope of economic analysis is seriously challenged"' (Beckerman 1972: 103).

In light of the preceding considerations it is all but surprising that Beckerman and Kapp take different sides on these crucial issues. The former is a stern defender of the orthodox approach. His exclusive concern is Pareto-efficiency.[15] Moreover, he sees no particular reasons to be suspicious about the reliability of individual assessments as a basis for setting the value of things. Due to the first reason, he cannot even make the first step in the direction of understanding why the problem may arise about making social rights and capitalistic property rights reciprocally compatible. Due to the second reason, he is not able to appreciate why a departure from the individualistic solution of free markets may be necessary in order to avoid outwitting and to protect social rights.

Had Beckerman acknowledged both these problems – inadequacy of both capitalistic property rights and free markets – he would have refrained from arguing that no major departure from traditional economics was necessary. In a nutshell this is what the debate between Kapp and Beckerman was about. We now have all the essential elements to judge who was closer to being right.

How to protect social rights?

It is all too obvious that to raise a problem is not to find a solution, especially when the problem is as serious as how to protect social rights without interfering too much with other basic principles of western democracies. In the preceding sections I have argued that social rights may conflict with two pillars of western civilization: capitalistic property rights and free markets. But the problem is, to some extent, more general because several solutions we can think of contrast with other defensible principles. To give an example: some solutions may be highly paternalistic and this, at least in some cases, may be at odds with the foundations of democracy. Moreover, there are social rights that conflict with one another, posing highly challenging moral problems.

The problem raised by Kapp, as I understand it, has not yet been solved, after all these years. Actually, no easy and perfect solution is within our reach.

To protect individuals and their social rights there are, in theory, several possible remedies which could also be combined with one another. Some of these remedies are preventive, others compensatory; some are paternalistic, others based on more democratic procedures; some work through the markets, others do not. But all face complex difficulties and none of them are free from imperfections and defects.

A first type of solution is a collectively decided compensation, typically of a monetary nature. Several provisions of the welfare state, as we know it, fit into this scheme and have been justified exactly on these grounds (Titmuss 1987). It is all too clear that with respect to some rights an ex post compensation is not

a good solution. More generally, it can be quantitatively inadequate and, above all, subject to the vagaries of political and economic tradeoffs. One has only to recall that the 'big' welfare state that many western countries built up in the Seventies and Eighties today is considered no longer affordable by a large majority; as a consequence – inefficiencies and waste aside – some social costs will not be compensated and the range of protected social rights will shrink.

Also because of this, one might wonder whether the inclusion of social rights in constitutional law would help to solve the problem. As is well-known, this has been one of the most debated issues in the process of writing a European 'constitution'. Another recent experience is that of South Africa, where the new constitution acknowledges some of those rights (Sunstein 2001). The crucial problem here is whether socio-economic rights are within judicial capabilities. As Sunstein puts it

> How can courts possibly oversee budget-setting priorities? If a state provides too little help to those who seek housing, maybe it is because the state is concentrating on the provision of employment, or on public health programs, or on educating children. Is a court supposed to oversee the full range of government programs, to ensure that the state is placing emphasis on the right areas? How can a court possibly acquire the knowledge, or make the value judgments, that would enable it to do that?
>
> (Sunstein 2001: 3)

This is a serious problem, which can be given only imperfect and quite complex responses.

To protect social rights, public expenditure is not the only solution, nor the most effective. An alternative is to keep in check business behaviour by means of monetary incentives or through the introduction of standards. All this is well-known to those who are familiar with the debate on environmental policy, where a distinction between price and quantity instruments is drawn (Weitzman 1974).

Kapp did not like monetary incentives, like taxes. He was much more in favour of standards that could be used not only to prevent environmental disruption but also to protect other social rights. He argued that environmental disruption cannot be defeated 'with customary ways of thinking and established methods of conducting business, modified only by a few and ad hoc controls' (Kapp 1970: 836).

There are disadvantages with standards. In particular they may not be very cost-effective, as shown by a large number of studies on the comparison between economic instruments and 'command and control' policies.

But Kapp's primary concern in this matter was a different one, not surprisingly related to distributive issues. He feared that ex post remedies, like incentives and disincentives via subsidies and taxes, might turn out to be ineffective (Kapp 1970: 836). A crucial question to him was: who is the ultimate bearer of the taxes levied on the firm? Will the market power firms enjoy allow them to pass on to consumers environmental taxes without any visible change in their attitude towards the environment?[16] These are difficult questions. Perhaps Kapp is much too suspicious about economic instruments and his conclusions are a bit hurried.[17]

But there is a more general and more fundamental question. How do we decide which social rights need protection and how wide this protection should be?

Do we need deliberative democracy?

Deciding which social rights are to be protected and the amount of resources that can be used to protect them is all but an easy problem. In a society institutionally characterized by capitalistic property rights and free markets – and nothing else – social rights would not be a problem because they are not even recognized. This is exactly what happens in the abstract world of more traditional economics – which is not, however, a good picture of reality.

Once we care about social rights we have several problems: defining them, protecting them and – most of all – avoiding that some social rights are protected at the expenses of others without being aware of that.

To take this stand is to go somewhat beyond the social cost problem as presented by Kapp. But his contribution is an essential starting point. In fact, while business firms are very often the primary creators of social costs, the most likely breakers of social rights, they are not the only ones.

Moreover, the infringement of social rights does not always further business interests. Think of some tragic choice like deciding whether a limited amount of resources is to be allocated to a hospital for organ transplants or to paediatric care. Two basic social rights here are clashing. Think also of the conflict that very frequently arises between the workers of a polluting factory and the inhabitants of the town where the factory is located – sometimes friends and relatives of those workers, if not the workers themselves, who may very well have substantial misgivings about the situation.

My point is that there are many instances in which the conflict is not simply between social rights and private profit – as in Kapp. There are several instances in which some social rights interfere with other social rights. If we have a solution for these tragic cases we are in a much better position to solve the social rights – private profit conflict. This is why it may be advisable to go beyond social costs to face up to the moral dilemma implicit in the clash between different social rights.

Unfortunately there is no ready-made solution, but neither are we at the starting point. Paternalism is a possible response. A team of experts or of politicians may be delegated to solve the most difficult problems. There are good reasons for advocating this solution in some cases. However, it would be wrong, in my opinion, to rely on paternalism only. Because even a benevolent paternalist makes mistakes – and more often than not paternalists are not benevolent.

Democracy, of the right type, is a better guarantee. What is needed is that kind of deliberative democracy which has been advocated also in recent years by several scholars (Elster 1998, Guttman-Thompson 1996). Deliberation is free discussion, which, as it unfolds, helps people to forge their responsible opinion, makes them ready to take difficult decisions and softens social conflict. There is no guarantee that everybody will share the same view, however, but there are

good reasons for believing that reciprocal understanding, and the propensity to cooperative action will be reinforced, with a low risk of free riding. In a deliberative democracy decisions are not simply aggregation of individual preferences, whatever the degree of information and knowledge of the persons involved.[18]

Such a procedure is to be preferred to the delegation of decisions to experts. On the contrary, experts should make their knowledge and information accessible to everybody, leaving the decision to a well-informed set of interacting individuals. Whatever advantage the expert enjoys in technical knowledge, he has no lead in the choice of how to compare two conflicting social rights.

It is reassuring to learn that Kapp made statements that go in this direction, showing to some extent that he too was going beyond the social costs problem, i.e. the case in which social rights are broken by profit-eager firms only.

The need for a collective decision is clearly stated in this quotation: 'the elaboration and acceptance of environmental goals call for a collective or social choice with a direct participation and expression of preferences by all members of society, even those outside the market and without reference to effective demand'[19] (Kapp 1963: 317).

In the following quotation, Kapp seems to be talking specifically about deliberative democracy:

> (...) in order to satisfy these human needs and to arrive at a substantive rationality in the utilization of society's scarce resources, these requirements (environmental requirements) will have been defined as objectively as our present knowledge permits and evaluated by means of a deliberate collective, i.e., political decision in comparison to other public goals to be pursued.
>
> (Kapp 1963: 317)

These are the final sentences of his reply to Beckerman, maybe the last stage in the development of his thought. How far all this is from traditional economics need not be stressed.

A society that gives social rights the importance they deserve should not fail to consider extensive recourse to deliberative democracy as the most important institutional progress it can make. It has been correctly said that deliberative democracy helps to sustain a conception of democracy with a capacity for moral improvement (Guttman-Thompson 1996: 9). And there can be no progress in social rights without moral improvement.

We should not forget, however, that deliberative democracy has its problems: it may be costly, it may take a lot of time to produce its decisions, and it cannot include all the persons concerned.[20] As I said before there is no simple and effective solution to this problem, so crucial to modern democracies, that Kapp so timely and so passionately pointed out to us. A problem, which lies outside the realm of traditional economics, making for the limited relevance of this theory to the true progress of our societies.

Concluding remarks

In this paper I have argued that the best way to look at social costs is to consider them as the result of the infringement of social rights of one type or another. Viewed in this perspective, social costs have much less in common with externalities – as they are predominantly interpreted in the economic literature – than one might think and the problem Kapp was struggling with is more aptly reformulated as one of difficult compatibility between the protection of social rights and free-market capitalism.

I sought to clarify why social rights are a challenge both to capitalistic property rights and to free markets, keeping these two building blocks of the system as separate as it is convenient to do. The result that Kapp reached taking social costs as his starting point – i.e. 'the scope of economic analysis is seriously challenged' – stands on firm ground, if regarded from the perspective I have suggested.

Approaching the problem as a Pareto-efficiency conundrum, which is what Beckerman and others do, leads to different landing-places and hides the fundamental reasons of the inadequacy of traditional economics and of the economic system it recommends.

The protection of social rights, once we acknowledge they exist, is a compelling reason to take very seriously questions which go to the heart of economics: do capitalistic property rights give the right incentives? Are individuals always able to protect their interests and their rights in a system where exchange is free in any meaningful sense? Are individual market-based solutions always adequate?

All these questions have positive answers in traditional economics. But Kapp knew and we all should know by now that such positive answers are very easy and not apt to allow our complex society to make civil and social progress. We need to take seriously the challenge of social costs and social rights, thinking of them as the starting point for designing richer institutional settings, where decision mechanisms are more collectively oriented than the market and property rights systems are more subservient to the fulfilment of social rights.

Notes

1 As it is well-known, Pareto-efficiency is fulfilled when there is no alternative to the established situation such that somebody is better off and nobody is worse off.
2 Talking of damages I consider negative externalities only. Of course it is also possible to have unpaid benefits, in which case we face a positive externality.
3 Papandreou (1994) carefully reconstructs the most important aspects of this debate.
4 In what follows I refer repeatedly to Kapp's writings. On Caffè's approach to externalities see, among others, Caffè (1990) and Ciccarone's remarks (1995).
5 The debate was originated by an article by Kapp (1970). Beckerman wrote a harsh reply (Beckerman 1972) to which Kapp reacted firmly but politely (Kapp 1972).
6 One has only to remember the title of his best-known book (The Social Costs of Business Enterprise) to understand the importance Kapp attached to this aspect. To be sure, he explicitly stated that this is one distinguishing feature of social costs, the other being their avoidability (Kapp 1963: 13–14).

7 Indeed Kapp criticizes this criterion by subscribing to Reder's remarks (Kapp 1963: 41).
8 Kapp asserts that the true problem are those macroeconomic extra-market and extra-industry cause and effect relations which are cumulative in character and that neoclassical economics has neglected, in contrast to the classics, Marx and Veblen (Kapp 1970: 842).
9 I come back to this assumption in the next section where I take up Kapp's criticism to individual rationality and the reliability of *willingness to pay* as a measure of welfare gains and losses.
10 By capitalist property rights I mean that the owners of capital – or their representative – enjoy the residual right of control and decision within the firm.
11 I said before that Coase was well aware of the practical relevance of those costs; thus he cannot be charged with neglecting them.
12 On this problem and many related issues, see also Basu (2000).
13 Kapp writes: 'The major emphasis of my reply will be on what I consider to be the central issue raised: namely the question of the adequacy of evaluating environmental goals and values in terms of the individual's willingness to pay or accept compensation' (Kapp 1963: 306). This statement can be generalized beyond environmental disruption, as he himself hints. (Kapp 1972: 90)
14 The first reason relates to the role that income distribution plays in distorting the outcome in favor of what rich people prefer; while the third is about the possibility that other solutions (in particular support to research) are crowded out by the monetization of environmental disruption. Both are relevant and still debated today. However, for a possible solution to the bias due to unequal income distribution see, among others, Pearce-Barbier (2000).
15 A recent and clear confirmation comes from his latest book. Beckerman criticizes some distinguishing aspects of sustainable development because they may run against Pareto-efficiency, which is considered a sort of sacred and inviolable principle (Beckerman 2002).
16 Kapp is afraid that these policies will achieve 'not much more than the passing on to consumers or to society as a whole the costs of "cleanliness" without really coming to terms with the serious problems raised by the current disruption of our environment'.
17 Unfortunately Kapp does not delve into the problem. He makes a scant reference to a short statement by Schaar-Wolin (1970). More specifically, the two authors wrote: 'It was predictable that the first target chosen by the government in its new zeal for nature would be that ancient enemy, Standard Oil, which had polluted the waters of New Orleans. It is also predictable that future policies will not be implemented—any more than the Sherman Act was—to transform the corporate structure. We may expect, instead, ingenious devices for passing on to the consumer the costs of cleanliness.'
18 As Guttman and Thompson (1996: 174) put it: 'aggregating what citizens want individually, which is what utilitarian policy analysis does best, does not necessarily produce the same result as asking citizens to consider together what they want collectively'.
19 It is worth noticing that by referring also to members outside the market, Kapp anticipates the modern notion of 'non use value'.
20 In the ongoing debate on deliberative democracy there is no scarcity of perplex or squarely critical views. See Macedo (1999).

References

Arrow, K. J. (1971) 'Political and Economic Evaluation of Social Effects and Externalities', in M.D. Intriligator (ed.) *Frontiers of Quantitative Economics*, Amsterdam: North Holland.

Basu, K. (2000) *Prelude to Political Economy*, Oxford: Oxford University Press.

——. (2003) 'Globalization and the Politics of International Finance: The Stiglitz Verdict,' *Journal of Economic Literature* 41: 885–899.

Bator, F. M. (1958) 'The Anatomy of Market Failure,' *Quarterly Journal of Economics* 72: 351–379.

Beckerman, W. (1972) 'Environmental policy and the challenge to economic theory', in *Political economy of the environment. Problems of method*; papers presented at the symposium held at the maison des sciences de l'homme, Paris: Ecole pratique des hautes etudes and Mouton & co.

——. (2002) *A Poverty of Reason. Sustainable Development and Economic Growth* Oakland, California: The Independent Institute.

Buchanan, J. and Stubblebine, W. C. (1962) 'Externality,' *Economica* 29: 371–384.

Caffè, F. (1990) *Lezioni di politica economica*, Torino: Bollati Boringhieri.

Ciccarone, G. (1995) 'La tipologizzazione dei regimi di Welfare State e il pensiero di Federico Caffè', in *Federico Caffè. Realtà e critica del capitalismo storico*, Catanzaro: Meridiana Libri.

Coase, R. H. (1960) 'The problem of social cost,' *Journal of Law and Economics* 3: 1–44.

Demsetz, H. (1967) 'Towards a Theory of Property Rights', *American Economic Review* 57: 347–59.

Elster, J. (1998) 'Introduction', in J. Elster (ed.) *Deliberative Democracy*, Cambridge: Cambridge University Press.

Gutman, A. and Thompson, D. (1996) *Democracy and disagreement,* Cambridge, Massachusetts: The Belknap Press of Harvard University Press.

——. (1963) *The Social Costs of Business Enterprise*, Nottingham: Spokesman.

——. (1970) 'Environmental disruption and social costs: A challenge to economics', *Kyklos* 23: 833–848.

Kapp, K. W. (1972) 'Social costs, neo-classical economics, environmental planning', in *Political economy of the environment. Problems of method* papers presented at the symposium held at the maison des sciences de l'homme, Paris: Ecole pratique des hautes etudes and Mouton & co.

Knight, F. H. (1924) 'Some fallacies in the interpretation of social cost', *Quarterly Journal of Economics* 38: 582–606.

Macedo, S. (ed.) (1999) *Deliberative Politics*, New York: Oxford University Press.

Meade, J. (1952) 'External Economies and Diseconomies in a Competitive Situation,' *Economic Journal* 62: 54–67.

Papandreou, A. (1994) *Externality and Institutions*, Oxford: Clarendon Press.

Pearce, D. and Barbier, E. (2000) *Blueprint for a Sustainable Economy*, London: Earthscan Publications.

Schaar, H. and Wolin, S. S. (1970) 'Where we are now', *The New York Review of Books* 14: 3.

Scitovsky, T. (1954) 'Two concepts of external economies,' *Journal of Political Economy*, 143–151.

Sunstein, C. R. (2001) 'Social and Economic Rights? Lessons from South Africa,' *University of Chicago, John M. Olin Law and Economics,* working paper no. 124.

Titmuss, R. (1987) *The Philosophy of Welfare,* London: Allen and Unwin.

Weitzman, M. (1974) 'Prices vs. Quantities,' *Review of Economic Studies* 41: 477–91.

4 Increasing complexity in the 'new' economy and coordination requirements beyond the 'market': blockages and lock-ins as social costs, and a new governance to mitigate them

*Wolfram Elsner**

Introduction

The characteristics of the 'new' economy are not evidenced so much by the hypermania of high-tech stock markets – which have turned out, with their recent deflations, to be rather conventional. They are evidenced by the more deregulated, clustered, digitized and net-based character of the real economy. This change in character entails a considerable intensification of direct interdependencies among economic agents, where the outcome of A directly depends also on the behaviour of B and vice versa. This has always been the case in real economies because they are socio-economies. It has been anticipated by institutionalists and other socio-economists, but it has become so intense, obvious and ubiquitous that it can no longer be ignored by mainstream economic theorists and real-world policy makers as well.

Direct interdependencies entail direct interactions of agents. Both direct interdependencies and interactions are difficult to conceptualize in mainstream economics' 'market' models. They cannot effectively be performed by the 'market' – with its arm's-length, short-run maximising individual behaviour, ideally based only on the indirect coordination through (equilibrium) prices. Ideally, these are dependent on the decisions of all other agents taken together, i.e. no direct interdependencies among agents. 'Markets' and prices are unable to reflect information and the formation of expectations required to effectively coordinate agents in situations of direct interdependencies and interactions. They fail in such complex situations – and in this way entail increasing social costs.

This paper is about social costs in the widest sense, i.e. about general collective blockages of action and more specifically about technological lock-ins on inferior paths of development.

Why have complex information, expectation and coordination problems gained such prominence in the 'new' economy?

First, the microelectronic/digital and net-based IT technologies that have come to dominate the technological basis of the modern economy have considerable informational and organizational impacts. With the nearly unlimited opportunities to produce, reproduce, process, store, diffuse and communicate data, the selection and use of information has become a major problem for effective choice and behaviour.

Further, electronic information has come to be available at near-to-zero marginal costs, adding yet another aspect to the collective-good character of information. This exists regardless of the fact that information is often used invidiously, i.e. opportunistic use of 'asymmetric' information.

Finally, given the ubiquitous external effects of the production of new information, and of technological decisions in general, in a world of microelectronic/digital net-technologies, any innovation is basically a collective good. This implies there is a considerable potential of technological lock-in among individual agents – if coordinated through 'markets' and prices alone.

Second, in a world of increasingly fragmented value-added chains and, consequently, ubiquitous technological complementarities, externalities of individual decisions have been pushed to the fore. Corresponding 'systemic' innovations become complex decision-problems that also call for more adequate forms of coordination than the 'market' can provide.

Specifically, manufacturing and services, also more fragmented, have become a directly interdependent relation that involves the problems of information-externalities, too. Service requires a collective-learning process together with manufacturing agents, and in this way other forms of coordination than those associated with 'markets' and prices.

Third, globalization is a politically and bureaucratically organized phenomenon involving regulated 'deregulations' influenced by the 'neo-liberal' 'market' ideologies. However, a global de-regulated economy cannot be, and indeed has demonstrated not to be workable in terms of societal effectiveness of different kinds – and in some respect even for the most advanced and powerful corporate agents. It has supported functional and spatial fragmentation and social/institutional disembedding to such an extent that the associated increase in turbulence and uncertainty cannot be effectively managed by even the most powerful individual agents through 'markets' alone. Hence, this situation, too, is characterized by new coordination requirements.

Common to these new dimensions of the 'new' economy is the fact that the requirement of more adequate coordination has become predominant and ubiquitous. Additional requirements for governing increased complexity implies 'markets' have become ever less effective institutional arrangements. Their capacity to manage complexity has decreased as 'neo-liberal' deregulation and disembedding have increased.

Subsequently, we experience widespread uncertainties, latent collective blockages of action and technological lock-ins, i.e. social costs in a wide understanding.

Complexity here not only depends on the number of agents involved and of relations among them, but also on the potential multiplicity of relations between

the same agents. Given the fact that collective-good and social dilemma problems have become ubiquitous, it can be demonstrated that different ways of behaviour are possible that allow for different types of relations among the agents. Here the outcome is basically undetermined, involving high levels of uncertainty at the outset. Specifically this dimension of complexity and increased uncertainty requires new forms of coordination in order to make effective outcomes feasible.

Given the currently increasing lack of social/institutional embeddedness on the global level, through which the conventional states had provided a minimum structure and stability of expectations and behaviour, the corporate economy worldwide is pursuing its own private substitute practices and arrangements to reduce complexity and uncertainty.

These mainly include monopolistic practices and power exertion, collusion, lobbying for protection and other hierarchical and restrictive practices. These mainly include hub & spoke-type networks. Indeed, they are increasingly protected and supported by states, politicians and public administration through standard-setting (heavily influenced by national and international 'champions'), permissive collusion-policies and the legislation of high 'intellectual property rights' fences. This entails considerable shortcomings of the socio-economic development.

However, those substitute practices also include local clustering. In other words, corporate agents are spontaneously developing new spatial forms of coordination among suppliers, assemblers, customers, service providers, public and intermediate agents, etc., based on proximity, with emergent local social/institutional embeddedness and long-run, recurrent and increasingly trust-based interaction and collective learning. In this way, even individualist practices, in a best-case development, may include more progressive attempts at coordination through non-restrictive, non-exclusive, i.e. predominantly innovative and future-bound, forms of clustering, and emergent cooperation.

It can be demonstrated that evolutionary processes in such more progressive arrangements that allow for collective learning among individual agents may lead to the emergence of rule-based coordinated behaviour, i.e. institutions of coordination and cooperation. Coordination then indeed may become possible in situations of direct interdependencies.

However, such ideal learned coordination through institutionalization is highly conditional with respect to its very emergence and sustainability. Emergence may be highly time-consuming, and existence may be fragile and prone to setbacks caused by changes in external circumstances or internal structure and path-dependent dynamics – if not blocked at all.

Therefore, more public rationality rather than private hierarchical rationality may be required, specifically a new policy approach which is oriented to the process and outcome of recurrent interactions and collective learning among individual agents. It is important for creating conditions in which private agents can collectively learn new, effective forms of coordination. This policy approach is oriented to promote the emergence of institutions of cooperation. Hence, it may be viewed as institutional policy. This altogether makes up for a new hybrid coordination arrangement.

Throughout his life, Karl William Kapp had strived to develop an analysis of the social costs of non-embedded forms of the economy and to analyse institutional arrangements that facilitate social embeddedness and sustainability. It is along these lines that an analysis of the costs of the global 'market economy', and of more adequate institutional arrangements – including new policies to facilitate their emergence – can be rendered more sustainable and better able to cope with the increasing complexity. Complexity is nothing more than a reflection of overall socialization of production and innovation in the widest sense. Only more adequate, i.e. specific hybrid institutional arrangements can enable agents to adopt a long-run and collective, i.e. sufficiently coordinated, perspective so that they can deal with the ubiquitous externalities and collective goods and to avoid otherwise predominant social costs.

The new features of the 'new' economy force us to reconsider these issues and to re(de)fine our conceptions. It is towards this end that the present paper is designed to contribute.

In the following, current tendencies towards increased complexities in the fields of globalization and local clustering, fragmentation of value-added chains, interactive services, digitized/microelectronic information, net-technologies and 'systemic' innovation are first considered. They include the increasing problems of direct interdependencies, externalities, collective goods, social dilemmas and strong uncertainty, the increasing probability of co-ordination failures and collective (technological) lock-ins in the framework of 'market' arrangements, and the requirement for more adequate coordination mechanisms. Next, we will consider private substitute strategies and arrangements to reduce complexity. Some conditions of the emergence and governance of sustainably innovative network forms of coordination are then discussed. It will turn out that they require, finally, a new policy approach to set conditions to initiate, accelerate and stabilize processes of collective learning and emergent institutions which, in total, forms a then hybrid coordination and cooperation arrangement.

Increasing complexity

De-regulated globalization, social disembedding, 'institutional disequilibrium' and uncertainty

Globalization is a long-run political and administrative project, regulated, on the one hand, by highly selective strategies of deregulation, liberalization and empowerment of capital and corporate concerns and, on the other hand, by national regulation, bureaucratization and authoritarian control of more general societal groups and concerns. The speakers of the transnational corporate economy, therefore, have always welcomed this one-sided, 'neo-liberal' construction of a socially exclusive global space (see e.g. Rupert 2000; Elsner 2000: 412; Perraton 2001: 678; Elsner 2003).

This global layer of exclusive activities thus has become highly disembedded from the historically developed arrangements of social institutions that used to

exist in the framework of the nation-states and in national and regional-local cultures. The 'neo-liberal' construction of the global space has deliberately removed any form of social control (see e.g. Altvater, Mahnkopf 1997: 112; Kingston 2000). It has become a system of social and spatial fragmentation and polarization that escalates inequality, unequal power and 'foot-loosening' of capital interests. As such, it is unable to act as an effective mechanism for civilizing the self-interest of the most powerful economic agents (Markusen 1996a; Amin 1999; Standing 2001; Biswas 2002). The 'neo-liberal' global system, thus, can be called a system in institutional disequilibrium, entailing an excess demand of international public goods through a decrease in their supply while demand is increasing (Padoan 2001).

However, being under-socialized, the global economy does not provide enough 'structure' to the environment in which decisions are made, even for the most powerful individual corporate agents. Hence, the corporate economy, being insufficiently coordinated by the 'market', cannot effectively deal with the increased level of uncertainty and turbulence that this very global deregulation involves (P. Armstrong 2001: 533 ff., 541 ff.). As a result, not only social insecurity (Standing 2001: 173 ff.) but also economic instability and transaction costs have increased. Subsequently, the global system thus has become a power-based redistribution mechanism rather than a mechanism for comprehensive and sustainable innovation and capacity enhancement.

Increased uncertainty, instability and insecurity are largely dysfunctional and counter-productive for effective individual economic decision-making. Thus, private agents have increased their demand for certainty and stability of future expectations, in order to reduce complexity and transaction (information) costs. These, in turn, can be provided only through an increased level and new way of coordination.

Fragmentation of the value-added chains, technological complementarities and interactive services

This very globalization has increased momentum towards vertical disintegration of value-added chains and the redefinition of the boundaries of corporate organization in an effort to reduce labour costs and to control an enhanced labour force world-wide. By doing so, globalization again unleashes unprecedented dynamics of income and wealth redistribution. It began with the outsourcing of services and low-skill physical production; now we are on the way towards virtual corporations that manage 'layered systems of contractual relations' (Raikes *et al*. 2000: 392 ff.). Functionally fragmented value-added chains have been also spatially fragmented by selecting labour and suppliers at optimal locations on a global level.

Coordination, therefore, had to extend beyond international trade forms that had been governed by conventional 'market' relations. Again, international restructuring, thus, is at least as much a struggle over the problem of uncertainty and transaction costs through new forms of coordination as it has been the provision

of cheaper labour and resources and new outlet opportunities (Ruigrok, van Tulder 1995; Raikes *et al*.). Global integration, first of all, has involved individualistic, power-led solutions on an unequal, hierarchical and authoritarian basis; this is what the transnational corporation and its centralized supplier networks stand for (see e.g. Jones 2000; Lazonick, O'Sullivan 2000).

Specifically, manufacturing and services have become directly interdependent. Production and innovation systems, under conditions of functionally and spatially fragmented chains, have come to be highly complementary systems. This is one reason why innovations are systemic, directly interactive and collective-learning processes between manufacturing and services, again requiring a high level of coordination (Delaunay, Gadrey 1992: 120 ff.; Rabach, Kim 1994; Antonelli 1998; Bennett *et al*. 2000).

In addition, information and technological solutions are increasingly user-, problem- and context-specific and tacit and so have to be developed (and learned) in a dense common interaction process. Externalities thus are ubiquitous also in these processes. Therefore, manufacturing-oriented service can be effectively provided only through a common learning process. This is not feasible unless governed by a coordination mechanism that can deal with complexity (Ruys 2000).

Digital microelectronic information, net-technologies, net-externalities and market failure

The 'new' economy is characterized by digital, microelectronic and net-based technologies. This has considerable impacts for the agents' decision-making, for policy-making, and for economic theorizing.

Information – in contrast to a conventional manufactured good – has the features of a collective good since information has always been characterized by non-rivalry in consumption. Regardless of the fact that generating and exploiting asymmetric information is a dominant and 'rational' opportunistic strategy in an individualist society, joint use (joint consumption) of information is welfare-enhancing, and even a basic necessity for nearly all kinds of social coordination (i.e. for the very existence of the socio-economy).

We are speaking of information, knowledge production and innovation in the widest sense here, i.e. making problem-solving collective action feasible.

It is well-known in conventional economic theory that the total utility of a collective good is the sum of the utilities of the individuals. In a similar way, the total societal benefit of information increases with the number of its users, independent of the individual(istic) preferability of asymmetric information. Information is 'complementary', or 'systemic', and often 'technically' indivisible, and it is normally collectively generated from billions of acts of behaviour and learning (Gallaway, Kinnear 2002).

But digital microelectronic technologies have added another characteristic to the collective-good property of information: the opportunities to produce, process, store, reproduce and disseminate information have drastically increased so that

the production of information takes place at near-to-zero marginal costs and ever-decreasing average costs, i.e. increasing returns or economies of scale. The 'new' economy thus has entered a stage of informational abundance which bears little resemblance to the conventional mainstream economic assumption of scarcity.

Further, the technical facilities for reproducing information have become so widespread that they are no longer under control of the original producer. Legal as well as semi- or non-legal ways of reproducing information are taking place at near-to-zero marginal costs. Consequently, digital microelectronic information has virtually become subject to non-exclusion, rendering information a full-fledged collective good subject to free-riders (Gallaway, Kinnear 2002).

Finally, digital microelectronics have made possible tele-information and -communication. The technological basis of the economy thus is a net-based system. The presence of competing technology suppliers with different technologies supplied means that problems of technical standards, interfaces and protocols have become ubiquitous. This can be seen as a core problem of the 'new' economy. 'Net-externalities' and direct interdependencies of technological decisions have come to govern the dynamics of the economy – and neoclassical 'efficient' individualist maximising the near-to-irrelevant exception (if it ever was a relevant reflection of real socio-economies) (Hutter 2001).

Since net-based technologies are the more useful the more agents are in the net, each agent with his technological decisions generates external effects on others. With ever-accelerating innovation and ever-competing (and initially not standardized) technologies of different supplier groups, uncertain, reluctant and passive, or even completely blocked users have become a ubiquitous latent feature of the economy (Tirole 1995: Ch. 10.6; Wettengl 1999). The introduction of colour TV, video-systems, high-definition TV and computer operating systems are instances from the recent industrial history that demonstrated the latent ubiquity of initial and latent collective blockages or impeded dissemination of new technologies. In other words, there are pervasive tendencies in the 'markets' to generate innovation at 'sub-optimal' levels and of complete mutual blockages of action among interdependent individual decision-makers (Miller 2001). Huge international private-public bureaucracies have been established to care for technological standard-setting and transfer protocols to prevent latent blockages to become actual (Piore 2001; Weitzel, Westarp 2002).

Generally, direct interdependencies among individual decision-makers cause strong uncertainty among agents, which is 'strategic' in the sense that, with new suppliers, customers and competitors, the individual agent cannot know at the outset, nor calculate with a given (subjective) probability, which 'strategic' choice the other agents will take (Dequech 1997; 2001: 919 ff.). We are talking of 'Knightian' uncertainty where no probability distribution for expected states does exist and thus no risk calculation is feasible.

The ubiquitous abundance of information as largely non-exclusive collective goods reduces the producers' abilities to collect revenue in the 'markets'. Their efforts to change conditions in order to gain and increase profits, in turn, endanger a continued process of social generation of new information, knowledge, intellectual

and cultural material under commercial conditions; here we may have to face 'the simple choice between profits and production' (Gallaway, Kinnear 2002: 443).

A culture of enforced individualist self-interest tends to impede the creation of rich and varied information and intellectual and cultural products. For instance, it supports non-legal distribution and reproduction of information that threaten information and cultural production as such. Even the often-praised internet communities, if not well-coordinated around an open-source supplier and around the common task to further develop this product in their own interest (to receive more effective software or to become famous with one's own contributions) (Raymond 2001; see below for the 'Linux-paradigm'), are in fact anonymous 'anti-communities' that 'only facilitate the free-rider problem' (Gallaway, Kinnear 2002: 445).

Generally, the very concept of private property and the 'market' becomes increasingly strained with abundant collective-good information and ubiquitous externalities and uncertainty. 'The need, then, is for new institutions ...' (op. cit.: 446). 'The limitations of information as a commodity now have come to the fore, both in economic analysis and in policy matters' and call for a 'thoroughgoing innovation in organizational design' to include 'a very high level of collaboration' (Lamberton 2001: 115, 117–118).

It is common knowledge in (industrial) economics that the net-based technological foundation of the economy causes coordination problems and renders economics as a science of (the collective learning of) adequate coordination mechanisms rather than a science of individual maximization, general equilibrium and 'optimality' (see e.g. J. Friedman 1994; Tirole 1995; Lamberton 2001). Specifically, the digital and net-based economy 'is social, dynamic, and multiple' and the use of digital and net-based technologies can only be comprehended as being 'varied, embedded, and emergent' (Orlikowski, Iacono 2000: 358, 360). This is to deal with increased complexity.

'Systemic' innovations and collective lock-ins

All in all, this boils down to two facts:

- Under conditions of microelectronic, digital and net-based technologies and of fragmented value-added chains, industrial production (physical as well as service and knowledge) and particularly industrial innovation have become complementary, 'systemic' and societal, and in most cases display the features of collective goods, or social dilemmas. Here, individual agents have to cooperate in order to generate a superior collective outcome, but at the same time have individualistic incentives to gain an extra one-shot profit by exploiting others, namely if these contribute to the collective outcome.
- Coordinated and cooperative behaviour and common social learning are required and cannot adequately be organized by 'markets' alone; markets can neither carry the necessary information nor generate or sustain the kinds of behaviour and future expectations required for this purpose.

Deregulated and socially/institutionally disembedded 'markets' are even more incapable to meet the requirements of coordination. Deregulated 'markets' are prone to failure and are only capable of applying and generalising the very characteristics of 'markets', i.e. power exertion, hierarchy and exclusive collusion, to try to meet those requirements of the 'new' economy (Block 2000: 55 ff.).

Net-based technologies and 'systemic' innovations with their complex technological, and organizational, complementarities make collective lock-ins a ubiquitous feature of the production, use and innovation of high-tech goods. Innovation processes have always been complex systems of technical interrelatedness, scale economies, network externalities, real- and human-capital interrelatedness and social learning processes with institutionalized coordinated behaviour adapted to a specific technology may be obviously inferior. Processes within complex systems are highly conditional and path-dependent and may easily result in lock-ins with sub-optimal technologies or spatial segregation of different technological paradigms (for the famous QWERTY case study and path-dependent technological processes, see Arthur *et al*. 1985; David 1985; Redding 1996; David *et al*. 1998; de la Mothe, Paquet 1999).

Against this increasingly problematic background, individual agents within the global corporate economy cannot but develop individualistic substitute 'solutions'. These largely apply and reinforce the 'market' principles of power exertion, exclusive collusion and hierarchy.

Spontaneous individualistic substitutes to cope with complexity

The global corporate economy, in order to establish solutions to compensate for the failures of, and the strong uncertainty in, disembedded 'markets', primarily deploys conventional and restrictive forms such as the use of market power and monopolistic positions, hierarchy, collusion, political lobbying for high 'intellectual property rights' and for international public-private standard-setting.

However, it also partly develops new kinds of coordination such as clustering and networking. These new forms acknowledge the fact that the collective, societal character of production and innovation is continuously coming to the fore.

Power exertion, collusion and 'intellectual-property-rights' fences

The features of the 'new' economy provide, through falling marginal costs and increasing returns of scale and scope, a textbook case of 'natural' monopoly. Companies indeed strive for monopoly positions, and often are able to keep them with their abilities to exert increased power, including political lobbying for financial support and legal protection. This 'solution' may be called the Microsoft paradigm, where an individual corporation commands 80 per cent or more of a relevant core technology and demand – and indeed receives the status of a national and transnational champion. It is well-known how Microsoft has been, and still

is, using its power vis-à-vis customers, suppliers and competitors – and how it also is able to use its power for lobbying and for the generalization of its technical standards on national as well as global levels, and for organising large hub & spoke structures throughout the global economy (Weitzel, Westarp 2002).

In most cases, however, 'markets' do not display monopolistic features in the narrow sense but have well-established oligopolistic structures. These structures do not so much involve conventional collusion but new and more comprehensive national and global public-private networks and organizations of alignment of interests and formal standard-setting.

As has been widely analysed, the new digital and net technologies have indeed pushed 'power-ization' and centralization throughout the 'new' markets (for telecommunication, Miller 2001; for the mass media industries, Champlin, Knoedler 2002).

Risk-taking, on the other hand, and in contrast to its everyday mythology, is not so much a feature of the 'new' economy (P. Armstrong 2001: 540 ff.).

Notably, monopolization, oligopolization and public-private collusion (in high-tech and knowledge-based areas) become feasible and emerge with the artificial 'construction of scarcity' of information (May 2002: 125 ff.) which otherwise would be a public good and largely available for free (Gallaway, Kinnear 2002).

The corporate 'new' economy, thus, is an enforced power structure 'that is increasingly at odds with technological reality' (Gallaway, Kinnear 2002: 446) which calls for basically new institutional arrangements.

Local clustering

Local clustering in 'new industrial districts', with repeated and long-termed interactions with nearby located suppliers, service providers, project partners, customers and diverse societal and public agents, has become the new kind of spatial organization of the most advanced, most high-tech, most globalized and most fragmented parts of the corporate economy (Maillat, Lecoq 1992; Feser 1998; Maillat, Grosjean 1999; Elsner 2000). In spite of its rhetoric of globalization, the corporate economy is striving for spatial and cultural proximity, some form of local re-embedding, structure, certainty and collectivity (Gertler *et al*. 2000; P. Armstrong 2001: 545 ff.).

Local/regional clusters are 'functional' systems of supplier relations that stem from, and still largely adhere to, 'markets' in the sense that they emerge from spontaneous, volatile arm's-length spot-relations (for the definition of clusters and networks, Elsner 2000: 413 ff.). At the same time they may outgrow the 'market' in the sense that they establish more stable and frequent (i.e. 'dense') supplier interactions. They partly drop 'market' relations as supplier relations get increasingly resistant against competitive low-price offers from outside the cluster. This is because their repeated and stable interactions lead them to develop parallel and similar ways of thinking, planning, expecting and behaving, better knowledge of each other – and in this way more stable expectations about the others' future actions, and mutual trust. Those stable expectations and mutual trust outweigh

low-price offers from outside. This would be an ideal process and desirable problem-solving outcome under individualistic conditions.

Without necessarily being conscious about this, the agents thus enter processes of collective learning of correlated behaviour that coordinates them in a new non-'market' way and helps them solve collective dilemma-prone decision problems more effectively (for a more detailed discussion of clusters, their mechanisms and effects, see Lazonick 1993; Ciccone, Hall 1996; Steiner 1998; Dupuy, Torre 1998; Rallet, Torre 1998; Keeble, Wilkinson 1999; Elsner 2000; Lambooy 2000; Torre, Gilly 2000; Breschi, Malerba 2001).

Although clusters can be viewed as coordination mechanisms of their own, they are only a first step beyond 'market' coordination and towards more effective kinds of coordination. Indeed, being 'functional' systems they are the basis for more consciously developed kinds of coordination, i.e. ('strategic') networks, formed by some set of firms in the cluster.

However, coordination mechanisms characteristic of 'markets', i.e. hierarchy and differential power, enter and eventually also 'conquer' clusters. So the effectiveness and innovativeness of clusters should not be overestimated. Global corporate players often organize local clusters in a hierarchical and power-based way around their local subsidiaries and often use them only in their very specific interests (Ruigrok, van Tulder 1995; Markusen 1996b).

Global corporate agents organize global and fragmented value chains, on global as well as local levels, on the basis of power, hierarchy and centralization. They control information flows, entry and access to resources and play their roles as key agents, independent of the specific spatial level, following the postulates of large-scale production under the conditions of global fragmentation of resources, production chains and customers (Lazonick 1993; Ruigrok, van Tulder 1995; Swyngedouw 1997; Oinas 2000; Raikes *et al*. 2000: 392 ff.).

Against this backdrop, local clusters adopt structures of hub & spoke, satellite platforms and other hierarchical and power-based types (Markusen 1996b). They may be highly innovative as long and as far as the powerful hub, e.g. a global player branch office forces innovations in its own interest. However, the more power-based, hierarchical and hub-centered the cluster structure the more the cluster will become risky, subject to life cycles, i.e. long-run downward swings, and accelerating its own ageing (Tichy 1998).

Networking, hub & spoke type

The corporate agents in the 'new' economy tend to surmount arm's-length and one-shot spot 'market' relations even more through networks. Networks are more or less consciously contracted and project-based multilateral mid-term oriented coordination mechanisms (for a definition, see Elsner 2000). Networks thus may more effectively assist in the solution of collective problems, promotion of collective learning and institutions-building.

Nevertheless, these private solutions still remain insufficient substitutes. As mixtures of progressive and regressive structures and practices, they indicate

the necessity of more comprehensive societal forms of organization and management. Being private, individualistic and thus power-based solutions, unregulated networks also display tendencies towards exclusion, collusion and hampering sustainable innovation, through their very hub & spoke structure and in the course of their life-cycle (Baker 1996; Pratt 1997; M. Armstrong 1998; Tichy 1998; Lamberton 2001). Thus, even intensely networking 'new industrial spaces' are as much 'symptoms of crisis' as they are problem solutions (Baker 1996).

Consequently, solving the problems of increased complexity through organizational innovation that allow for more sustainable coordination and cooperation, requires a set of characteristics of the structure and governance of networks – and a public policy framework to support private agents in their collective learning of equally based, open and sustainably innovative institutionalized co ordination.

Networking, open-source type – conditions of progressive network-governance and sustainable cooperation

However, other network types have inspired recently fashionable contentions about the possibility of self-organized and self-sustained cooperation and networks. The success – and the myths – of the Silicon Valley, for instance, have often been ascribed to its superior capability of network self-organization (Saxenian 1994). Specifically, the Hewlett-Packard corporation has been said to have established open-source practices from the early fifties on. Following the Hewlett-Packard paradigm, many successful networks in Silicon Valley have internally practiced open-source policies and cross-licensing among the network-participants. By doing so, they have established a less restrictive network where new knowledge is openly communicated, taking into account that in this way outsiders might get easy access to it. The condition of their success, thus, was not so much informational restriction but the speeding-up of innovation that they yielded through an open and trust-based network culture. This culture includes that even the smallest and most specialized sub-supplier is not only equally integrated in the information flow but becomes part of a system of parallel and interconnected technological innovations and future expectations of all network members. Competitors who should try the strategy of reverse-engineering and imitation would be shaken off by the superior speed of innovation.

Nevertheless, a closer look at this paradigm indicates that these networks are also more or less hub & spoke type, well-organized by and around powerful big assembler corporations. (In addition, their success is also based on the fact that the complete Silicon Valley public higher education, public services and private-public-networks infrastructures have been designed to produce positive external effects to these networks.) We have discussed the shortcomings of those network structures and the risks for their abilities to innovate and life-cycles above.

A more recent, and perhaps more promising network type has been established as what we may call the Linux paradigm. Its core philosophy is a more radical open-source strategy vis-à-vis the whole internet public, more decentralization and a low level of power exertion by the hubs (McKelvey 2001; Raymond 2001).

Indeed, the strongly decentralized, largely public and highly communicative, even near-to-ideally anarchistic processes have brought to the fore one of the big success stories of the digital economy with an unprecedented and sustainably high speed and high quality of innovation which obviously exceeds that of the main competitive system from the large and powerful hierarchical structure of Microsoft, i.e. the MS-DOS/Windows operating system. And, indeed, after Netscape has moved over to an open-source strategy more recently, Microsoft seems to be one of the last big names of the digital economy that still have not.

Indeed, a core finding of 'hackerdom' and recent radical open-source networking is that low levels of power and hierarchy, open information flows and non-exclusion are network properties favourable to a culture of effective learning of coordination and, subsequently, to fast and sustainable innovation (on the economics of information openness, see Foray 1998; Cohendet *et al*. 2001; McKelvey 2001; Raymond 2001).

A large amount of literature has recently developed general progressive network properties and 'good governance' principles or rules favourable to effective collective action so that networks neither adopt a restrictive/collusive character nor become vulnerable to sharp external changes nor prematurely age and become retarding rather than innovative (Ostrom 1990). Among these principles are openness, guaranteed and continuous entry, diversity/pluriformity of agents, low levels of hierarchy and power, parallel and redundant actions among network participants, the exertion of the voice mechanism irrespective of differences of size and power of participants, reciprocity, and others, i.e. superimposing network relations with collectively learned patterns of institutionalized behaviour, i.e. a network culture or ethics (for a more comprehensive discussion see Maggioni 1997: 238–49; Bogason 2000: 49 ff., 64 ff.; Elsner 2000: 450 ff.). These principles are all to ensure a continuous learning and institutionalization of cooperation and subsequently an abundance of systemic innovations (Powell Smith-Doerr 1994; de Bruijn, ten Heuvelhof 1995; Bond, Pyle 2001; Lazonick 2001; Rantisi 2002).

Sustainably effective, or 'dynamically efficient', networks of this kind could well be ineffective in the short-run, i.e. 'statically inefficient', specifically in the view of powerful individual agents (Maggioni 1997: 249).

However, as has been mentioned, in the reality of 'market' economies those ideal networks will sooner or later be dominated by powerful corporate agents.

Even the more recent hacker-culture and open-source networks have already been working in a more or less hub & spoke structure, too, with low levels of power and a greater decentralization and several smaller hubs, though. Private ownership of the open software is still well-defined and the process of innovative interaction, with the incentives of superior effectiveness of everyone involved, is well-organized by the hubs/owners (Cohendet *et al.* 2001; McKelvey 2001: 222 ff.; Raymond 2001: 73 ff.). In this way, it is a vulnerable structure, too (McKelvey).

With the recently further increased power of Microsoft and its intention to buy up Linux and the increasing dominance of conventional powerful industrial players

and state-administrative control in favor of commercial B2B- and B2C-interests in the internet, the future of largely self-regulated cooperations and networks in the internet has become highly questionable (Dolata 2002).

Finally, there is no guarantee with private-agents networks that the collective goods that networks are dealing with, and possibly are successful to generate, are confined to the limits of these networks. In contrary, it will normally be the case that those goods are socially, functionally and/or spatially more far-reaching than the boundaries of a private-agents network.

All in all, the societal character of production and innovation processes indeed requires an integration of networks in a larger and more public environment where the collective goods involved are evaluated from the perspective of its public relevance, and private networks may have to get interconnected, i.e. networks of networks established (Bogason 2000: 73 ff.; Elsner 2000: 435 ff.). Therefore, '(p)ublic policy should establish (...) an environment favorable to innovative behavior' (Maggioni 1997: 250) so that 'the development within collective public action will take place within networks based on hybrid forms of organization, backed to some degree by the public power' (Bogason 2000: 76 ff.; for 'multiple', i.e. hybrid forms of coordination consisting of some combination of networks with learned institutions, some form of moderate hierarchy and specific public policy intervention, see Baudry 1998).

Before we develop the argument of a hybrid coordination arrangement, however, we have to take a closer look at the complexity problem and the opportunity of self-sustained private coordination.

Complexity and coordination beyond the 'market' – the strength and weakness of 'self-sustained' evolutionary institutions-building

Complexity and the problem of coordination defined

Complexity is often defined as a system property based on a large number (N) of interrelated agents, i.e. with a large number of direct and mutual relations [R(N)] among them. Ideal 'markets' are not complex in the sense that they are conceptualized as systems of isolated individual agents with man-good relations only, specified through a (equilibrium) price vector that is determined and made public by an auctioneer instance. In ideal 'markets' there are no direct mutual agent-to-agent relations, i.e. no direct interdependence [R(N)=0]. Agents are only indirectly interdependent in the sense that the (equilibrium) price vector depends only on the quantities of supply and demand of all other agents taken together (see fig. 1). This removes complexity in a way that the formal logic of individual(istic) constrained maximization can be applied. It has been long-proven (but not really acknowledged by the bulk of the economic profession) that isolated maximizing individuals acting in the institutional framework of a system of 'markets' which are coordinated by an auctioneer instance do not provide sufficient structure to the

S_i: suppliers $i = 1, ..., n$; D_j: demands $j = 1, ..., m$; p_k^*: equilibrium prices $k = 1, ..., l$
q_k: quantities of supply and demand for k goods $k = 1, ..., l$

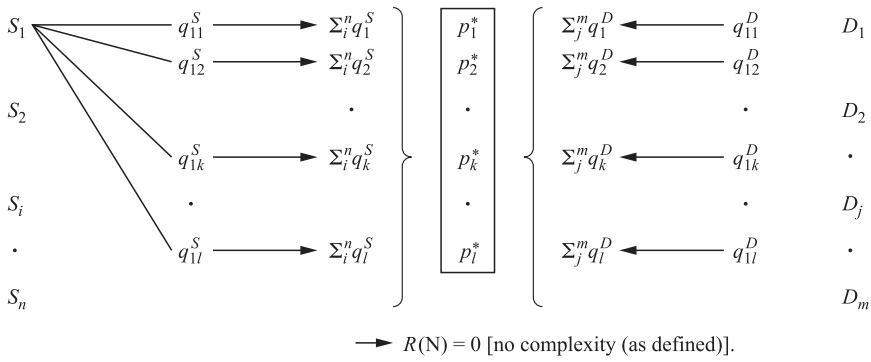

Figure 1 The Ideal World of 'Markets'

economic system to guarantee in any reasonable sense the 'optimality' of a 'market' economy, i.e. the uniqueness, let alone the stability, of a general equilibrium (the well-known 'Mirowski-debate' and its aftermath, see Mirowski 1989; for a recent comprehensive analysis of the impossibility of a unique and stable rigorous general 'market' equilibrium in a physical analogy, see Wellhoener 2002).

With the acknowledgement of the real-world direct interdependencies and direct interactions, specifically in the 'new' economy, complexity becomes an essential property of both the object and theory of economics (Dequech 2001: 912 ff.; Delorme 2001; Maggioni 1997; Rycroft, Kash 1999). This is related to the fact that R(N) becomes a drastically increasing function so that, with moderate numbers N already, real-world ('bounded rational') individual agents will come to perceive 'chaos' with the potential (maximum) number of relations R (see fig. 2).

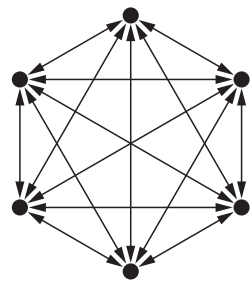

Figure 2 Direct Interdependencies (Lattices)

	C	NC
C	a, a	d, b
NC	b, d	c, c

with $b>a>c>d$ and (b + d)/2<a (in order to prevent alternating sequences of co-operation and defection being successful): $R_i = [R_1(C, C), R_2(NC, NC), R_3(NC, C), R_4(C, NC)]$ [complexity (as defined)].

Figure 3 Multiple Relations – The Example of a Prisoner's Dilemma (Collective Good) Structure

However, in view of collective-good problems and social dilemmas, real worlds of direct interdependence are complex in an additional dimension, i.e. multiple relations (R_i) among the same agents. This can be illustrated by a simple Prisoners' Dilemma (PD) structure with its different kinds of relations among the agents (see fig. 3). Individual agents are mutually interwoven in their decision-making, i.e. 'strategically' interdependent in a coordination problem to solve a collective-good problem, so that individual decision-making becomes an individualistically ('rationally') non-solvable problem. This again reflects the fact that the economy is indeed a socio-economy and all production and innovation have a collective, or societal character (for a discussion of a potential solution see below).

Complexity of the economy, specifically in the sense of multiple interrelations, is connected with conceptions of strong uncertainty of the individuals. As has been indicated earlier, this stems from the fact that an individual cannot know initially what action the other individual will take, while being aware, nevertheless, that the other's action will have an impact on his/her own welfare (Dequech 2001: 919; Delorme 2001, 84 ff.). There may be strategic choice, even creativity, novelty and surprise in real-world processes of direct interactions.

In addition, strong uncertainty may also reflect the fact that individuals have a relatively limited competence in processing information ('bounded rationality'). This may refer to limited information and knowledge capacity in a longitudinal section, i.e. limited memory about the ways the individual has interacted with other individuals in the past, as well as cross-sectional limited information processing capacity, i.e. limited attention or perception (in transparency) of current interactions among the other members of the socio-economy.

Since complexity of the real socio-economy is reflected in the way the economy survives by coordination of its individual parts, coordination is indeed the central problem of real socio-economies – and of economics as well (J. Friedman 1994). Economics thus has to become a science of organizational forms of social coordination.

Individualistic solutions to this problem do either not come into existence or are (Pareto-) inferior, respectively. In practice this is reflected by widespread

blockages of action and technological lock-ins on inferior technological paths (Block 2000).

And coordination here is indeed feasible only by way of collective complexity reduction. The conventional device to reduce complexity in real-world 'market' economies is hierarchy. However, this seems to be a very limited tool to coordinate agents in terms of equal base, non-collusion, good governance and sustainable innovation, in lattices of direct interdependencies. Such lattices of direct interdependencies, under the principle of hierarchy, are transformed into purely hierarchical hub & spoke networks (with $R(N) = N - 1$, i.e. reduced complexity) which may reduce complexity, but do so by implying the very shortcomings already discussed.

With collective-good, or social dilemma structures, respectively, however, i.e. with multiple relations, complexity reduction also has to do with the collectively learned reduction of the number of potential relations R_i. A more 'complex' complexity-reduction device, i.e. complexity reduction that deals with real-world direct interdependencies, thus is collectively learned institutions of superior coordination. Collective learning of institutions of coordination (or cooperation, respectively) then may lead to reduced complexity: $R_i \rightarrow R_1 = R(C,C)$.

This solution of the coordination problem in a complex socio-economy with mutual direct interdependencies is not reducible to isolated 'rational' decisions of the individuals (Delorme 2001: 102 ff.). It must be 'systemic' in the sense of establishing a super-individual context that prevents individuals from exclusively acting in their own self-interest.

This 'systemic', or societal, context has to do with recurrent interaction, learning and developing expectations (about others' actions and expectations), i.e. adaptation to the actions and expectations of the others, historical time, path-dependent evolution and futurity (Delorme 2001; Maggioni 1997).

This solution consists in the collective learning of a behaviour that habitually excludes or restricts the strive for short-run maximization of the individualist interest, i.e. the emergence of a social institution of coordination, even in view of continuing incentives to defect. The reconciliation among these 'mixed-typed' individual interests and incentives and the common restriction of a pure individualistic short-run maximizing behaviour is only conceivable in a process of emergence of a commonly accepted habitual rule, i.e. an institution (Dequech 2001: 922 ff.). Institutions work as inter- and super-individually existent social informational devices that inform the individual about what he/she normally can expect and what normally is expected from him/her.

This is nothing else than the very process of establishing social order and reducing complexity to a level where individuals can reasonably be expected to act effectively, i.e. to manage the existing, and now reduced, level of uncertainty and in this way to become capable, and inclined, to innovate in the widest sense, i.e. to establish problem-solving through future-bound behaviour. The individual is indeed incapable of being innovative in the face of a level of turbulence that is too high, and of strong uncertainty (for a conception similar to the idea of blockages and lock-ins, i.e. hysteresis, in relation to strong uncertainty, see Setterfield 1996).

Also 'markets' are workable in any reasonable sense only if well-embedded in institutional arrangements that care for a way of coordination adequate to manage the problems of ubiquitous externalities and collective goods (MacEwan 2000: Ch. 4: 'The Social Construction of Markets'; Elsner 2003).

Evolutionary game theory of social dilemmas, as has been indicated, provides some crucial insights to this problem. We will turn to this in the following.

The possibility, and improbability, of spontaneous coordination through institutional emergence

Many economic problems can be represented as coordination failures in institutional settings that we call 'markets', specifically if these are poorly 'embedded'. Coordination failure is reflected by the fact that many, or even most, of the individual private agents involved would be willing to act in a sustainably problem-solving manner, but cannot do so because they do not, and cannot know at the outset whether the others will do the same, i.e. act in a coordinated way in turn; and problem-solving here requires the common, coordinated and cooperative action of all agents. The agents are 'locked-in' or blocked.

The view that 'market' failure is a coordination failure has indeed been utilized to explain a wide range of phenomena related to economic decline (Cooper, John 1988; Ball, Romer 1991).

Coordination failure is represented in a social setting characterized by some 'system' properties that we will assume in the following:

- Direct interdependence of the individual agents, i.e. we assume a genuinely social situation.
- A social dilemma structure of the PD type (this implies the adoption of some neoclassical assumptions, such as individual agents maximising the expected utility, who are well-informed about the pay-off structure; this is to assume a largely individualistic culture at the outset – for the sake of a simple formal analysis).

The solution of the problem of coordination in complex decision settings generally depends on 'system' properties that we also will assume:

- Recurrent interactions, i.e. infinitely or indefinitely repeated direct interactions among the same agents.
- 'Sequentiality' of rounds of decision-making. This provides the opportunity for processes of collective learning and institutionalization.

A PD-setting seems to be a relatively simple setting which has, however, a relatively high explanatory power and empirical significance. It reflects why 'markets' fail – namely, in a rather individualistic culture – when there are direct interdependencies with strong uncertainty and 'collective-good' problems. (It should be noted, however, that the general argument is not dependent on the specific

form of a PD-dilemma. In technical terms our argument referring to coordination can also be developed on the basis of other social problem structures, such as a coordination game.)

As is well-known, 'market failure' is demonstrated in a one-round PD or a PD with a given and commonly known finite number of rounds. Here, a stable formal equilibrium exists which reflects a socially 'unsatisfactory' ('Pareto-inferior') situation. The collective good involved is inaccessible to the private agents in this individualistic setting ('market').

If we define a private good as a good that can self-sustainingly be produced by the interacting private agents in a process taking place in this 'market' setting, then the transition from a finite-round PD to a PD-supergame (i.e. a PD being repeated infinitely or in an indefinite number of rounds) implies the potential transition from a collective good (which was inaccessible to the private agents before) to a 'private' good.

A simple static solution of a two-persons/two-strategies PD-supergame with given pure strategies (cooperation and defection) is the following. The setting (i.e. the pay-off-matrix of two persons with two strategies) was given in fig. 3.

(It should also be noted that the argument developed below is independent of the formal attribute of the pay-offs being symmetric, as it will be given here. The symmetric pay-off structure is used for reasons of simplification.)

This implies that the incentive to defect (b) while the other one cooperates is greater than the incentive to commonly cooperate (a), which in turn is greater than the pay-off for common defection (c), which for its part is greater than the pay-off for cooperation (d) while the other one defects.

While cooperative agents will normally gain in an infinite series of rounds,

$$C/C = a + \delta a + \delta^2 a + \ldots = \frac{a}{1 - \delta} \tag{1}$$

(where δ is a common discount parameter for pay-offs in future interaction rounds), a defecting agent may gain, in an infinite series of rounds, as a maximum (i.e. while the other one cooperates in the first round and also defects from the second round on, i.e. the well-known tit-for-tat strategy)

$$NC/TFT = b + \delta c + \delta^2 c + \ldots = \frac{c}{1 - \delta} + b - c. \tag{2}$$

Cooperation pays if C/C > NC/TFT. This can be postulated as (1) – (2) > 0, or

$$\delta > (b - a) / (b - c). \tag{3}$$

This equation provides a simple formal condition for the superiority of cooperation over conventional individualistic (short-term, or one-round) maximizing behaviour (in a world with the fixed pure strategies mentioned). It shows

that the superiority of cooperation (for individualistic agents) depends on the relative pay-offs of (a), (b) and (c), i.e. the incentive structure, and on the discount parameter (δ), i.e. futurity. Its result indicates that if the discount parameter, i.e., the weight of future pay-offs, is greater than a certain combination of pay-offs [with (b – a) being the 'relative incentive to defect'], cooperation pays, while it is obviously unfavourable for cooperation if (b) and (c) are relatively high and (a) and (δ) are relatively low.

We will return to this result later to discuss some of the not-so-familiar policy implications involved.

However, because this result is based on a static, so-called single-shot, supergame with fixed, idealized strategies in which only one single calculation is to be made by the agents, and because we have, in contrast, assumed a process of sequential rounds, we additionally have to show how cooperation emerges in a process of interaction among agents who are assumed at the outset to be conventional individualistic short-term maximizers. This means showing how individuals may change their behaviour through (collective) learning (i.e. individual learning in direct interactions).

As a first step we might refer to the well-known Axelrod solution (Axelrod 1984, 1997), derived from his famous computer tournaments in which more than sixty strategies were submitted, and thus given, to be applied in PD-supergames. Axelrod employed a simple replicator mechanism with a differential increase for those agents whose strategies were scoring higher in two-strategy interactions (which were organized in great numbers). The well-known result showed the clear and stable superiority of the relatively simple cooperation strategy, 'tit for tat'. Axelrod has interpreted these results to show the emergence of the institution of cooperation in an evolutionary process. This may be justified insofar as we can interpret a replicator mechanism as a sort of learning process. It presumes the differential selection of the physical individuals. In formal models a new individual who assumes the place of an old one, who has ceased to exist, can also be interpreted as an individual who has changed his behaviour through learning.

Cultural evolution, however, requires an explicit conception of the search and learning mechanisms of individuals and social groups (Stein 1997). Needless to say, the object of differential selection and differential reproduction here is the type of (individual and collective) behaviour (i.e. the institution) and not the physical agent (Field 1994).

Schotter (1981) was one of the first to develop a model for sequential PD-supergames with stochastic search and learning. Today there are many approaches and models to formalize cultural-evolutionary processes which employ mechanisms of 'selection', 'crossing', 'mutation' and individual adaptation through learning (from one's own experience, through imitation, etc.). They formally show that, and under which conditions, cultural evolution – particularly in dilemma-prone (PD-)situations – involves the emergence of institutions of cooperation (Liebrand, Messick 1996; Fudenberg, Levine 1998; Kirman 1998; Offerman, Sonnemans 1998; Oltra, Schenk 1998).

A variety of approaches and models thus have demonstrated that, in a cultural-evolutionary process with emergent institutions, a social dilemma basically can be overcome through socially learned, institutionalized cooperation, and a collective good can be transformed into what we have defined as a private good. Many neoclassicists and 'neo-liberal' policy advisors hold great expectations for the evolutionary 'efficiency' of the 'market' economy with those results.

However, a problem that remains and which cannot be solved through individualistic rationality, is the continuing existence of the basic social-dilemma stemming from, and specifically prevailing in, a relatively individualistic 'market'-related culture. Indeed, the dilemma-prone structures always remain in the background. This implies that the process of solving a 'collective-good', or a social-dilemma problem through cooperation may be highly time-consuming and unstable. The more individualistic the culture is – or the stronger the dilemma-structure in terms of the relations of (a), (b), (c) and (δ) – the greater the incentive to defect, and specifically to deviate from an already-established institution of cooperation will be, and continue to be existent.

Social lab experiments and model simulations of evolutionary processes based on social coordination problems and related settings, namely repeated PD-supergames, have indeed illustrated that hundreds or thousands of interactions may be necessary to establish cooperation as a rule of behaviour, and that even then cooperation may be fragile and occasionally collapse, sometimes because of mere small external changes or some internal dynamics in more complex models (Lindgren 1997).

'Neo-liberal' 'market' conceptions seem to have made cultures more individualistic in reality in the past decades (compared to the traditional 'welfare-state' socio-economies). Thus the individualistic behaviour represented in a PD, and assumed at the outset, can indeed be taken as a 'worst-case' view of a reality in which the social institutions of cooperation, social cohesion and social solidarity have in fact become weakened. It can indeed be argued

> that neoclassical rationality (...) is simply a case in which the parental bent is severely repressed. It could arise only in a cultural setting in which there is a trained incapacity to consider the impact of one's behavior on the well-being of others.
>
> (Bush 1999: 139)

It thus seems to be necessary to introduce a supra-individualistic rationality, or mechanism, into this process – specifically, a public-policy intervention in order to either initiate (i.e. de-block, un-lock) or accelerate and stabilize the process of institutionalization of cooperation, which cannot be brought forth with sufficient certainty, speed and stability by individualistic rationality alone. Private power exertion and hierarchy as organising media would entail a high vulnerability to the whole process and its outcome. Consequently, there is a fundamental requirement, and also ample opportunity, for a more comprehensive and collective solution, i.e. an active policy role to be played vis-à-vis the system of private agents.

Thus, we are talking about multiple, or hybrid institutional arrangements to reduce complexity and solve the coordination problem – not so much as combinations of institutions of cooperation (ethics, cultures) and hierarchy, normally introduced into those lattices by private agents, but of institutions and a public rationality.

An institutional policy to promote a hybrid coordination arrangement

The 'collective good' problem has conventionally been regarded in economics as implying a purely public task, but in an era of 'neo-liberal' policy conceptions it is dealt with by trying to empower the 'markets' and enlarge their ranges through 'de-regulation', 'privatization', 'transaction-costs' reduction, etc. This in fact tends to worsen the problem of provision of these goods by rendering the latent coordination-failure problem ubiquitous. Concurrently, collective goods have been ever more dismissed from the area of public responsibility.

The conventional 'collective good' view, however, unduly directed collective responsibility away from the private agents, who indeed have – in pursuing their own individual economic interests – considerable interests, as well as some opportunities, to contribute to solving the problem of provision of the collective good, as the simple PD structure makes obvious.

Our arguments have been leading to design a relation between the private interaction process and the emergent institution of coordination and cooperation on the one hand, and specific economic policy instruments aimed at promoting the emergence of this institution on the other.

Meritorization

First, we assume that the potential outcome of that private process can be related to a public, i.e. economic-policy objective in such a way that it can be subject to social valuation or meritorization. In other words, we assume that the private agents could indeed collectively produce a 'good' (rather than a 'bad') that has a potential public value in addition to its private values.

The conception of the merit good has been revived since the Eighties, and its basis has been slightly changed, even by Richard Musgrave himself (Brennan, Lomasky 1982; Musgrave 1987), from one of 'wrong individual preferences' to one that substantiates meritorization (i.e. a positive public valuation) on the basis of 'community preferences' that have evolved from historical processes of interaction outside of the 'market' (Musgrave 1987: 452). This implies a social evaluation of the outcomes of the 'market' through some kind of a social decision-making broader than, independent of and superior to that 'market'.

For our purpose we will define a merit good as a good basically resulting from the spontaneous private 'market' interaction process described above (i.e. a 'private good' as defined above), which is evaluated through a social decision-making process on the grounds of its complete lack, or deficient quantity, quality, relative

price, and – as new dimensions discovered in deficient individualistic interaction processes explained above – the time span needed for its production as well as the certainty or stability of acquiring it through that private interaction (Ver Eecke 1998).

Against this background, one of the problems of (mainstream) economics is that it has not developed a conception of social decision-making that could be regarded as an allocation mechanism independent of and with priority over the 'market', on the basis of which the results of the 'market' could be evaluated or '(de-)meritorized'. Neoclassical 'public choice', restricted to areas that the 'market' leaves aside, basically faces the same individualistic problems of social coordination that are faced in individualistic dilemmas in the 'market'.

Institutionalists, by contrast, claim that democratic and participative socio-political decision-making should continue to be relatively independent of the 'market' allocation and should have priority over it. The institutionalist conception of the negotiated economy has been elaborated to show that the 'market' has to be deliberately embedded in a wider socio-political process, and that this indeed is workable (Commons 1934/1990: 612 ff., 649 ff.; Nielsen, Pedersen 1988; Ramstad 1991; Nielsen 1992; Shipman 1999: 214 ff.).

We will not go into this discussion any deeper here (for a more detailed argument, Elsner 2001) but will simply assume an economic policy agent who is legitimized through a process of participatory democratic decision-making. In this very process, public policy objectives can be developed which provide the criteria for the 'meritorization' required.

Paradigm change towards a lean policy

One of the basic insights of the PD-structure is that the private agents receive a specific benefit from, and thus have a specific interest in, the production of the 'good' (which is originally 'collective', later possibly 'private' and finally becomes 'merit', as defined), regardless of the fact that it generally cannot be adequately produced by them (because of the coordination failure inherent in their spontaneous 'market' interactions). At the same time that the public agent is making the 'good' more accessible through 'meritorization' and through promoting coordination and cooperation, therefore, he/she can also call the private agents in to contribute to this social solution. In contrast to the conventional neoclassical collective goods analysis, we come to recognise that the state is by no means exclusively responsible for producing (and financing) this social solution. This implies a 'hybrid' kind of production. And, basically, it also allows for a leaner policy approach.

A related simple insight from the PD-analysis is that the 'collective good' problem, or the dilemma structure, is a gradual problem. This implies that the complete dissolution of the dilemma structure is not necessary for the public agent in order for the private agents to be able to solve the coordination or cooperation problem. For instance, if the public agent subsidises cooperative behaviour to such a degree that $a > b$ and $d > c$ in the above-mentioned PD, so that cooperation plainly

is more attractive to the private agents than non-cooperation, this might well be a costly public policy. (Besides, it would also be a trivial one in a theoretic sense.)

Social problem-solving, by contrast, can be promoted by gradually weakening the dilemma structure and in this way allowing for a more collective and coordinated culture. For instance, if the incentives for cooperation, i.e. the pay-off (a), are increased, the opportunity costs of cooperation [which are (b – a)] will be gradually reduced.

Let us take a numerical example, say $\delta = 0.9$, $b = 4$, $c = 2$ and $d = 1$. Equation (3) then implies that cooperation would already be superior to non-cooperation with $a = 2.2$ (instead of $a > 4$, as one would spontaneously assume, and where the PD-structure would be dissolved). (Above that a would have to be > 2.5 in order to meet the additional condition $(b+d)/2 < a$.) In other words, small rewards for cooperation may already be effective. More generally spoken, with gradual, relative changes in the incentive structure the probability of cooperative solutions may increase, i.e. cooperation will come into existence with increased probability, speed and stability.

Generally speaking, a leaner policy becomes feasible because the approach allows for a clearer definition of the relative interests, or benefits, as well as the relative responsibilities, or costs, of the private and public agents with respect to that cooperative social solution – as opposed to fuzzy 'public-private partnerships' so much en vogue currently.

Instruments

The public policy agent has to publicly identify the specific characteristics of the 'good' he/she wants the private agents to coordinatedly and cooperatively produce, i.e. the public objective or 'merit good'. But he/she must also establish incentives to enable the private agents to learn to cooperate in order to achieve the objective of common interest. For this purpose, he/she has to define the instruments to be employed in order to promote that coordination and cooperation.

Rewards of cooperation, pecuniary and non-pecuniary

The first complex of instruments is rather obvious; it aims at changing the incentives ('pay-offs') in order to increase the relative rewards for cooperation, or the opportunity costs of defection (or decrease the opportunity costs of cooperation and the net benefits from non-cooperation).

Equation (3), however, shows that the more successful the public agent is in integrating the private agents into a future-bound process – i.e. the higher the discount parameter δ is – the less the increase of the relative rewards for cooperation needs to be.

However, this trade-off between the rewards for cooperation (a) and the 'shadow' of the common future (δ) does not imply a contrast between quantitative (namely, pecuniary) and qualitative instruments, i.e. offering pecuniary subsidies as opposed to promoting the development of more favourable expectations among

the agents. The incentives from the public policy agent which reward cooperation may largely consist of non-pecuniary benefits (for this new policy approach, see Klein 1990; de Bruijn, ten Heuvelhof 1995: 173 ff.; Elsner 2001: 76–79).

Enlarging the 'shadow of the future'

The second complex of instruments indeed refers to the 'futurity' of the basic interactive process, i.e. the probability for private agents of meeting again. This can be related to the discount parameter. This can be interpreted not only as the weight allocated to a future pay-off but also as the probability of a future interaction among the same agents. According to equation (3), cooperation can be promoted if the discount parameter can be increased, i.e. if future interactions become more probable.

Although it does not seem to be so obvious, this condition for the success of the interactive process can also be subject to policy control. As Axelrod (1984) has already pointed out, the public agent can indeed increase the importance (i.e. the probability, or weight) of future interaction, e.g. by organising cooperation through frequent events, or make it permanent, e.g., by organising meetings with a greater frequency, dividing projects into several sub-interactions, connecting different projects so that the same agents will meet in different arenas and become aware of their common future, etc.

Obviously, there is some opportunity for the public agent to deliberately shape the conditions of interaction to promote cooperation in a variety of subject fields that private agents are jointly interested in (for practical innovation, industrial and regional policy applications, Hart 1992; Cossentino *et al*. (eds) 1996; Redding 1996: 468 ff.; Mizrahi 1998; Belussi 1999: 737 ff.; de la Mothe, Paquet 1999: 104 ff.; Rycroft, Kash 1999: 211–223; Elsner 2001).

Interactive and 'institutional' policy in the 'new' economy

To sum up, this new policy approach includes pecuniary and non-pecuniary incentives and specific framework measures, which help to change the behaviour and expectations of the private agents, considering the pay-off structure and time horizon of their interactions. This policy interacts with the interaction process of the private agents, i.e. interactive policy. It aims at supporting collective learning and institutions-building, i.e. institutional policy (for the basics of the 'institutional' approach to economic policy, Tool 1979; Hayden 1994; Elsner, Huffschmid 1994). It is also a case for publicly initiated institutional change (i.e. new institutions and related valuations emerging among private agents; Bush 1987) and deliberate 'institutional design' of framework conditions of interactions (Calvert 1995), all in the framework of a democratic and participatory negotiated economy conception with a publicly determined structure of merit goods. This policy approach, finally, with information, expectations and institutions-building at its core, has come to be widely developed and advocated recently (Klein 1990; Jessop 1994; Amable, Petit 1996; Block 2000; Yu 2000; Elsner 2001; Bryant 2001).

Conclusion: a hybrid coordination arrangement

Most of the specific features of the 'New' economy contribute to a considerably increased complexity, with direct interdependencies, ubiquitous externalities, collective-goods problems, and related social dilemmas in direct interaction processes. Innovations have become 'systemic', a fact that largely reflects the increasing collective character of all production. Problems of strong uncertainty vs. stable, trust-based future expectations about the kind of actions other individuals can be expected to take have come to the fore. The problem of coordination thus becomes the dominant issue of the modern economy, and therefore, also of innovation theory and economics in general.

The 'market' with its dominant price-based information and coordination is insufficient to solve those problems. This is increasingly the case since the 'markets' have been deregulated and in this way disembedded from their historically developed institutional structures, under 'neo-liberal' postulates. These institutional structures had provided for reduced uncertainty and stable expectations, collectively learned and habitualized cooperation, reduced complexity and a high level of coordination. In their absence, or reduced effectiveness, the 'market' loses its ability to yield sustainably high levels of innovation and basically becomes a mechanism of power-exertion and wealth redistribution rather than welfare enhancement.

In a world of deregulated globalization, social/institutional disembedding, fragmentation of value-added chains, locations and 'markets', net-technologies, complex interaction between manufacturing and services, and overall increasing direct interdependencies, the global corporate economy has developed different spontaneous private/individualist substitute coordination forms in order to reduce complexity to manageable levels. Among these are local clustering, networking, public-private organizations for standard setting, and more conventional collusive and power-based strategies. These, by and large, are incapable of adequately dealing with complexity and in this way to sustainably perform high degrees of innovation. Clusters and networks, specifically, turn out to be largely hub & spoke-typed and thus highly dependent on the interests of powerful hubs. Inadequately structured and governed networks tend to prematurely age, shorten their life cycles and lose innovative capacity.

Evolutionary game-theoretic arguing can be employed to analyse processes of direct interaction and to demonstrate that self-sustaining processes may indeed lead to the emergence of institutions of cooperation and coordination through collective learning. While this is basically feasible, it also tends to be highly time-consuming and fragile, if not blocked at all or locked-in on an inferior path.

A new 'hybrid' system of effective coordination, a 'New New Deal' (de la Mothe, Paquet 1999: 85; Storper, Salais 1997) with complexity reduction and 'good' (self-) governance of cooperative (network-) arrangements, including a new policy approach that refers to that interaction process of the private agents and shapes specific conditions to support that collective learning process of institutionalized

coordination, has been outlined here. The policy approach employs different kinds of incentives and shapes the expectations of the private agents about the future of their interaction so that the future counts high. This approach thus aims at collective learning and procedural, evolutionary trust- and institutions-building among private agents, i.e. 'institutional' policy. This hybrid system allows for the combination of strengths through a clear-cut allocation of responsibilities of private and public action and benefits of private and public agents. It thus also allows for a leaner policy.

This hybrid approach starts from an evolutionary conceptualization of interaction processes among private agents, with its potential of collectively learned institutions of coordination. It relates (through the societal merit-good structure), in a second, specific interaction process, an adequately shaped public policy system (with the public rationality and public objectives defined in a negotiated-economy framework) to that 'private' (and now meritorized) interaction system.

By overcoming the complexity challenge which is at the core of the modern 'new' economy, this hybrid approach of coordination mechanisms (i.e. social institutions plus public rationality) also allows for sustainable systemic innovation in the widest sense since it allows for the emergence of frequent, recurrent, repeated interactions in a more stable framework, for common collective learning processes, the stabilization of future expectations, the emergence of cultural proximity, common tacit knowledge and trust. In this way it may help to avoid ubiquitous actual or potential social costs of the 'new' economy.

Economic policy in the 'new' economy thus has to be interactive and process-oriented, information- and expectations-sensitive and coordination-oriented, future-bound and headed to institutions-building. It has to be 'institutional policy' with a transparent structure of the private and 'merit' dimensions of the goods of the economy, with transparent public objectives related to that societal goods structure, provided for in the framework of a participatory 'negotiated' economy. The new problems of the 'new' economy with its widespread cluster structures, network governance issues, high-speed and 'systemic' innovation requirements and, finally, new policy requirements can be fruitfully discussed in this framework, and solutions can be advanced.

Note

* The author is grateful to the discussants at the 2002 Trento conference on social costs. He also is grateful to the discussants of the 2002 annual conference of the iiso Institute for Institutional and Social-Economics, University of Bremen, for commenting on an earlier version of the paper.

References

Altvater, E. and Mahnkopf, B. (1997) *Grenzen der Globalisierung*, Muenster: Westfaelisches Dampfboot.

Amable, B. and Petit, P. (1996) 'New Scale and Scope for Industrial Policies in the 1990s', *International Review of Applied Economics* 10, 1: 23–41.

Amin, S. (1999) 'Capitalism, Imperialism, Globalization', in R.M. Chilcote (ed.) *The political economy of imperialism: critical appraisals*, Boston, Dordrecht London: Kluwer.

Antonelli, C. (1998) 'Localized technological change, new information technology and the knowledge-based economy: The European evidence', *Journal of Evolutionary Economics* 8: 177–98.

Armstrong, M. (1998) 'Network Interconnection in Telecommunications', *The Economic Journal* 108: 545–64.

Armstrong, P. (2001) 'Science, enterprise and profit: ideology in the knowledge-driven economy', *Economy and Society* 30, 4: 524–52.

Arthur, W. B., Ermoliev, Y. M. and Kaniovski, Y. M. (1985) 'Strong Laws for a Class of Path-Dependent Urn Processes', in *Proceedings of the International Conference on Stochastic Optimization,* Munich: Springer.

Axelrod, R. (1984) *The Evolution of Co-operation*, New York: Basic Books.

——. (1997) *The Complexity of Co-operation, Agent-Based Models of Competition and Collaboration*, Princeton, N.J.: Princeton University Press.

Baker, P. (1996) 'Spatial Outcomes of Capital Restructuring: 'New industrial spaces' as a symptom of crisis, not solution', *Review of Political Economy* 8: 263–78.

Ball, L. and Romer, D. (1991) 'Sticky Prices as Co-ordination Failure', *American Economic Review* 81, 3: 539–52.

Baudry, B. (1998) 'Trust in Inter-Firm Relations: Multiple Forms of Co-ordination', in N. Lazaric and E. Lorenz (eds) *Trust and Economic Learning*, Cheltenham, UK, and Northampton, MA, USA: Elgar.

Belussi, F. (1999) 'Policies for the development of knowledge-intensive local production systems', *Cambridge Journal of Economics* 23: 729–47.

Bennett, R. J., Bratton, W. A. and Robson, P. J. A. (2000) 'Business Advice: The Influence of Distance', *Regional Studies* 34: 813–28.

Biswas, S. (2002) 'W(h)ither the Nation-state? National and State Identity in the Face of Fragmentation and Globalization', *Global Society* 16: 175–98.

Block, F. (2000) 'Disorderly co-ordination: the limited capacities of states and markets', in L. Burlamaqui, A. Castro and H. J. Chang (eds) *Institutions and the Role of the State*, Cheltenham, UK, Northampton, MA, USA: Elgar.

Bogason, P. (2000) *Public Policy and Local Governance. Institutions in Postmodern Society*, Cheltenham, UK, Northampton, MA, USA: Elgar.

Bond, M. A. and Pyle, J. L. (2001) 'Diversity Dilemmas at Work', in *Approaches to Sustainable Development*, Forrant F., Pyle J. L., Lazonick W. and Levenstein C. (eds) Amherst: University of Massachusetts Press.

Brennan, G. and Lomasky, L. (1982) 'Institutional Aspects of 'Merit Goods' Analysis', *Finanzarchiv* 41: 183–206.

Breschi, S. and Malerba, F. (2001) 'The Geography of Innovation and Economic Clustering: Some Introductory Notes', *Industrial and Corporate Change* 10: 817–33.

Bryant, K. (2001) 'Promoting innovation: an overview of the application of evolutionary economics and systems approaches to policy issues', in J. Foster and J. St. Metcalfe (eds) *Frontiers of Evolutionary Economics, Competition, Self-Organization and Innovation Policy,* Cheltenham, UK, Northampton, MA, USA: Elgar.

Bush, P. D. (1987) 'The Theory of Institutional Change', *Journal of Economic Issues* 21: 1075–116.

——. (1999) 'Veblen's 'Olympian Detachment' Reconsidered', *History of Economic Ideas* 7: 127–51.

Calvert, R. L. (1995) 'The Rational Choice Theory of Institutions: Implications for Design', in D. L. Weimer (ed.) *Institutional Design*, Boston, Dordrecht, London: Kluwer.

Champlin, D. and Knoedler, J. (2002) 'Operating in the Public Interest or in Pursuit of Private Profits? News in the Age of Media Consolidation', *Journal of Economic Issues* 34: 459–68.

Ciccone, A. and Hall, R. E. (1996) 'Productivity and the Density of Economic Activity', *American Economic Review* 80: 54–70.

Cohendet, P., Creplet, F. and Dupouet, O. (2001) 'Organizational Innovation, Communities of Practice and Epistemic Communities: the Case of Linux', in A. Kirman and J. B. Zimmermann (eds) *Economics with Heterogeneous Interacting Agents*, Berlin: Springer.

Commons, J. R. (1934) *Institutional Economics. Its Place in Political Economy*. New York: Macmillan; reprinted (1990) in New Brunswick: Transaction Publishers.

Cooper, R. and John, A. (1988) 'Coordinating Co-ordination Failures in Keynesian Models', *Quarterly Journal of Economics* 103: 441–63.

Cossentino, F., Pyke, F. and Sengenberger, W. (eds) (1996) *Local and regional response to global pressure: The case of Italy and its industrial districts*, Geneva: ILO.

de Bruijn, J. A. and ten Heuvelhof, E. F. (1995) 'Policy Networks and Governance', in D.L. Weimer (ed.) *Institutional Design*, Boston, Dordrecht, London: Kluwer.

David, P. A. (1985) 'Clio and the Economics of QWERTY', *American Economic Review* 75, 2: 332–37.

——, Foray, D. and Dalle, J. M. (1998) 'Marshallian Externalities and the Emergence and Spatial Stability of Technological Enclaves', *Economics of Innovation and New Technologies* 6: 147–82.

de la Mothe, J. and Paquet, G. (1999) 'Structural competitiveness and interdependencies: regional patterns', in G. Boyd and J. H. Dunning (eds) *Structural Change and Co-operation in the Global Economy*, Cheltenham, UK, Northampton, MA, USA: Elgar.

Delaunay, J. C. and Gadrey, J. (1992) *Services in Economic Thought. Three Centuries of Debate*, Boston, Dordrecht, London: Kluwer.

Delorme, R. (2001) 'Theorizing complexity', in J. Foster and J. St. Metcalfe (eds) *Frontiers of Evolutionary Economics,* Cheltenham, UK, Northampton, MA, USA: Elgar.

Dequech, D. (1997) 'Uncertainty in a Strong Sense: Meaning and Sources', *Economic Issues* 2: 21–43.

——. (2001) 'Bounded Rationality, Institutions, and Uncertainty', *Journal of Economic Issues* 35: 911–29.

Dolata, U. (2002) 'Internet: Das Ende der Anarchie', *Blätter fuer deutsche und internationale Politik* 47: 1385–87.

Dupuy, J. C. and Torre, A. (1998) 'Co-operation and Trust in Spatially Clustered Firms', in N. Lazaric and E. Lorenz (eds) *Trust and Economic Learning*, Cheltenham, UK, Northampton, MA, USA: Elgar.

Elsner, W. (2000) 'An Industrial Policy Agenda 2000 and Beyond – Experience, Theory and Policy', in W. Elsner and J. Groenewegen (eds) *Industrial Policies After 2000*, Boston, Dordrecht, London: Kluwer.

——. (2001) 'Interactive Economic Policy: Toward a Cooperative Policy Approach for a Negotiated Economy', *Journal of Economic Issues* 35, 1: 61–83.

——. (2003) 'Global Industrial Policies', in P. D. Bush and M. R. Tool (eds) *Institutional Analysis of Economic Policy*, Boston, Dordrecht, London: Kluwer.

—— and Huffschmid, J. (1994) 'Industrial Policy', in G. M. Hodgson, W. J. Samuels and M. R. Tool (eds) *The Elgar Companion to Institutional and Evolutionary Economics,* Aldershot, UK, Brookfield, USA: Elgar.

Feser, E. J. (1998) 'Old and New Theories of Industry Clusters', in M. Steiner (ed.) *Clusters and Regional Specialization*, London: Pion.

Field, A. J. (1994) 'Game Theory and Institutions', in G.M. Hodgson, W.J. Samuels and M.R. Tool (eds) *The Elgar Companion to Institutional and Evolutionary Economics,* Aldershot, UK, Brookfield, MA, USA: Elgar.

Foray, D. (1998) 'The Economics of Knowledge Openness: Emergence, Persistence and Change of Conventions in the Knowledge Systems', in N. Lazaric, and E. Lorenz (eds) *Trust and Economic Learning*, Cheltenham, UK, Northampton, MA, USA: Elgar.

Friedman, J. W. (1994) 'Introduction and Overview', in J. Friedman (ed.) *Problems of Co-ordination in Economic Activity*, Boston, Dordrecht, London: Kluwer.

Fudenberg, D. and Levine, D. K. (1998) *The Theory of Learning in Games*, Cambridge, MA, USA, London: The MIT Press.

Gallaway, T. and Kinnear, D. (2002) 'Free Ride: An Institutionalist Analysis of Information in the Internet Age', *Journal of Economic Issues* 36: 441–47.

Gertler, M. S., Wolfe, D. A. and Garkut, D. (2000) 'No place like home? The embeddedness of innovation in a regional economy', *Review of International Political Economy* 7: 688–718.

Hart, J. (1992) 'The Effects of State-Societal Arrangements on International Competitiveness: Steel, Motor Vehicles and Semiconductors in the United States, Japan and Western Europe', *British Journal of Political Science* 22: 255-300; reprinted in W. Grant (ed.) (1995) *Industrial Policy*, Aldershot, UK, Brookfield, USA: Elgar.

Hayden, F. G. (1994) 'Institutional Theory of Economic Policy', in G. M. Hodgson, W. J. Samuels and M. R. Tool (eds) *The Elgar Companion to Institutional and Evolutionary Economics*, Aldershot, UK, Brookfield, USA: Elgar.

Hutter, M. (2001) 'Efficiency, Viability and the New Rules of the Internet', *European Journal of Law and Economics* 11, 1: 5–22.

Jessop, B. (1994) 'Changing forms and functions of the state in an era of globalization and regionalization', in R. Delorme and K. Dopfer (eds) *The Political Economy of Diversity. Evolutionary Perspectives on Economic Order and Disorder*, Aldershot, UK, Brookfield, USA: Elgar.

Jones, M. T. (2000) 'The competitive advantage of the transnational corporation as an institutional form. A reassessment', *International Journal of Social Economics* 27: 943–958.

Keeble, D. and Wilkinson, F. (1999) 'Collective Learning and Knowledge Development in the Evolution of Regional Clusters of High Technology SMEs in Europe', *Regional Studies* 33, 4: 295–303.

Kingston, W. (2000) 'A spectre is haunting the world – the spectre of global capitalism', *Journal of Evolutionary Economics* 10: 83–108.

Kirman, A. P. (1998) 'Economies with interacting agents', in P. Cohendet, P. Herena, H. Stahn and G. Umbauer (eds.) *The Economics of Networks* Berlin, Heidelberg, New York: Springer.

Klein, D. B. (1990) 'The Microfoundations of Rules vs. Discretion', *Constitutional Political Economy* 1: 1–19.

Lamberton, D. McL. (2001) 'The Knowledge Based Economy: Better Living or Bigger Profits?' in G. Sweeney (ed.) *Innovation, Economic Progress and the Quality of Life*, Cheltenham, UK, Northampton, MA, USA: Elgar.

Lambooy, J. (2000) 'Learning and Agglomeration Economies: Adapting to Differentiating Economic Structures', in F. Boekema, S. Bakkers, R. Rutten and K. Morgam (eds) *Knowledge, Innovation and Economic Growth. The Theory and Practice of Learning Regions*, Cheltenham, UK, Northampton, MA, USA: Elgar.

Lazonick, W. (1993) 'Industry Clusters versus Global Webs: Organizational Capabilities in the American Economy', *Industrial and Corporate Change* 2: 1–24.

——. (2001) 'Organizational Integration and Sustainable Prosperity', in F. Forrant, J. L. Pyle, W. Lazonick and C. Levenstein (eds) *Approaches to Sustainable Development*, Amherst: University of Massachusetts Press.

—— and O'Sullivan, M. (2000) 'Maximizing shareholder value: a new ideology for corporate governance', *Economy and Society* 29: 13–35.

Liebrand, W. B. G. and Messick, D. M. (eds) (1996) *Frontiers in Social Dilemmas Research*, Berlin: Springer.

Lindgren, K. (1997) 'Evolutionary Dynamics in Game-Theoretic Models.', in W.B. Arthur, S.M. Durlauf and D. Lame (eds) *The Economy as an Evolving Complex System II*, Reading, MA, USA: Addison-Wesley.

MacEwan, A. (2000) *Neo-Liberalism or Democracy? Economic Strategy, Markets, and Alternatives for the 21st Century,* Dhaka, Bangladesh: The University Press Ltd.

Maggioni, M. (1997) 'Firms, uncertainty and innovation policy', in G. Antonelli and N. De Liso, (eds) *Economics of structural and technological change*, N. London, New York: Routledge.

Maillat, D. and Lecoq, B. (1992) 'New Technologies and Transformation of Regional Structures in Europe: The Role of the Milieu', in N. Hansen, K. Button and P. Mijkamp (eds) (1996) *Entrepreneurship and Regional Development*, 4: 1–20. Reprinted in *Regional Policy and Regional Integration*, Cheltenham, UK, Brookfield, USA: Elgar.

Maillat, D. and Grosjean, N. (1999) 'Globalization and Territorial Production Systems', in M. M. Suarez-Villa and M. Steiner (eds) *Innovation, Networks and Localities*, Berlin: Springer.

Markusen, A. (1996a) 'Interaction between Regional and Industrial Policies: Evidence from Four Countries', *International Regional Science Review* 19: 49–77.

——. (1996b) 'Sticky Places in Slippery Space: A Typology of Industrial Districts', *Economic Geography* 72: 293–313.

May, C. (2002) 'Unacceptable Costs: The Consequences of Making Knowledge Property in a Global Society', *Global Society* 6: 123–44.

McKelvey, M. (2001) 'The Economic Dynamics of Software: Three Competing Business Models Exemplified Through Microsoft, Netscape and Linux', *Economics of Innovation and New Technologies* 10: 199–236.

Miller, E. S. (2001) 'The Impact of Technological Change on Market Power and Market Failure in Telecommunications', *Journal of Economic Issues* 35: 385–93.

Mirowski, P. (1989) *More Heat than Light. Economics as Social Physics, Physics as Nature's Economics*, Cambridge, Massachusetts.

Mizrahi, S. (1998) 'Regional Co-operation and Innovative Industries: Game-theoretical Aspects and Policy Implications', in M. Steiner (ed.) *Clusters and Regional Specialization. On Geography, Technology and Networks*, London: Pion.

Musgrave, R. A. (1987) 'Merit Goods', in J. Eatwell, M. Milgate and P. Newman (eds) *The Palgrave. A dictionary of economics*, London, Basingstoke: Macmillan.

Nielsen, K. (1992) 'The Mixed Economy, the Neo-liberal Challenge, and the Negotiated Economy', *The Journal of Socio-Economics* 21: 325–51.

—— and Pedersen, O. K. (1988) 'The negotiated economy: Ideal and history', *Scandinavian Political Studies* 2: 79–101.

Offerman, T. and Sonnemans, J. (1998) 'Learning by experience and learning by imitating successful others', *Journal of Economic Behavior & Organization* 34: 559–75.

Oinas, P. (2000) 'Distance and Learning: Does Proximity Matter?', in F. Boekema, S. Bakkers, R. Rutten and K. Morgan (eds) *Knowledge, Innovation and Economic Growth. The Theory and Practice of Learning Regions*, Cheltenham, UK, Northampton, MA, USA: Elgar.

Oltra, V. and Schenk, E. (1998) 'Evolution of Co-operation with Local Interactions and Imitation', in P. Cohendet, P. Herena, H. Stahn and G. Umbauer (eds) *The Economics of Networks*, Berlin, Heidelberg, New York: Springer.

Orlikowski, W. J. and Iacono, C. S. (2000) 'The Truth Is Not Out There: An Enacted View of the 'Digital Economy', in E. Brynjolfsson, and B. Kahin (eds) *Understanding the Digital Economy. Data, Tools, and Research*, Cambridge, London: MIT Press.

Ostrom, E. (1990) *Governing the Commons: The Evolution of Institutions for Collective Action*, New York: Cambridge University Press.

Padoan, P. C. (2001) 'Globalization, Regionalism and the Nation State: Top down and Bottom up', in M. Franzini and F. R. Pizutti (eds) *Globalization, Institutions and Social Cohesion*, Berlin, Heidelberg, New York: Springer.

Perraton, J. (2001) 'The global economy – myths and realities', *Cambridge Journal of Economics* 25: 669–84.

Piore, M. J. (2001) 'The Emergent Role of Social Intermediaries in the New Economy', *Annals of Public and Cooperative Economics* 72, 3: 339–50.

Powell, W. W. and Smith-Doerr, L. (1994) 'Networks and Economic Life', in N.J. Smelser and R. Swedberg (eds) *The Handbook of Economic Sociology*, Princeton New York: Princeton University Press and Russell Sage Foundation.

Pratt, A. (1997) 'The Emerging Shape and Form of Innovation Networks and Institutions', in J.M. Simmie (ed.) *Innovation, Networks and Learning Regions?*, London: Kingsley.

Rabach, E. and Kim, E. M. (1994) 'Where is the chain in commodity chains? The service sector nexus', in G. Gereffi and M. Korzeniewicz (eds) *Commodity Chains and Global Capitalism*, Westport, CT, USA: Greenwood Press.

Raikes, P., Jensen, M. F. and Ponte, S. (2000) 'Global commodity chain analysis and the French filière approach: comparison and critique', *Economy and Society* 29: 390–417.

Rallet, A. and Torre, A. (1998) 'On Geography and Technology: Proximity Relations in Localized Innovation Networks', in M. Steiner (ed.) *Clusters and Regional Specialization. On Geography, Technology and Networks*, London: Pion.

Ramstad, Y. (1991) 'From Desideratum to Historical Achievement: John R. Commons's Reasonable Value and the 'Negotiated Economy' of Denmark', *Journal of Economic Issues*, 25: 431–39.

Rantisi, N. (2002) 'The Local Innovation System as a Source of Variety: Openness and Adaptability in New York City's Garment District', *Regional Studies* 36, 6: 587–602.

Raymond, E. S. (2001) *The Cathedral and the Bazaar. Musings on Linux and Open Source by an Accidental Revolutionary*, Sebastopol, CA, USA: O'Reilly.

Redding, S. (1996) 'The Low-Skill, Low-Quality Trap: Strategic Complementarities Between Human Capital and R&D', *The Economic Journal* 106: 458–70.

Ruigrok, W. and van Tulder, R. (1995) *The Logic of International Restructuring*, London, New York: Routledge.

Rupert, M. (2000) *Ideologies of Globalization: Contending Visions of a New World Order*, London, New York: Routledge.

Ruys, P. H. M. (2000) 'Service-Specific Governance Regimes', paper presented at the Annual Conference of the International Society for New Institutional Economics, Mimeo, Tüebingen, Germany.

Rycroft, R. W. and Kash, D. E. (1999) *The Complexity Challenge. Technological Innovation for the 21st Century,* London, New York: Pinter.

Saxenian, A. (1994) *Regional Advantage. Culture and Competition in Silicon Valley and Route 128*, Cambridge, Massachusetts: Harvard University Press.

Schotter, A. (1981) *The economic theory of social institutions*, Cambridge, UK: Cambridge University Press.

Setterfield, M. (1996) 'Hysteresis and Uncertainty, Complementary or competing visions of evolving economic systems?' in S. Pressman (ed.) *Interactions in Political Economy*, London and New York: Routledge.

Shipman, A. (1999) *The Market Revolution and its Limits*, London, New York: Routledge.

Standing, G. (2001) 'Global Flexibility: Economic Integration, Social Disintegration?' in M. Franzini and F. R. Pizzuti (eds) *Globalization, Institutions and Social Cohesion*, Berlin, Heidelberg, New York: Springer.

Stein, J. (1997) 'How Institutions Learn: A Socio-Cognitive Perspective', *Journal of Economic Issues* 31: 729–40.

Steiner, M. (1998) 'The Discreet Charm of Clusters: An Introduction', in M. Steiner (ed.) *Clusters and Regional Specialization. On Geography, Technology and Networks*, London: Pion.

Storper, M. and Salais, R. (1997) *Worlds of Production. The Action Framework of the Economy*, Cambridge MA, USA, London: Harvard University Press.

Swyngedouw, E. (1997) 'Neither Global nor Local. 'Glocalization' and the Politics of Scale', in K. R. Cox (ed.) *Spaces of Globalization. Reasserting the Power of the Local*, New York, London: Guilford Press.

Tichy, G. (1998) 'Clusters: Less Dispensable and More Risky than Ever', in M. Steiner (ed.) *Clusters and Regional Specialization. On Geography, Technology and Networks*, London: Pion.

Tirole, J. (1995) *Industrieökonomik* Munich, Vienna: Vahlen.

Tool, M. R. (1979) *The Discretionary Economy: A Normative Theory of Political Economy*, Santa Monica, CA, USA: Goodyear Publishing Company.

Torre, A. and Gilly, J. P. (2000) 'On the Analytical Dimension of Proximity Dynamics', *Regional Studies* 34, 2: 169–80.

Ver Eecke, W. (1998) 'The Concept of a 'Merit Good'. The Ethical Dimension in Economic Theory and the History of Economic Thought or the Transformation of Economics Into Socio-Economics', *Journal of Socio-Economics* 27: 133–53.

Weitzel, T. and von Westarp, F. (2002) 'From QWERTY to Nuclear Power Reactors: Historic Battles for the Standard', in K. Geihs, W. König, F. von Westarp (eds) *Networks. Standardization, Infrastructure, and Applications*, Heidelberg, New York: Physica.

Wellhoener, V. (2002) *Oekonomik. Physik. Mathematik. Das stabilitaetstheoretische Forschungsprogramm der Allgemeinen Gleichgewichtstheorie im Kontext seiner konstitutiven Metaphern*, Frankfurt/M.: P. Lang Publishing.

Wettengl, S. (1999) *Initiierung technologischer Systeminnovationen. Wege zur Vermeidung von Abwarteblockaden in Innovationsnetzwerken*, Göttingen: Vandenhoeck & Ruprecht.

Yu, T. F. L. (2000) 'A new perspective on the role of the government in economic development. Co-ordination under uncertainty', *International Journal of Social Economics* 27: 994–1012.

5 Policy for social costs: Kapp vs. neoclassical economics

*James A. Swaney**

Introduction

K. William Kapp was an extraordinary social scientist who contributed much to an interdisciplinary systems approach to the social sciences, particularly in relation to the biophysical environment. His work anticipated core contributions to ecological economics, and he was a major contributor to social economics and institutional economics. Kapp attempted to open social scientists' eyes to the shortcomings of the market system, especially its relentless drive to shift costs from the powerful to the weak, off the books of going concerns onto workers or the biophysical environment. Kapp also attempted to alert scientists and policy-makers to the mounting threats that modern economies pose to basic ecosystems that support all life. Although his arguments were well-reasoned and supported with substantial empirical evidence, his limited success in these educational efforts is not surprising. Cost-shifting is a radical concept, undermining capitalism's legitimacy. Ecological education is not easy either, as a significant portion of social scientists, including a large majority of economists, continue to believe that the biophysical environment is immune from significant degradation.[1] Despite widespread frequent publication of scientific findings that we are wrecking earth's life-support systems – systems that provide air to breathe, water to drink, climate to support agriculture, and so on – many economists remain oblivious, 'plugging away' in what Thomas Kuhn called 'normal science' (Kuhn 1962).

Kapp also emphasized that market outcomes are unlikely to coincide with community norms, which should be determined democratically.[2] Even in the idealized unreality of perfect competition, where no cost escapes the market's accounting and all market participants are price-takers (having no economic power), the market's efficient results are unlikely to satisfy community standards for basic human material needs, necessary public goods, or intertemporal objectives such as sustainability. Although serious enough, these problems are less threatening to neoclassical economists than the concept of cost-shifting or the realization that economies are undermining their biophysical foundations. Orthodox economists acknowledge that society has to be responsible for a few public goods and has the right to tinker with market rules, and occasionally with market outcomes, but hasten to add that such tinkering is ill-advised because it nearly always shrinks

the economic pie. They also contend that discussions about the distribution of income and wealth lie outside economic science, falling on the wrong side of the line separating 'positive' (legitimate) economics from 'normative' (illegitimate) discourse.[3]

Among his many contributions to our understanding of the interconnected social and biophysical systems, Kapp's most important concept is cost-shifting. Cost-shifting suggests, contrary to conventional economics, that competitive capitalism contains a built-in tendency to reduce (accounting) costs, not only through development and implementation of better management or technology (as touted by neoclassical economists), but also through whatever other means might be available, including deliberate transfer or shifting of costs onto workers, the natural environment, or the community at large. In other words, cost-shifting is predatory behaviour, where the powerful seize benefits for themselves and impose costs onto others. Kapp's earliest discussion of cost-shifting is found in *The Social Costs of Private Enterprise*, originally published in 1950

> Indeed, generally speaking, capitalism must be regarded as an economy of unpaid costs, 'unpaid' in so far as a substantial portion of the actual costs of production remain unaccounted for in entrepreneurial outlays; instead they are shifted to, and ultimately borne by, third persons or by the community as a whole.
>
> (Kapp 1971a: 231)

When costs are unaccounted for, economic growth is 'over-measured' (overstated or inflated). Many environmental costs do not show up in the national income accounts, a problem that has been recognized for decades. The architect of the US national income accounts in the 1930s, Simon Kuznets, in the 1950s called for an expansion of these accounts so as to at least begin to measure environmental costs.[4] Not only are costs not properly accounted for, but many avoided costs end up damaging the natural environment or people's health, leading to additional spending on remediation or health care, which in turn further inflates economic growth.

Kapp understood that capitalist institutions constituted a system for shifting costs from businesses onto workers, the environment, and the community at large, meaning that we 'grow our economy' with unsustainable and ultimately counter-productive practices. After eleven chapters analysing and documenting various types of social costs, Kapp concluded

> What these losses have in common and what makes them truly social costs is the fact that they do not enter into the cost calculations of private firms. They are shifted to and are paid for in one form or another by individuals other than the entrepreneur or by the community as a whole or by both.
>
> (Kapp 1971a: 229–30)[5]

Kapp realized that property relations and other institutions of capitalism not only permit many different types of costs to be transferred from businesses

to workers, society, and the environment, but actually encourage this cost-shifting. Why, asked Kapp, should capital be treated as a fixed resource and labour as variable, so that a recession hurts working people more severely than it hurts bankers? Why should many environmental resources, such as water and air, be treated as free, leading to wasteful overexploitation and destruction? Why should rivers and oceans and the atmosphere be treated as free dump stations? Why should poor, largely minority households be exposed to far higher levels of hazardous wastes and other industrial contaminants than those of the middle and upper classes? (Yes, in 1950 Kapp was concerned with what we now call 'environmental justice.')

A cogent expression of Kapp's environmental policy prescriptions is found in chapters seven and eight in the John E. Ullmann-edited *Social Costs, Economic Development, and Environmental Disruption* (Kapp 1971b), and 'Environment and Technology: New Frontiers for the Social and Natural Sciences' (Kapp 1977). In this article, Kapp lays out three sets of policies for environmental improvement:

1) 'strict public control of the use of noxious inputs and the disposal of dangerous residuals;'
2) 'systematic development and promotion (...) of technologies with a low ecological impact in order to reduce the degradation of the human environment by production and consumption;'
3) 'aim at increasing the natural environment's capacity to assimilate residuals, on the one hand, and developing new ways of recovering and reusing waste materials, on the other' (Kapp 1977: 534–35).

These policies, while having implications for population and resources, are focused not on the input (resource) stream but on the output (waste, residual) stream. This emphasis, found throughout Kapp's environmental work, sets him apart from scholars with a more 'Malthusian streak,' providing the key to understanding Kapp's contributions to ecological, social, and institutional economics.

Kapp's ecological economics understands that the market system is increasingly deficient in directing economic activities as the earth becomes both 'fuller,' with human numbers and power bringing an end to resource abundance, and 'smaller,' with residuals stressing and degrading earth's life systems and reducing their capacity to cycle nutrients, to clean the air and water, and to produce soils, stable climate, and other ecological life support services.[6] Kapp's social economics rejects the legitimacy of the current distribution of income, wealth, rights, and responsibilities, arguing instead for a bill of economic rights. Kapp calls this 'existential minima', arguing on moral grounds that we are our sisters' and brothers' keepers; that we are collectively responsible for providing basic necessities for all citizens (Kapp 1965). In light of the 'reverse public good' problems arising from ecological degradation, we will do well to protect the environment for future generations while providing basic needs to people today. Along with basic nutrition and healthy air and water, Kapp's existential minima list might also include personal security, health care, and educational opportunities. Kapp's institutional economics, which was probably influenced more by Myrdal than Veblen, advocates a participatory

democratic process informed by interdisciplinary systems analysis to develop and pursue the community's objectives (which Kapp believes should include existential minima and essential public goods).[7]

In my judgement, Kapp's ecological and institutional economics suggest that he would agree with mid-20th century American ecologist Aldo Leopold that basic existence rights should be extended to all species. Leopold's term for this is the 'land ethic' (Leopold 1949). Leopold explains that ethics (rules differentiating social from antisocial behavior) between individuals developed earlier than ethics for relations between individual and society. Although these two ethics continue to evolve, Leopold argues that increases in population density and the efficiency of tools require us to develop a third ethic

> dealing with man's relation to land and to the animals and plants which grow upon it (...) The extension of ethics to this third element in human environment is, if I read the evidence correctly, an evolutionary possibility and an ecological necessity.
>
> (Leopold 1987: 203)

Kapp was human-centered, and to my knowledge did not cite Leopold. Nonetheless, Kapp's discussions of our ignorance of ecological systems and his concern for the plight of future generations are consistent with Leopold's land ethic on two levels. On the moral level, the fullness of life and the welfare of future generations would be served by extending biotic rights to all species. On the practical level, a land ethic would provide substantial protection from the grave threats to life-support services produced by ecosystems.

Also like Leopold, Kapp realized that proper care of our ecosystems could not possibly be accomplished by government controls alone; that without fundamental change in our understanding of and attitudes toward the environment, it would be impractical to develop or enforce regulations extensive enough to save Earth's ecosystems from irreversible degradation. As Leopold said, a land ethic, including new attitudes and explicit norms of behavior, is an 'ecological necessity'. Kapp put it like this

> This new code of norms and responsibilities towards the environment will have to be based upon the scientific analysis of the interrelationships which connect man's social and economic systems with those physical and biological systems upon which human life on this planet depend.
>
> (Kapp1971b: 142)

Kapp's ethic is similar to Leopold's, as both recognized that extending biotic rights would protect ecosystems from degradation and that neither market or state, alone or in combination, are up to the task.

Before elaborating on the 'full Earth / small Earth' perspective of ecological economics (Section II), followed by a 'Kapp-inspired' critique of the two neoclassical approaches to social cost (Section III), a clarification of Kapp's 'social cost'

concept is in order. To neoclassical economists, accounted-for costs plus externalities equal total (sometimes, 'social') costs. Their term for the uncompensated side-effects or 'spillovers' of market transactions is 'externality' or 'external cost.' Kapp did not like the word 'externality,' so he uses different terminology. To Kapp, accounted-for costs plus social costs equal total cost to society. In other words, what to a neoclassical economist is an externality is to Kapp a social cost.

Kapp objects to the terms 'external cost' and 'externality' on the grounds that this terminology suggests incorrectly that such costs are accidental and infrequent, occurring only in exceptional circumstances unrelated to, and certainly in no way caused by, the system of investment for profit. On the contrary, argued Kapp: Uncompensated costs are both pervasive – the rule rather than the exception – and systemic – predictable and widespread rather than incidental (Swaney and Evers 1989: 8). 'Externality' suggests that the problem is outside the system. Kapp insisted the problem originated inside the system; that these costs are endemic, arising from and produced within – the product (result) of a fundamental design flaw. In systems language, competitive capitalism's relentless pressure to reduce accounting costs by whatever means possible produces a positive feedback loop of deteriorating standards in the treatment of everyone and everything with which the profit-pursuing enterprise comes in contact. A system that affords decision-makers opportunities to avoid responsibility for costs, while simultaneously placing those same decision-makers in a competitive environment with constant pressure to reduce accounting costs, is a system designed to erode community standards. In evolutionary terms, the system selects for the ruthless gene and against the responsible gene. It is this deep-rooted systemic flaw that Kapp termed 'cost-shifting.'

Full, small Earth

'Empty-full' is about Earth's capacity to supply natural resources: On empty Earth nature may be stingy, but potential resources appear unlimited; whereas on full Earth people are bumping up against resource constraints and threatening the fecundity of forests, fisheries, and other natural systems that produce flows of natural resources. On empty Earth human-produced capital is relatively scarce and natural capital abundant; whereas on full Earth natural capital has become the limiting factor.[8] On empty Earth frontiers permit the consequences of resource overuse to be ignored; whereas on full Earth 'free for the taking' is restricted to nuisance species that thrive in degraded ecosystems. On empty Earth human population is far below carrying capacity; whereas on full Earth we are approaching carrying capacity.[9] Early 21st century Earth is not completely 'full,' but is well on its way.[10] As of the mid-1980s, approximately one quarter of Earth's photosynthetic ('net primary') product (NPP), and 40 per cent of terrestrial (land-based) NPP, were being appropriated by homo sapiens (Vitousek *et al*. 1986). Even with the encouraging projections that Earth's human population will stabilize in a few decades, little of what will remain could be called 'nature,' and if conventional economic growth continues, wildness will nearly disappear and the survival of Earth's remaining biodiversity will be the project of zoos. Earth's fullness indicates

that human economies are limited, less by the quantity and quality of equipment and technology (conventional or 'human-made' capital), and more by Earth's capacities ('natural' capital) (Daly 1996).

'Large-small' is related to 'empty-full', but is less about resources and more about environmental health – about Earth's capacity to assimilate wastes and provide ecological services (Swaney 2003). On large Earth production of ecosystem goods and services such as soils, stable climate, protection from damaging solar radiation, nutrient cycling, and waste assimilation appear inviolable and are taken for granted; whereas a shrinking Earth raises questions about the health of ecosystems and their ability to produce life-support services. On large Earth, ecological disruption is sporadic, isolated, and temporary; whereas on small Earth ecological disruption is regular, widespread, and chronic, with long-lasting or permanent ecosystem degradation. On small Earth, the ecological principles, 'Everything goes somewhere', and 'Everything is connected to everything else', grow in economic importance as more and more market activities impact third, fourth, fifth, and nth people who had no say in the market exchanges that impose costs on them. For example, since the beginning of the industrial revolution, atmospheric concentrations of carbon dioxide have increased by one-third, to 375 ppm in 2003 (from 280 ppm). Because much industrial era CO_2 has been absorbed by Earth's 'carbon sinks', this increase is far less than would be expected based on fossil fuels burned and lands deforested. Although much carbon dioxide has been absorbed by oceans, science has much to learn about Earth's carbon sinks. An acceleration in the rate of increase in carbon dioxide concentrations in the past couple of years, however, suggests that positive climate change feedbacks are accelerating, possibly because oceans are approaching their carbon storage capacity; that is, the ocean 'sink' is 'filling up'. Along with climate change, other global problems such as ozone depletion and species loss suggest human commandeering of more and more NPP and human impacts on Earth's ecosystems are 'shrinking' Earth at an accelerating rate, imposing untold costs on future generations. The smaller Earth becomes, the more pervasive and severe are the social costs, and the more difficult it becomes to properly account for these spiralling social costs. On small Earth, new technologies that 'grow' the resource base, thereby counteracting the trend toward full Earth on the resource side, frequently exacerbate threats to the Earth's life-support ecosystems. For example, advanced search and drilling technologies permit economic extraction of increasingly hard-to-find and difficult-to-extract oil deposits, but produce additional water and air pollution and exacerbate climate change.

People have always changed ecosystems, but as Earth becomes smaller, ecosystem degradation increases, reducing services provided by ecosystems. On large Earth degradation is confined to small regions and Earth's life-support functions appear unchanged; whereas on small Earth degradation is commonplace and substantial, altering important systems on a global scale. Stratospheric ozone depletion and climate change provide evidence of small Earth conditions caused by residuals from industrial economies, and the accelerating decline in biological diversity (biodiversity) over so much of the globe substantiates the pervasive, deleterious, and extremely costly consequences of the combined effects

of residuals and increasing demands for resources. In the words of designer William McDonough, on our small Earth 'away went away'.

Empty Earth is linked with large Earth (and full Earth with small Earth) through the basic ecological principle, 'Everything goes somewhere'. This is the same concept as the First Law of Thermodynamics in physics and the materials balances approach in economics. Also known in physics as the Law of Conservation of Matter (and Energy), this means that all material and energy inputs into economies eventually become residuals. Green design, for greater durability or easy (minimal consequence) recycling, can reduce the volume and impact of the economy's residuals, but ultimately the volume of residuals depends on the volume of material and energy inputs from the environment into the economy.

In his 1966 essay, 'The Economics of the Coming Spaceship Earth', Kenneth Boulding described empty-large Earth as the 'cowboy economy' and full-small Earth as the 'spaceman economy' (Boulding 1993). The cowboy economy is an open system where Earth is perceived as a limitless plane with limitless resource inputs and unlimited capacity to assimilate wastes. In the cowboy economy the typical outcome, if not the primary objective, is to maximize 'throughput', as measured by GDP. On an empty, large Earth, a cowboy economy does not seem inappropriate. Everyone has plenty of 'elbow room', and the frontier offers ample natural resources and pristine environments should lumber or fish become scarce, soil erode, or water become scarce or unfit to drink. The 'spaceman economy', on the other hand, is a closed system where both resources and waste assimilation are limited and human survival over time requires careful design and planning because all material waste must be recycled. Behaving like a cowboy in the spaceman economy is suicidal because ecological life-support services are degraded by 'throughput'. In the spaceman economy, residuals that are free of toxins can be composted, providing nutrients for natural systems, but other residuals (except for unrecoverable waste heat) must be reused indefinitely. Although 'spaceship Earth' receives energy inputs from the sun and radiates (some) waste heat into space, all materials stay in the system; that is, all inputs necessarily become residuals, so the objective is to maintain stocks by minimizing material inputs and using the design principle 'waste = food' to prevent death by starvation, poisoning, or both (McDonough and Braungart 1998). On spaceship Earth, the consequences of shifting costs are apparent and often severe; that is, it becomes necessary to fully account for all costs over time. In other words, capitalism can no longer survive as an economy of unpaid costs. Cost-shifting must come to an end; capitalism must undergo a radical transformation.

While physics (First Law of Thermodynamics), economics (material balances), and history all indicate that empty Earth corresponds to large Earth and full Earth corresponds to small Earth, distinguishing between input and output (residual) issues highlights important differences between resource problems on the one hand (full Earth 'running out' concerns) and life support problems on the other (small Earth ecosystem degradation problems including climate change and loss of biodiversity). This dual model also suggests the logical possibility of two other combinations of resource and ecological conditions: 'full-large Earth' and 'empty-small Earth'.

Could Earth be full and large simultaneously? That is, could we butt up against resource constraints, yet face no ecological problems from residuals? Malthusians are often portrayed in this framework. A shortage of food could be caused by agricultural productivity limitations unrelated to, for example, soil erosion or climate change. Conceivably, with enough understanding of Earth's waste assimilation and life-support functions, environmentally benign design and good environmental stewardship could prevent ecosystem degradation, keeping the world large. Nevertheless, once human population grew to command 100 per cent of net primary product, carrying capacity would be reached and Earth would be full.

The other logical possibility is an Earth that is empty, yet small. Many institutionalists as well as architect William McDonough fit this 'empty-Earth, small-Earth' model. Institutionalists consider the Earth empty because they believe humans have limitless problem-solving potential. We are not about to run out of resources because human creativity converts 'neutral stuff' into 'resources'. In the words of geographer Erich Zimmermann, 'Resources are not, they become', and 'Knowledge is truly the mother of all other resources'. People are extremely talented at developing substitutes and entirely new resources, permitting continued economic growth. But economic growth has always meant more inputs, and rising inputs necessarily produce rising residuals. Technology can be miraculous, but even the most unimaginable innovations are quite unlikely to overturn the First Law of Thermodynamics! As throughput expands, at some point residuals that were benign or even nutritious at lower levels begin to degrade ecosystems. That is, applying human creativity to solve resource problems increases throughput, and increasing throughput inevitably overloads nature's sinks and reduces both the stability and resilience of key ecosystems. Far from 'running out' of resources, the problem is that we have been too successful; that is, we have created too many resources that end up as ecology-degrading residuals. Our success in avoiding full Earth has apparently expanded carrying capacity, but this expansion is fleeting because residuals' effects are pulsing through Earth's ecosystems like a slow tsunami through an ocean. Our focus on growing our economies has not kept the Earth empty, but has (so far) managed to postpone fullness, yet in the process we have been undermining Earth's capacity to assimilate waste and support life. We urgently need to refocus human creativity toward small Earth problems if we are to avoid a global epidemic of Malthusian outcomes.

Generally, ecological economists are occupied with both full Earth and small Earth problems, but their focus on full Earth problems often diverts attention from the more difficult and perilous small Earth problems. Kapp maintained a focus on small Earth problems for a quarter-century beginning with *The Social Costs of Private Enterprise* in 1950. The same cannot be said for his contemporaries. Even after Boulding's 'Spaceship Earth' essay in 1966, economists were slow to pick up on the critical importance of waste assimilation and life support services produced by the natural environment. Ecological and institutional economists began making significant contributions in the 1970s, two decades after Kapp first addressed the issue.[11] Unfortunately, many American institutionalists, especially of the Ayresian school, remain unconcerned with small Earth problems.[12]

The two neoclassical approaches to addressing social costs

Conventional economics offers two approaches to dealing with a divergence between private and total costs: tax/subsidy and property rights. The tax/subsidy approach was developed by A. C. Pigou in the early 20th century (Pigou 1970). The seminal article for the property rights approach was by Ronald Coase (Coase 1960). This section critiques these two approaches from Kapp's perspective, using Alan Coddington's elaboration of Kapp as the starting point (Coddington 1970).[13]

From the voluminous literature of conventional economics during the first half of the 20th Century, Kapp considered Pigou's *The Economics of Welfare* to be the only substantive contribution to the general analysis of social costs. Kapp wrote that neoclassical economists have 'no room' for the analysis of social costs 'with the exception perhaps of Pigou's' conceptual system (Kapp 1971a: 232). The basic idea underlying Pigou's approach is really quite simple: When some costs escape market accounting, a tax is levied on the product to raise the price to where it 'should be'; that is, to where all costs are included. The 'correct' tax is just enough to account for all unpaid costs, providing consumers the 'correct' price. At this higher price, less is purchased, leading to lower production and, consequently, fewer unpaid costs. Remaining unpaid costs can be at least partially compensated with revenue from the tax. Just as taxes are levied on products involving external costs, subsidies are provided for products involving external benefits, and for research and development of technologies that promise to reduce external costs.

In recent years, Pigou's idea has gained currency in popular U.S. environmental literature under the label 'green taxes'. Two prominent examples are Paul Hawken and Lester Brown. Hawken advocated 'getting the prices right' in *The Ecology of Commerce* (Hawken 1993). Brown, long-time critic of economics and director of the respected WorldWatch Institute, gave green taxes a central role in addressing many of the Earth's most pressing environmental problems in *Eco-Economy* (Brown 2001).

By raising prices of environmentally damaging economic activities, green taxes reduce the competitive advantage enjoyed by environmentally damaging (brown) industries such as fossil fuels and encourage research and development of green products and markets such as wind, solar, and other renewable energy resources. Participants in markets do indeed respond to incentives, and perverse incentives (including 'artificially' cheap brown goods) do indeed make problems worse. It follows that green taxes – getting the prices right as regards the environment – should be an integral part of environmental policy, but it is a mistake to use them as the foundation.

Kapp argued correctly that environmental policy should not be based on green taxes or other 'indirect controls'. '[I]ndirect methods fall short (...) although they may become part of a comprehensive anti-pollution program' (Kapp 1971b: 129). Environmental policy must be constructed on a foundation of knowledge from the biophysical sciences, particularly ecology. With our dearth of knowledge (relative to what we need to know to establish rights and rules that protect ecosystems and biodiversity), the precautionary principle needs liberal application.[14] But even if

prices are adjusted both to reflect known environmental costs and to insure against bad outcomes resulting from scientific ignorance, neither of which are easy tasks, private property 'free' markets still come up short. A severe problem with market allocation is that value in the marketplace is synonymous with the aggregate willingness to pay of individuals. This is morally reprehensible because it excludes the environmental needs of the poor and the unborn. Nor is this the end of debilitating shortcomings in market-driven allocation of environmental goods and services. Those who have money to register their wants in the market face conflicts between what they want as individuals and what they want as members of the community. For example, most Americans, having no viable transportation alternative to the automobile, are frequently frustrated with speed limits and traffic lights. We would like to be exempt from the rules, but we certainly wouldn't want everyone to be exempt. Or, more to the point, as individuals most Americans want to be free to drive comfortable, massive vehicles, but as members of the community we want to reduce both pollution and dependence on oil imports. These problems are compounded by still another set of problems, the inherent short-sightedness of markets combined with the reality of environmental irreversibilities. Even with good information, markets essentially ignore all consequences for generations beyond our grandchildren, and both ecosystems and species lost are lost forever.[15]

Inappropriate prices 'explain' only a fraction of the divergence of individual and social interests as regards the environment. Indeed, 'free-market' prices are symptomatic of fundamental inadequacies of a market approach to environmental policy. Whether or not we are all genetically 'wired' with a fundamental need for interaction with other species (E. O. Wilson's biophilia hypothesis [Wilson 1984]), nearly everyone would prefer to bequeath all of Earth's flora and fauna to future generations. Markets in which to register this preference do not exist, nor are they merely 'missing'. Markets require divisible, transferable property rights that spell out the benefits (and duties or obligations) of the holder (owner) of the property (Bromley 1978). With still-unidentified species on Earth numbering in the millions, the most fundamental knowledge prerequisites for property rights are lacking. Even with the miraculous bridging of this knowledge gap, efficient markets still could not be constructed because species inhabit niches in interconnected, interdependent ecosystems. Even where private ownership could be established for economically valuable ecosystem components, free markets in such resources would fail miserably because dividing and transferring their economic components will nearly always degrade the ecosystem. As Aldo Leopold wrote

> A system of conservation based solely on economic self-interest is hopelessly lopsided. It tends to ignore, and thus eventually to eliminate, many elements in the land community that lack commercial value, but that are (as far as we know) essential to its healthy functioning. It assumes, falsely, I think, that the economic parts of the biotic clock will function without the uneconomic parts.
>
> (Leopold 1987: 213)

Even if we could somehow miraculously protect ecosystem integrity and correct for our dearth of knowledge about ecosystem functioning, the market system's fundamental intergenerational shortcoming remains. Competitive markets are notoriously short-sighted, as the time value of money coerces market participants into discounting future values, effectively ignoring intergenerational consequences. Renewable resources that grow at a rate below the market interest rate are treated as nonrenewable, stock resources. For example, white oak is a valuable slow-growing tree that, when harvested, is replaced by faster-growing species. The market acknowledges no biotic rights for the white oak or species that depend on it. For example, the white oak provides prime roosting sites and insect-rich foraging canopy for the endangered Indiana bat (http://www.nj.gov/dep/fgw/ensp/pdf/end-thrtened/indianabat.pdf, accessed 12.12.2004). Although less competitive markets permit somewhat longer time horizons, markets are simply incapable of addressing long-term environmental threats such as climate change and loss of biodiversity.

In 1970, British economist Alan Coddington published a three page-essay that builds on Kapp's 1950 classic, *The Social Costs of Private Enterprise* (Coddington 1970). Here Coddington emphasizes a fundamental difference between goods ('commodities') and bads or residuals ('discommodities'), providing valuable insight into problems with the property-rights approach to the environment

> A commodity is something you want to acquire; a discommodity is something you want to get rid of. To acquire something is to acquire it from someone (...) But to get rid of a discommodity is not necessarily to get rid of it to anyone in particular (...) The discommodity then becomes the (unwanted) possession of society at large ...
>
> (Coddington 1970)

This is Coddington's fundamental asymmetry. A general term for acquiring commodities is benefit-capturing, and a general term for getting rid of discommodities is cost-avoiding. Benefit-capturing is familiar to economists. As Douglas North explains, when benefits are anticipated, people work to obtain ownership of ideas and property, including the creation of new private property where none previously existed (North 1991). That is, capitalism's 'design mechanism' for capturing benefits is property rights, specifically the ability of those who produce (or anticipate producing) commodities to obtain ownership privileges so that those who wish to benefit can be excluded if they refuse to pay. That is, those who produce benefits (or are considering going into the business of producing a particular set of benefits) have an incentive to petition the government for property rights so as to capture the benefits – so as to make money by selling the good to those who are willing to pay for it. If no property rights exist, the commodity producer will petition the government or legislature or the courts. For example, in Diamond V. Chakrabarty (1980), the US Supreme Court ruled that genetically modified organisms (GMOs) could be patented. Had the Court ruled instead that GMOs,

like other life forms, are gifts of nature that belong to everyone, today's world would be much different.

Curiously, most economists remain unfamiliar with the cost-avoiding side of the equation. Just as surely as people will work to establish private property rights and market rules to capture benefits, people will also resist institutional innovation to capture costs. The anticipation of costs elicits either no response or a parallel response in the opposite direction. Anticipation of costs under conditions of no property (open access) elicits no response, because the discommodity can be released ('gotten rid of') to no one in particular (and therefore to 'society at large').[16] Where private property rights exist, anticipation of costs elicits efforts to actively avoid (shift) costs. Such cost-avoidance occurs through either institutional or technological change. In some cases such cost-avoiding changes actually reduce costs from the social as well as the private accounting stance, but in the absence of appropriate mechanisms (institutions) of social control, such outcomes amount to no more than fortunate happenstance.[17] As Kapp argued half a century ago, the incentives of private free markets lead to cost-avoiding institutional and technological changes that reduce private costs, not by actually eliminating them, but by shifting these costs on to someone else (Kapp 1971a). That is, the cost-avoiding, discommodity side of Coddington's asymmetry is synonymous with Kapp's cost-shifting.

On empty, large Earth, with few people, feeble technologies, ample capacity for waste processing by natural systems, and little or no understanding of ecology or of the health effects of various pollutants, costs could be avoided or shifted easily and often without notice. The context today is, to understate the case, 'somewhat different'. The pervasive impacts of discommodities and the need to address shifted costs grow exponentially as we grow our numbers and our economies. As discussed in the previous section, on full Earth fish harvest is limited by the number of fish, not by fishing effort or technology, and on small Earth fisheries are being degraded not only by overfishing, but also by ecologically unsound fishing technologies, pollution, ocean warming, deforestation, and other ecological problems.

For decision-makers faced with competitive (or bureaucratic) pressure, every cost is a discommodity. Whatever paths are available for reducing costs are paths that every business manager has incentive to follow – the incentive to 'release' garbage onto whomever or whatever is powerless to resist. The greater the disparity in the distribution of income, wealth, and knowledge, the easier it is for the powerful to find recipients for cost-shifting. Also, the less robust the democracy, the less trouble it is to shift costs onto the politically disenfranchised. While these factors determine the relative ease with which costs are shifted, Coddington's fundamental asymmetry underlines Kapp's general conclusion, quoted in the introduction, that capitalism is an economy of unpaid costs.

The property rights approach, widely attributed to Ronald Coase (Coase 1960), does not ignore social costs. If costs are going unpaid by market participants, the rights of people bearing these costs are being violated, and these people will use the legal system to protect their rights or obtain compensation for their losses.

This approach assumes that rights are clearly established and that political and economic power are equitably (if not equally) distributed. If rights are not clearly defined, a likely circumstance for discommodities, or if those bearing the burden of shifted costs have little power, this approach fails. Property rights advocates acknowledge that government has a role in specifying rights and entitlements, and in providing a judicial system through which people can pursue protection or compensation when their rights or property have been violated, but most advocates of the property rights approach are adamant that government's role should be strictly limited. The effect is to promote cost-shifting by denying rights to those with little power. That is, the potential of this approach is undermined because many problems that might be addressed effectively from this general perspective require a substantial government effort in defining rights and responsibilities and in constructing institutions to allocate these rights and responsibilities. A good example is the sulphur allowance system established by the 1990 US Clean Air Act Amendments. Although highly successful in reducing total sulphur emissions, the US Congress doled out free sulphur allowances in proportion to historic emissions; that is, the biggest polluters were given the most allowances. For the first decade, they could continue polluting at their historic levels without penalty. If they reduced emissions, they could profit by selling their unused allowances. Also, there is no provision in the sulphur allowance system for regional 'hot spots'.

As Coddington pointed out, because of the asymmetry between commodities and discommodities, the incentive to establish property rights in commodities is lacking in discommodities. Few if any private parties will be petitioning the government to set up property rights in discommodities, and such property rights will, as a general rule, have to protect entire communities, geographic regions, states, or even the entire Earth. Such property rights will seldom be similar to the private property rights that protect providers of commodities

> The whole notion of property rights is quite different for discommodities from what it is for commodities. Property rights in the case of commodities protect the owner of the commodity (...) Property rights in the case of discommodities must protect not a single owner, but society at large.
>
> (Coddington 1970)

Advocates for the neoclassical property-rights approach cannot bring themselves to admit that property rights to protect against discommodities will be social and community rights, not private rights, and that governments will have to be much more involved than in rights for commodities. For example, although *Free Market Environmentalism* (Anderson and Leal 1991, 2001) cited neither Pigou nor Coase in the first edition, the authors clearly favour the neoclassical property-rights approach and equate it with the free market.[18] Free-market environmentalists believe that technology and entrepreneurs will solve problems of unpaid costs, as illustrated by their stories of cattle branding and barbed wire. In the American West farmers consistently suffered large losses from crop predation, particularly on the Great Plains where the combination of large fields and a scarcity

of wood made fence building to exclude unwanted animals prohibitively expensive. The invention of barbed wire dramatically reduced the cost of fencing. Not only were unpaid costs (losses) reduced, fortunes were made in the barbed wire industry. This provides a nice illustration of how technology can reduce social costs when the problem is the high cost of enforcing clear, established, transferable, divisible, private property rights; that is, when the only impediment to efficient property rights is that exclusion of unauthorized uses is expensive. Cattle branding is a similar example, except the resource is mobile. There are probably many historic examples where new institutions or technologies overcame impediments to efficient private property. Unfortunately, today's environmental problems are not amenable to such solutions. Most social costs are associated with resources or ecosystem services where few if any conditions of efficient private property can be realized. The best hope for reducing many of these social costs is the development of common property institutions, about which a considerable literature has developed in recent decades.[19]

Anderson and Leal offer several hypothetical examples of how high-tech devices may allow private property to be established in new areas. For example, wolves that venture out of their designated area could be sedated by collars that also serve as tracking devices, allowing them to be transported back into their territory. Similarly, whales could be auctioned off to wildlife advocates after equipping each whale with an identification device that would allow its owner not only to track it, but also to somehow monitor violations of whale rights. One cannot help but wonder what recourse a whale owner has when ecological degradation wipes out that species' food supply. If this high-tech 'solution' is applied to polar bears, what recourse does the polar bear owner have as climate change shrinks Arctic sea ice, removing the bear's hunting vehicle? Generally, ecosystems simply cannot be made to fit the private property model. Divisible, transferable property rights institutions simply do not fit ecosystem realities.[20] The inability of free-market advocates to produce even one single illustration of free-market environmentalism for a discommodity supports Coddington's contention, following Kapp, that discommodity problems typically require not private property rights, but rights protecting society.

Anderson and Leal acknowledge that to establish a market in pollution, the government must first set aggregate emission limits. The government must also decide on an initial distribution of the ownership of rights to pollute. Only after these two decisions have been made and the appropriate institutional framework has been constructed can the allocation question be addressed (Daly 1999). With sufficient scientific knowledge, a ceiling below environmental carrying capacity can be established. If democratic agreement could then be obtained regarding an institutional framework and an initial distribution of rights, a pollution rights (permits, allowances) scheme like the sulphur allowance system established in the US by the 1990 Clean Air Act Amendments (CAAA) could reduce the costs of bringing emissions down to the safe (ceiling) level. Costs are reduced because permit owners who figure out how to reduce emissions will then need fewer permits and can sell their extra allowances on the market. Innovation pays, and

anyone who can devise a way to reduce sulphur emissions can profit from their innovation. The sulphur allowance scheme effectively created property rights in a discommodity, but such schemes do not fit the definition of a 'free market'. A free market requires (efficient) private property rights as well as freedom of exchange. To work properly, property must be separable from its environment and divisible, so that it can be transferred (exchanged) without consequences to others who have rights or own property.[21] Because the private sector has no incentive to organize and develop property rights for discommodities, public sector involvement is much broader than for commodities.[22] The political process has to define the 'property' (e.g. the right to emit one ton of sulphur in a particular year), determine a schedule that specifies the total number of permits available for each year going forward, and decide how these permits are to be allocated/ distributed.

For most discommodities, it is also necessary to consider local and regional differences in consequences. For example, coal-burning electric generation stations and oil refineries are major emitters of sulphur, subjecting people directly downwind to high concentrations of poisonous sulphur dioxide and acid rain. Under such circumstances, the sulphur allowance cost to the company is often far below the discommodity cost to the community, suggesting that additional mechanisms be implemented. One approach in such 'hot spots' is to overlay the allowance system with direct controls such as requirements to install specific pollution control technologies. Another approach is to add a punitive tax on emissions to fund health care, home air filtration, or mobility of residents to a healthier location. As neoclassical economists argue, 'market-like' or 'indirect,' incentive-based approaches have significant efficiency advantages, particularly over time. But especially when addressing discommodities, we must bear in mind that efficiency is driven by willingness to pay, which is less democratic and less fair the greater the disparity between the 'haves' and the 'have nots.'

Everyone should have a right to healthy air to breathe and water to drink, which Kapp called 'existential minima' (Kapp 1965). We must also remember that clearly specified, efficient property rights are the foundation of efficient markets, and that air and water simply cannot be made to conform to these conditions. If appropriate protections from discommodities are specified and enforced, if social objectives are incorporated into the permit system, and if the initial distribution of rights is fair, then 'market-like' programmes can be a successful component of environmental control and improvement.

On small earth, fewer and fewer economic activities have insignificant consequences for individuals beyond the marketplace, for communities, and for the human prospect. Like his contemporary in the environmental movement, Aldo Leopold, Kapp understood that human survival, let alone property, required an extension of ethics to all members of the land community. As Coddington's essay makes clear, environmental protection on small Earth requires community rights and protections, suggesting a transition away from private property toward common property institutions. Consistent with Kapp, institutionalists emphasize that these common property institutions should be constructed through broadly participatory democratic processes.

Notes

* Thanks to Al Schmid for comments on an earlier draft.

1 This belief is an outmoded habit of thought I have called the 'quiescent Earth mentality' (Swaney 1985). Herman Daly expresses this as an 'assumption of nondisruptability' (Daly 1996: 59).

2 It is this aspect of his work that aligns Kapp with the heterodox academic traditions of social economics and American Institutionalism.

3 Never mind that orthodoxy's concept of value is restricted to 'willingness to pay' as demonstrated by marketplace purchases, or that this extremely limited notion of value is entirely dependent on who has the money to go register their willingness to pay.

4 The US lags several European countries in reforming our national income accounts to more accurately *add* economic activity that contributes to well-being and *subtract* environmental remediation activities and ecosystem degradation that reduce well-being. Probably the best contemporary effort in the US to adjust the GDP to better account for shifted costs is the Genuine Progress Indicator (GPI), which is an outgrowth of Daly and Cobb's ISEW (Index of Sustainable Economic Welfare) (Daly and Cobb 1989). In 1999 a National Academy of Sciences report, *Nature's Numbers: Expanding the National Economic Accounts to Include the Environment,* recommended the development of a separate, parallel set of accounts to monitor the environment. Progress on such reform is unlikely in the current political climate.

5 Although *The Social Costs of Private Enterprise* is a remarkably comprehensive account of the wide scope of social costs in US capitalism, more forceful and concise explanations of the cost-shifting concept are found in later works, especially a pair of *Kyklos* articles (1969, 1970) and a new introduction for the re-issue of *The Social Costs of Private Enterprise* (1971a). (See Swaney and Evers 1989).

6 'Full earth' and 'small earth' are explained more fully below.

7 This democratic process produces evolving values and priorities as the community develops new knowledge and technology to solve its problems. Many Veblenian institutionalists, especially those influenced by C. E. Ayres, who used Dewey's work on democracy to add ethical content to Veblen's 'institutions vs. technology' dichotomy, call this the 'instrumental process' (Ayres 1962). Although Kapp was a consistent advocate for democracy, my reading of his work does not suggest he embraced the 'instrumental process' concept. This is explained more thoroughly below.

8 On empty Earth, the quantity of natural resources that the economy is able to use is limited by knowledge, its application (technology), and the supply of productive equipment (tools, machinery, computers, etc.). That is, economic activity is limited by the availability of human-produced equipment ('man-made capital'). The Earth was empty until quite recently. On our now increasingly full Earth, the quantity of natural resources is limited by nature's abundance, consisting of stocks of nonrenewable resources (e.g. deposits of valuable metals, fossil fuels) and flows of renewable resources (e.g. fish, timber). That is, economic activity is limited by nature's stocks and flows ('natural capital'), making the scale of economies (in relation to nature's capacities) a critical issue (Daly 1996).

9 Although some technologies expand carrying capacity, increasingly they are offset by residuals that disrupt the production of ecosystem services, thereby reducing carrying capacity. This is the principal link between full Earth and small Earth.

10 Earth's 'fullness' has been debated at least since the time of Malthus, but as human numbers have grown exponentially, as Malthus predicted, so has food production, contrary to Malthus' prediction. At the turn of the 20th century, when it seemed that human population was approaching Earth's carrying capacity, German chemist Fritz Haber developed inexpensive nitrogen fertilizer, and agricultural output soared. A few decades later the Green Revolution led to another dramatic increase in agricultural productivity,

allowing the world to feed even more people. Agriculture is now on the cusp of another productivity revolution via genetically modified organisms (GMO). While the full consequences on Earth's ecosystems will not be known for decades, there is a significant nonzero probability that the net effect of GMO organisms in the environment will actually reduce carrying capacity.

11 See, for example, Daly (1968), Wolozin (1971), and Solo (1974).

12 Naive technological optimism wedded to conflation of the environment's resource and life-support functions dates from mid-20th century University of Texas institutionalist Clarence E. Ayres. Marc Tool and Anne Mayhew, prominent institutionalists schooled in the Ayresian tradition, have discarded this interpretation (Tool 1979; Mayhew 1981). Tool advocates 'environmental compatibility' and cites Leopold approvingly. Mayhew cites the developing science of climate change, arguing that this, like all scientific knowledge, is technology that should be further developed and applied to problem-solving (e.g. stabilizing climate). For examples of 'large Earth' Ayresian institutionalists who retain a technological optimism uninformed by ecology, see Lower (1987) and DeGregori (2002). For an institutional discussion of the high-profile 'boomster-doomster' debate between biologist Paul Ehrlich and business economist Julian Simon, see (Swaney 1991).

13 The property rights discussion in this section is confined to the neoclassical position, rooted in private property and the Coase Theorem. Both Veblen-Myrdal institutionalists and 'new' institutionalists, including Coase, have explored other institutions and conditions.

14 The precautionary principle states that 'ecosystem services can be used on a truly sustainable basis only if the intensity of use is substantially less than the apparent maximum' (Marten 2001: 153).

15 For Christians, this means that markets are indifferent to destroying God's Creation.

16 Open access is a condition of no property, and is not synonymous with common property, which is a group of people managing a resource. Conflating these two entirely different conditions has led to many errors. (See Swaney 1990 and 2003).

17 This is not to argue that we lack mechanisms of social control, which have been in place in the US since the early 1970s. In the case of air pollution, the 1990 Clean Air Act Amendments directed the US EPA to impose additional controls whenever a new air health hazard is established scientifically.

18 An excellent critique of *Free Market Environmentalism* is provided by Herman E. Daly: 'Free market environmentalism: turning a good servant into a bad master', reprinted in *Ecological Economics and the Ecology of Economics*, 1999. (Northampton, MA: Edward Elgar). This discussion parallels Daly's discussion, borrowing liberally on occasion.

19 See Swaney 2003 for explanation and citations of some of the important literature.

20 Another alleged application of free-market environmentalism involves deteriorating visibility in Grand Canyon National Park. Air pollution was known to decrease visibility, but the source of the increase in concentration of small particles in canyon air was not established until a tracer material added to the coal used to generate electricity downwind from the canyon was detected in canyon air. How this example illustrates free-market environmentalism is more than a little puzzling, as the only private parties involved were the owners of the coal-burning electricity-generating facilities. The Grand Canyon is a World Heritage site that, if it can be said to be 'owned' by anyone, is owned by the US government or, collectively, the American people, but with stewardship responsibility for all of Earth's citizens, now and in the future. A non-governmental organization, the Environmental Defense Fund (now Environmental Defense), played a critical role in negotiations between the power company and several government agencies. The negotiations eventually led to an agreement wherein the power company committed to a gradual implementation of additional pollution control technologies.

The tracer element technology helped to identify a major cause of Grand Canyon haze, but the government still has to define rights, allocate rights, and enforce rights. Furthermore, the rights being protected, while benefiting specific individuals, are clearly not private property rights. Private property and free markets are both conspicuously absent from this 'example', so to suggest that private property was the solution is specious.

21 Ownership also requires a practical ability to manage, including exclusion of others from encroaching on, using, or degrading the property. For a cogent explanation of efficient property rights, see Tietenberg 1984, pp. 39–40.

22 Although Coase himself objects, 'free market' advocates tell us that the Coase Theorem says that ordinary citizens will band together and bargain with polluters for their clean air and water, but this seldom happens (except through the political process) because those harmed by the pollution are usually poorly informed, not knowing what the pollution is doing to their health, or harmed by a small amount relative to the costs of organizing and bargaining. That is, transaction costs are very high and the Coase Theorem does not apply. Coase's point has always been that transactions costs are significant, and his research program, complementing Douglass North's, is to figure out how transactions costs can be reduced.

References

Anderson, T. L. and Leal, D. R. (2001) *Free Market Environmentalism*, revised edition, New York, NY: Palgrave; first edition published in 1991 by Pacific Research Institute, San Francisco.

Ayres, C. E. (1962) *The Theory of Economic Progress*, New York: Schocken Books; originally published 1944.

Boulding, K. (1993) 'The Economics of the Coming Spaceship Earth', in E. H. Daly and N. Kenneth (eds) *Valuing the Earth: Economics, Ecology, Ethics*, Townsend, Cambridge, Massachusetts: MIT Press; originally published 1966.

Bromley, D. W. (1978) 'Property Rules, Liability Rules, and Environmental Economics', *Journal of Economic Issues* 12, 1: 43–60.

Brown, L. (2001) *Eco-Economy*, New York: Norton.

Coase, R. (1960) 'The Problem of Social Cost', *Journal of Law and Economics* 3, 1–44.

Coddington, A. (1970) 'The Economics of Ecology', *New Society* 5: 595–597.

Crocker, T. (1996) 'Structuring of Atmospheric Pollution Control Systems', in H. Wolozin (ed.) *The Economics of Air Pollution*, New York: Norton.

Dales, J. H. (1992) 'The Property Interface,' in A. Markandya and J. Richardson (eds) *Environmental Economics: A Reader*, New York: St. Martins Press.

——. (1968) 'On Economics as a Life Science', *Journal of Political Economy* 76, 3: 392–406.

Daly, H. E. (1996) *Beyond Growth*, Boston: Beacon.

——. (1999) *Ecological Economics and the Ecology of Economics*, Northhampton, Massachusetts: Elgar.

Daly, H. E. and Cobb, J. B. Jr. (1989) *For the Common Good*, Boston: Beacon Press.

DeGregori, T. R. (2002) *The Environment, Our Natural Resources, and Modern Technology* Ames, Iowa: Iowa State Press.

Diamond v. C., 447 US 303, US 1980.

Hanley, N., Shogren, J. F. and White, B. (1997) *Environmental Economics In Theory and Practice*, New York: Oxford University Press.

Hawken, P. (1993) *The Ecology of Commerce*, New York: Harper Collins Publishers, Inc.

Kapp, K. W. (1961) *Toward a Science of Man in Society-A Positive Approach to the Integration of Social Knowledge*, The Hague: Martinus Nijhoff.

——. (1965) 'Economic Development in a New Perspective: Existential Minima and Substantive Rationality', *Kyklos* 18: 49–79.

——. (1971a) *The Social Costs of Private Enterprise*, New York: Schocken Books; originally published 1950.

——. (1971b) 'Social Costs, Economic Development, and Environmental Disruption', in J. E. Ullmann (ed.) *The Humanization of the Social Sciences*, Lanham: University Press of America.

——. (1976) 'The Open-System Character of the Economy and Its Implications', in K. Dopfer (ed.) *Economics in the Future*, Boulder, CO: Westview Press.

——. (1977) 'Environment and Technology: New Frontiers for the Social and Natural Sciences', *Journal of Economic Issues* 11, 3: 534–35.

Kuhn, T. (1962) *The Structure of Scientific Revolutions*, Chicago: University of Chicago.

Leopold, A. (1987) *A Sand County Almanac*, New York: Oxford University; originally published 1949.

Lower, M. (1987) 'The Concept of Technology within the Institutionalist Perspective', *Journal of Economic Issues* 21, 3: 1147–1174.

Marten, G. G. (2001) *Human Ecology*, London: Earthscan Publications

Mayhew, A. (1981) 'Ayresian Technology, Technological Reasoning, and Doomsday', *Journal of Economic Issues* 15, 2: 513–20.

McDonough, W. and Braungart, M. (1998) 'The Next Industrial Revolution', *The Atlantic Monthly* 282, 4: 82–92.

New Jersey Department of Environmental Protection (2004). *Indiana Bat*, retrieved December 12, from http://www.nj.gov/dep/fgw/ensp/pdf/end-thrtened/indianabat.pdf.

North, D. (1991) 'Institutions', *Journal of Economic Perspectives* 5, 1: 97–112.

Pigou, A. C. (1970) *The Economics of Welfare*, London: Macmillan.

Simon, J. L. (1981) *The Ultimate Resource*, Princeton, New Jersey, USA: Princeton University.

Solo, R. (1974) 'Problems of Modern Technology', *Journal of Economic Issues* 8, 4: 859–76.

Swaney, J. A. (1985) 'Economics, Ecology, and Entropy', *Journal of Economic Issues* 19, 4: 853–65.

——. (1990) 'Common Property, Reciprocity, and Community', *Journal of Economic Issues* 24, 2: 451–462.

——. (1991) 'Julian Simon versus the Ehrlichs: An Institutionalist Perspective', *Journal of Economic Issues* 25, 2: 499–509.

——. (2003) 'Are Democracy and Common Property Possible on Our Small Earth?', *Journal of Economic Issues* 37, 2: 261–290.

—— and Evers, M. A. (1989) 'The Social Cost Concepts of K. William Kapp and Karl Polanyi', *Journal of Economic Issues* 13, 1: 7–33.

Tietenberg, T. (1984) *Environmental and Natural Resource Economics*, Glenview, Illinois: Scott, Foresman and Company.

Tool, M. R. (1979) *The Discretionary Economy*, Santa Monica, California: Goodyear Publishing.

Vitousek, P. M., Ehrlich, P. R., Ehrlich, A. H. and Matson, P. A. (1986) 'Human Appropriation of the Products of Photosynthesis', *BioScience* 34, 6: 368–73.

Wilson, E. O. (1984) *Biophilia*, Cambridge, Mass.: Harvard University Press.

Wolozin, H. (1971) 'Environmental Control at the Crossroads', *Journal of Economic Issues* 5, 1: 26–61.

Zimmermann, E. W. (1951) *World Resources and Industries*, New York: Harper and Brothers.

6 Improved allocation through environmental taxes? Theory and reality: the example of Germany

Gustav M. Obermair and
Lorenz Jarass

> 'Society must be prepared to transform social costs into private costs through political intervention.'
>
> Kapp (1950)

Introduction: objectives and instruments

It is certainly true that basic judgements on the standards of and the requirements for sustainability, i.e. for life on this earth, lie outside the realm and scope of the 'disembedded' economic system – and hence to a large extent also outside the focus of the prevailing economic theory. Yet the continuing deterioration of the environment in many parts of the world, the imminent danger of a drastic change in climate and the destructive effects of continual high unemployment have created a new interest for those economists who have since the early 20th century analysed these evils inherent in market economies and proposed some solutions. One possible instrument to fight pollution and increase employment will be the main subject of this article: environmental taxes[1].

Unfortunately some of the early proposals like the Pigou tax or the Coase contracts have for a long time just remained beautiful toys for the kind of mathematicians that are called theoretical economists.

In reality we are confronted not with the standard textbook example of one paper mill upstream and one fisherman downstream who risks losing his job, but with millions of polluting activities with an impact on 400 million people in Europe alone and with 20 million unemployed in the EU15.

We do not know and have no way to measure the marginal social cost of each of such activity as required by theory in order to determine the optimum Pigou tax rate, nor can we set up Coase contracts between millions of partners.

Kapp's work, on the other hand, was always aimed at a realistic political concept of economics beyond mere mathematical games; yet his general concept of social costs was to a large extent ignored by mainstream social science.

But may be this is changing: the threat of global warming, the political instability resulting from high structural unemployment, the abundance of financial capital not finding interesting investment opportunities, and, last but not least, the foreseeable end of cheap oil and gas – all this acts as a pressure on governments to reduce the dependence on fossil fuels and to pave the way for new, labour-intensive worldwide investments in renewable energy, energy saving and all kinds of solar technology.

Certainly the advocates of such policies are growing in number all over the globe; yet they face enormous obstacles: the resistance of traditional big industries, often in collusion with governments and sometimes even with labour unions, but also the more subtle mental reservations of the common economic actors described in the next section.

Rational man – irrational society or: the prisoners' dilemma

In many cases the individual economic subject has insufficient information about his cost situation with respect to public goods like environmental quality or adequate levels of employment. Two simple examples may serve as an illustration for such situations:

'I would, of course, not dump my garbage into the ditches. But, I suspect, all the others will continue to do it. Hence not doing it myself, I will only hurt my own interests without helping to keep ditches clean.'

'I would, of course, hire some more workers. But, I suspect, all my competitors will invest in production automats. Hence not investing that way myself, I will only hurt my own interests without improving employment.'

In abstract language: on the basis of his insufficient information the individual has no better choice than to maximize his private benefit, even knowing that he might thus hurt the benefit of all. A situation like this is what game theory describes as 'The Prisoners' Dilemma'. In other contexts it has simply been called 'harmful competition'! It is also the result of unfair or harmful competition when governments do not tighten their environmental protection laws, because other countries with (even) lower standards would gain a cost advantage for their products.

> In all these cases the defence of the public goods that are not covered by the price mechanism of the market requires institutional solutions, some kind of regulation organized by a superior authority. Here Kapp's admonition, quoted as the motto above, comes in: 'Society must be prepared to transform social costs into private costs through political intervention'
>
> (Kapp 1950)

Essentially two types of instruments are available for such political intervention:

- Normative (non-fiscal) instruments, i.e. rules with limit values and sanctions.
- Price instruments by government regulation, e.g. tax incentives and disincentives.

In many cases the two types of instruments should and do in fact act together as in the following case:

A (world-wide) trade in pollution licences, in particular with respect to CO_2 emissions, as an instrument against the greenhouse effect, is generally advocated by the leading climatologists (Schellnhuber 2003). It may be described as a combination of normative and price instruments: after strictly controlled upper limits for greenhouse gas emissions have been fixed by international agreement for each national economy (normative aspect), the price that the high-emission industrialized countries have to pay to low-emission developing countries for additional emission rights follows supply and demand (market aspect). The revenue for the poor countries ought to be invested (again by supranational normative control) mostly in CO_2-reducing technology.

Does 'global rationality' as a prerequisite for the completion and supervision of such treaties have a chance? Or will islands be drowned in rising oceans and hundreds of millions of people starve or fill new refugee camps before such comprehensive international agreements can be pushed through? These are questions that will not be pursued further in this essay. Instead we return to what could be achieved on a national or – at best – European level.

Norms are of course indispensable and quite effective for the protection of the biosphere: limits on the emission of toxic or noxic agents, radioactive material etc. into air, water and soil. Ideally they should be internationally standardized (also to avoid the above-mentioned unfair competition). At least for radioactive isotopes this has been achieved through UN institutions. In many other cases we are far from such standardization. Moreover norms alone may also prove insufficient and may indeed impede technical innovation: once the prescribed values are achieved there is no incentive for further improvement and for the development of technical alternatives. As a remedy against unemployment normative solutions have, at least if used on a large scale, not turned out very helpful – meaningful jobs are not created by norms on levels of employment; government-financed meaningless work where there is no demand for the results of the work, practiced not only formerly in the COMECON states, but also in some areas of Western Europe, has been a disaster both for public finance and for the work ethic of the population.

This is where the second instrument comes in: price instruments furnish additional information. The cost expectation for socially desirable economic activities can be lowered, that for socially undesirable behavior increased, e.g. by appropriate taxation. Taxes should hence be fixed at such levels that socially or environmentally preferred standards become competitive on the market (hence called standard-price taxes).

There is a counterargument against such tax instruments which is quite popular throughout the political spectrum from right to left: government tax policy should always remain strictly fiscal and not begin to interact with the economy (in German: 'Nicht steuern mit Steuern!').

Now, whoever argues like this must be stricken with blindness: the fiscal policy of governments is not in a state of economic innocence, on the contrary, in the last

25 years the tax policy of most continental European states has interfered strongly with the allocation of production factors. They have shifted much of the tax load away from capital income and wealth onto the production factor labour, (see Jarass/Obermair 1999). For example, Germany: the effective tax rate on both labour and capital income was around 30 per cent up to 1980; in 2002, however, the load of taxes and contributions on labour amounted to nearly 40 per cent, that on capital income to less than 20 per cent. Taxes on natural resources and environment remained low and amounted to less than 3 per cent of GDP in nearly all countries. Expressed in real purchasing power the price of gasoline was lower in 2000 than in the 1960s before the first oil price crisis. The most polluting activity – air traffic – is still subsidized by the tax exemption for kerosene etc.

And the price mechanism of the market did in fact work: as the investment in labour-saving technologies is subsidized whereas the load of taxes and contributions on labour went up, so did unemployment; as the real price of fossil energy went down, so did the efforts to curb its consumption.

Thus, even if inadvertently, government tax policies have to a large extent contributed to the present socially undesirably allocation of the production factors labour, capital and natural resources: 'rational man – irrational society!'

To reverse this trend at least to some degree the new German government has tried since 1998 to introduce a serious environmental tax.[2] Before we describe this project and its positive and negative consequences, some matters of principle will be taken up: the problem of the monetarization of the social costs of environmental damage and the related problem of fixing the appropriate rates of environmental taxes in such a way, that they may lead to a new 'equilibrium' between the forces of the economy and the requirements of sustainability.

Monetarization and the quest for the 'equilibrium level' of tax incentives and disincentives

Talking of costs, we talk about money; talking of social costs, we talk about monetarization, we must write out a price tag. But what is the monetary value, of course at today's market price, of all the specific damage, deterioration or loss that human activities – industry, traffic, energy conversion and our daily life style – do to the biosphere, to the habitat of plants and of animals of all kind including humans?

For the social cost of air pollution the German Federal Agency for the Environment ('Umweltbundesamt') has arrived some years ago at such a figure (€ 45 billion per annum) by adding up a rather inhomogeneous set of entries: financial loss in forestry due to dying trees, reduced yield in agriculture, health expenses and loss of working hours due to increasing respiratory diseases, repair costs for eroded stone or steel and concrete both on historical buildings and on modern technical construction etc. But such lists can hardly be complete: when a sinking tanker spills thousands of tonnes of oil one may be able to estimate the financial loss for fishery, but how many seabirds will be killed and what is their monetary value?

And the social cost of continuous unemployment – can it be measured in terms of the social assistance payments for those who are not or no longer covered by unemployment insurance? What about depression, alcoholism, neglected children or the development of slums?

What this brief list of examples shows is that monetarization of social costs may at best result in a crude first approximation that gives an indication of the order of magnitude of the problems involved. Moreover, even if one could do better, it would not really be helpful: in order to determine the rates of tax incentives and disincentives that would compensate the social costs, one would also need all kinds of elasticities with respect to the deployment of production factors. In the end one would be caught up in mere mathematical games based on some rather unrealistic assumptions:

- That the outputs of an economy, the social product, the level of employment, the quality of life etc. are continuous and differentiable functions (in the mathematical sense) of the inputs in terms of relative factor prices etc.
- That one is never far away from some kind of equilibrium.
- That the velocities of adjustment are very high and all transaction cost very low and so on.

So let's leave neoclassical idealizations and turn briefly to a somewhat different point of view that originates in the theory of optimization.

The hard way towards a greener valley

There seems to be a broad consensus at least in European societies that one should fight unemployment and struggle for the improvement of the environment. Yet most concrete measures to improve the situation in this respect are confronted with objections, impediments or with vetoes from influential groups. A paradoxical situation: something that a large majority holds to be desirable finds active support at best from small minorities.

One could describe this as a case of split collective consciousness or quote the 'innate conservatism' of human nature, but such psychological explanations, even if plausible, are not very helpful, because they deal with symptoms, not with causes. The cause of innate conservatism lies in the fact that an established economy finds itself, due to the negotiating margin of its participants, at any time near some dynamic or 'flow' equilibrium, i.e. a local or relative optimum (=maximum with constraints) of the benefit function of its members. By virtue of the negotiation process leading to such equilibria they are stable against not-too-big perturbations: after an excursion the system is pushed back to a state close to the original one (like a ball always rolling back to the lowest directly attainable point of a valley) because a deviation from the equilibrium results in decreased benefits for most participants. The unexplained 'innate conservatism' thus appears as a completely rational reaction based on long collective experience.

So is there no way out? Are we really trapped in the present mode of production where the return on capital is the prime motor of the economy, where the means of production are selected to maximize productivity at the expense of employment and environment? In the attempts to formulate answers to this crucial question we invoke an analytical tool that has been developed in the last decades: optimization theory.

Interlude on optimization

There exists now a vast literature on the mathematics of optimization, covering a wide range from Operations Research to Statistical Mechanics, from the traveling salesman problem (Lin/Kernighan 1973), to the physics of glasses (see Kirkpatrick *et al.* 1983). (For an updated list of references see Beutner *et al.* 2001.) The two prototype examples of this range are in fact the traveling salesman, who tries to visit N-many clients in one tour with the shortest or fastest possible route (while maybe trying to visit the most important clients at short regular intervals) and the arrangement of N-many molecules in an amorphous material like glass where the attractive forces and the hard core repulsion between the atoms bring about structures that represent a relative minimum of the total energy.

It is characteristic of such systems with a very large number of independent state variables (N steps in the tour, N molecular positions in the sample) that besides the true absolute minimum of the benefit function (of the total length of the tour, of the total potential energy of the sample) there are many 'good' relative minima that represent excellent approximations of the true minimum. However, in the space spanned by the state variables (the sequence of individual steps, the positions of individual molecules) these different relative minima are as a rule far away from each other and separated by regions with very unfavorable values of the benefit function. It is like in a huge mountain range: there are many deep, warm, green valleys, separated by high saddles; if a group of people want to go from 'their' valley to another warmer and greener one, they all have to go through a lot of hardship.

This structural insight, not the (futile) attempt to do any quantitative computations, motivates this digression: we may gain some heuristic arguments that explain the evident stability of an untenable state of affairs.

Modes and means of production

The adjective 'capitalist' has more or less vanished from public debates and even from the social science vocabulary; the traditions of the University of Trento and the title of this conference represent a remarkable exception. For the general public the word 'capitalism' is a no-no word; for most economists the capitalist mode of production represents the 'natural' order of human affairs or the final destiny of history and is therefore not even worth being explicitly mentioned. From this point of view a market economy cannot be anything but a capitalist economy and hence there is only one equilibrium which is – apart from some minor deficiencies – simply identical with the present state of affairs.

But this is not true – even the cursory sketch of optimization theory given above indicates that any system with a large number of variables must exhibit several (locally stable) relative optima of the total social benefit, all maintained under the same conditions of free-market exchange; only the composition and weight of individual goals entering the 'benefit function' will differ greatly between these different equilibria. There is one, which we know, in which the predominant social goal is to maximize the yield for the givers of capital; there is another free-market equilibrium: Vogt has shown its hypothetical existence and stability under the usual mathematical assumptions of neoclassical theory (Vogt 1986) where the quality of life of the working class dominates the mix of individual benefits and a similar one (Jarass *et al.* 1989) in which also the quality of the ecosphere – call it 'sustainability' – is given priority over profit (see also Daly 1977).

Yet here we are: even given a 'volonté generale' to realize a sustainable economy, how do we get from here to there? The new 'greener valley' lies far away, there is no easy way to get there due to the very property of local stability: small- or medium-size changes of the variables, in our case of the relative deployment of production factors, will only lead to non-optimal conditions both for the profit rate of capital and for the quality of life for labour and will not yet produce an essential ecological improvement. Here we have a model explanation of the paradoxical situation described above: the new equilibrium cannot be reached spontaneously. Attempts to get there by drastic political measures meet broad disapproval based on the (correct) assessment that both labour and capital will for a long time suffer considerable reductions of their benefits.

Experienced inhabitants of mountainous regions know how to react: do not go over the tops, but over the lowest saddle and decrease the efforts to overcome it. This is where environmental taxes come in: fiscal incentives in favor of ecological and labour-intensive technologies, shifting the tax load from the production factor labour to increased costs for the input of environmental quality, may ease the transition and improve, even in the short term, the benefits of ecological means of production. This, then, is the expected effect in our metaphorical language: the profit-oriented optimum becomes less profitable, the saddle less steep and the green valley even greener. To do this job, the pace of the cost shift must follow typical economic and technological adjustment times; we even have an experimental clue for such adjustment times: after the oil price increase in the 1970s, the decreased fuel consumption of new cars and the improved efficiency of industrial and public utility energy conversion has nearly compensated this increase within five years. (Don't underestimate the versatility of capital.)

The German 'ecotax': a large-scale socio-economic experiment

The tax shift proposed around 1990 by the former EC President Delors and elaborated in the *White Book* (EC 1994) and the subsequent studies was never realized on the European level. Instead, targets for the reduction of green house gases from 1990 to 2008/12 were set for each of the EU15 countries which would result in an overall reduction of 8 per cent, with higher targets for some developed

industrialized member states (−21 per cent for Germany and Denmark, −12.5 per cent for the United Kingdom) and increased limits for some southern European countries (+15 per cent for Spain and +27 per cent for Portugal). This target sharing is to take into account the differing geographic, climatic, economic and social points of departure. The specific instruments for meeting these targets are left for the national governments to decide.

More specific plans were developed step by step in Germany since the early 1990s, mostly by NGOs and research institutes. In 1994 Greenpeace Germany commissioned a first comprehensive study on the measures and results of an 'Ecological Tax Reform (ETR)'. The scenario was: increase the tax on carburants by 7 per cent annually over a ten-year period, increasing the consumer price by a total of more than 50 per cent, while keeping the total state revenue constant by redistributing the increment to the citizens by a lowering of taxes and, in particular, social security contributions, on labour – Delors' concept of a tax shift.

The study (DIW 1994), performed by the German Institute of Economic Research gave a projection of the main effects of these measures:

- The overall reduction of primary energy consumption would be 20 per cent over the ten year period, that of CO_2 emissions 21 per cent.
- Between 400 and 800 thousand new, additional jobs could be created due to reduced labor costs and the price incentives for energy conservation and renewable energy technologies.

This is the 'double dividend' hypothesis; its practical test, 1998–2003, will be described below.

When spokesmen of the Green Party advocated a final price of DM 5 or € 2.50 for the litre of gasoline or diesel oil in 1998 there was a public outcry (this price may indeed not be to far from the actual social cost of automobile and truck traffic), but the predicted election disaster did not occur and the newly elected red-green government was able to pass the ecotax-legislation and start a (considerably downsized) program of ecological taxation:

- Beginning in April 1999 the tax on carburants went up in five yearly steps of € 0.0307 per litre.
- An extra tax on consumer electricity was introduced: + € 0.01 per kWh_e in 1999 and additional steps of + € 0.0026 per kWh_e each year from 2000 to 2003.
- Single increases in 1999 on natural gas: + € 0.0016 per kWh_{th} and on light heating fuel oil: + € 0.002 per litre (equivalent to approx. € 0.02 per kWh_{th}).

But, of course, exemptions had to be made for powerful lobby groups, for energy-intensive industries and for agriculture.

On the other hand the extra revenue estimated at € 14 billion for 2002 (i.e. about 3 per cent of total tax revenue excl. social security contributions) was redistributed in the form of a reduction of employers' and employees' social security contributions by overall 1.7 percentage-points; the state pension insurance rates

could be lowered for some time from 20.3 per cent to 19.1 per cent of gross wage; and an additional government expenditure of € 200 million could be provided for renewable energies.

Thus it appears that, in spite of the exceptions and exemptions (a reduction of € 4 billion of revenue with respect to the full rates now paid only by consumers and small industries), the double dividend concept did in fact work for some time:

- Emissions from traffic decreased by about 2 per cent annually since 2000; the total CO_2 emission decreased by nearly 20 per cent since 1990 for a variety of reasons, mostly the deindustrialization of Eastern Germany after 1990, but lately also due to the ecotax effect.
- Against a long-term downward trend the number of passengers in the public transport system began to go up again with a total increase of 3 per cent since 1999.
- The increased demand for energy-efficient cars and for improved energy saving in households (heating, cooling, thermal insulation etc.) shows a favourable consumer reaction.
- It is estimated that up to 250,000 additional jobs will have been created by 2003 due to reduced labour costs and to investments in energy saving.

Thus it may appear that the ecological tax reform is a real success story. Is it really?

The figures just quoted do indeed indicate some significant allocation effects; even more important a qualitative change of long standing trends, in particular with respect to energy consumption.

In spite of this well-documented impact public opinion on this reform is rather negative; populist mass media have done their best to counteract against a government publicity campaign; they emphasize the increased gasoline and electricity bill and neglect to mention the relief in social security contributions ('Freedom of press is the freedom of rich people to have their opinion printed', Kurt Tucholsky, around 1925). Unfortunately it is also true that the lowest income group has the lowest relief and the highest relative extra cost with respect to their disposable income.

There is also severe criticism from environmentalists: some of the worst polluters have not even been touched – German hard coal is still highly subsidized, air traffic kerosene still untaxed. It can be argued that a much higher fraction of the additional revenue ought to be directly invested for environmental recovery.

Last but not least it has been argued that the temporal relief on social security contributions has only obscured the tendential bankruptcy of the existing system and impeded the efforts for the necessary fundamental reform, by which all tax payers, not only the wage earners, would have to contribute to the maintenance of a civilized society.

The saddle towards the greener valley has not yet been surmounted, but we may, with Albert Camus, 'imagine Sisyphus as a fortunate person: the struggle against mountains may fill human hearts with happiness'.

Notes

1 For the authors' contributions to this subject cf. Jarass/Obermair (1993), and Jarass/Obermair (1999); for social cost arguments in infrastructure planning see e.g. Jarass *et al.* (1989).
2 A stocktaking of taxes in the EU that could be labelled 'environmental', compiled by the authors in 1996 for the European Commission, Jarass/Obermair (1996), showed that the economic impact of these relatively low taxes, often restricted to narrow sectors and a few special products, cannot have been too high.

References

Beutner, J., Bauer, G., Obermair, G. M., Morgenstern, I. and Schneider, J. (2001) 'Optimization of the Time-dependent Traveling Salesman Problem', *Physical Review* 64, 036701.

Daly, H. E. (1977) *Steady State Economics,* San Francisco: W. H. Freeman and Company.

Deutsches Institut für Wirtschaftsforschung (1994) *Wirtschaftliche Auswirkungen einer Ökologischen Steuerreform*, Berlin: German Institute for Economic Research (DIW, www.diw.de).

European Commission (1994) *White Book on Growth, Competitiveness and Employment*, Brussels: European Commission.

Jarass, L., Nieβlein, E. and Obermair, G. M. (1989) *Von der Sozialkostentheorie zum umweltpolitischen Steuerungsinstrument – Boden- und Raumbelastung von Hochspannungsleitungen,* Baden-Baden: Nomos-Verlag.

—— and Obermair, G. M. (1993) *More Jobs, Less Pollution: A Tax Policy for an Improved Use of Production Factors,* Report on Tax Incentives in the Framework of Statutory Levies, commissioned by the European Commission, DG XXI, revised version in *Natural Resources Tax Review*. *Tax Analysts* (1994) Arlington, Virginia: part I: 1911–57; part II: 2219–42; part III: 23–38.

—— and Obermair, G. M. (1996) *Manual – Statistics on Environmental Taxes,* commissioned by European Commission, DG XI – Environment, Nuclear Safety and Civil Protection and DG XXXIV – the Statistical Office of the European Communities Online (www.jarass.com).

—— and Obermair, G. M. (1999) *More Jobs, Less Tax Evasion, Cleaner Environment: Options for Compensating Reductions in the Taxation of Labour – Taxation of Other Factors of Production,* Report to the European Commission, DG XXI Online (www.jarass.com).

Kapp, K.W. (1950) *The Social Cost of Private Enterprise*, Cambridge, Mass: Harvard University Press.

Kirkpatrick, S., Gelatt, C. D. and Vecchi, M. P. (1983) 'Optimization by Simulated Annealing', *Science* 220: 671–680.

Lin, S. and Kernighan, B.W. (1973) 'An Effective Heuristic Algorithm for the Traveling Salesman Problem', *Operations Research* 21: 498–516.

Schellnhuber, H. J. (2003) 'Gefahr für die Zivilization (Danger for Civilization*), Der Spiegel,* Chronicle of 2003, Hamburg.

Vogt, W. (1986) *Theorie der kapitalistischen und einer laboristischen Ökonomie*, Frankfurt/New York: Campus Verlag.

7 Unemployment as a social cost[1]

Paolo Ramazzotti and
Marco Rangone

Introduction

Over the last decade, unemployment has increasingly been dealt with by resorting to active labour market policies. Economists and policy makers have focused on measures (such as training, incentives and information) that are expected to raise the efficiency of the market. Even though unemployment relief systems devised under the (post-) Keynesian framework are still at work alternatives have been introduced which consist of making welfare help people to gain self-sufficiency through work. Welfare-to-work or workfare – a milestone in President Clinton's and in Prime Minister Blair's reforms – provides support to the unemployed only insofar as they are willing to accept a job.

Despite their alleged superiority to Keynesian measures, these programmes do not prove to be particularly successful. We believe the reason lies in their theoretical underpinnings. They assume that institutional constraints prevent labour supply from adjusting to the (technologically determined) requirements of labour demand. We argue that institutions, far from being constraints, are major determinants of supply and demand. When a mismatch occurs, this is due to the different institutional requirements underlying supply and demand. Unemployment is not determined by institutions, it is determined by the mutual inconsistency of institutions. More precisely, it is caused by business requirements that are inconsistent with social efficiency.

We base our discussion of unemployment on Kapp's notion of social costs. We contend that unemployment is one such cost. Kapp defines social costs as 'all direct and indirect losses sustained by third persons or the general public as a result of unrestrained economic activities' (Kapp 1963: 12). Although this definition recalls that of a conventional externality (Feldstein 1978), it is in no way 'a minor and exceptional disturbance' but 'a characteristic phenomenon of the market economy' (Kapp 1963: 9). The difference lies in what Kapp views as key features of the economic system.

First, contrary to conventional economics, agents are not atomistic, they interact and, above all, they can change the rules and the payoffs of the economic 'games' they play: this is the upshot of Kapp's notion of *cost shifting*[2]. Second, interaction also occurs between the economy and the surrounding – natural and social – environment. It generates feedbacks and an overall process of

cumulative causation whereby 'the 'economic' and the so-called 'noneconomic' are intrinsically interrelated' (Kapp 1963: 12)[3]. Some of the implications of this notion of social cost will emerge as we carry out our analysis. They include distribution, democracy, measurement and methodology. Here, we only wish to point out that, owing to the systemic openness of the economy, a metric (relative prices) that is restricted to the market cannot function as a general unit of measurement when 'noneconomic' (including ethical) issues are concerned[4]. The value judgement required to assess social efficiency may be frustrating for whoever seeks an 'objective' unit of social accounting. It is, nonetheless, a necessity[5].

The point of departure for our analysis is a brief discussion of active labour market policies. In the section that follows, we argue that this kind of policy implicitly assumes that 'working for a market wage' is a socially relevant value, independently of the specific features a job has. Quite to the contrary, we contend that when individuals look for a job, they generally want to take into account its specific features. In section three we discuss this issue and argue that when the social evaluation of jobs (in short, occupational status) is taken into account, the usual economic metric, i.e. money, will not serve its purpose. Status cannot be traded, it usually is complementary to income, it determines lifestyles and life possibilities. Moreover, it cognitively influences the way preferences – about or related to jobs – are formed and it provides differential social resources that affect future choices and courses of action.

These arguments lead us to suggest, in section four, that attempts to improve one's status (e.g. through social mobility) may be positively valued, regardless of their partly ceremonial substance and of social scarcity problems. This raises a problem, however: labour demand requirements do not necessarily fulfil the aspirations of workers; different interests and goals are at stake.

Labour demand requirements are generally seen as binding because they reflect efficient behaviour and technical constraints. These aspects are dealt with in section five. We dispute these claims on two grounds. First, business behaviour may be driven by concerns other than competitiveness; more specifically, distributive and strategic concerns may underlie the division of labour as well as the level of employment. Business 'efficiency' cannot be taken to be a measure of social efficiency. Second, we show the institutional nature of technological and scientific research. In particular, technology is not independent of business goals. Both science and technology are – either explicitly or tacitly – associated to business goals. Technology cannot therefore be used as a benchmark to assess the efficiency of business conduct.

In section six, we suggest that Sen's notion of capabilities may constitute an appropriate benchmark to assess the social efficiency of the economic system. Based on this premise, we argue that unemployment should be viewed as a cost in itself, inasmuch as it reveals that important aspects of wellbeing have been endangered. Accordingly, we are less interested in the negative consequences of unemployment than in the problems that underlie the institutionalized working of capitalism.

This leads us to a few policy implications, which are discussed in section seven. Owing to the inconsistency between the requirements of workers and of business, we argue that a value judgement is required. The 'capabilities benchmark' leads us

to stress the importance of *freedom to choose how to conduct one's life*. Acting in favour of freedom involves the understanding of how business strategies affect learning patterns and available choice sets. It also involves the assessment of policy issues – such as cooperation between the scientific community and business, scientific freedom, educational goals and their institutional implementation, and unemployment relief systems – which may influence the relation between business strategies and social learning.

Tackling unemployment?

From economic growth to individual employability

Since the Eighties economic theory has changed the way it explains unemployment. Attention has shifted from external determinants to the internal functioning of labour markets (Blanchard 1999; Henry and Snower 1996). In the post-war boom period up to the Seventies external, mainly macroeconomic shocks were imputed. Unemployment was claimed to arise when aggregate demand does not allow labour supply to be fully employed in productive activities. This Keynesian view implied that unemployment ought to be tackled by increasing or stabilizing aggregate demand through fiscal and monetary measures, while social damages could be addressed by public intervention through protective measures directed to individuals and families.

In recent years, this view has increasingly been challenged. Focus is on the 'rigidities' that affect the competitive working of the labour market: unemployment occurs *because* the labour market cannot operate as it should. As a consequence, policy concerns also changed. Mainstream analysts point to two obstacles: incomplete information and distorted incentives.

Information problems may entail high search costs; this produces inefficiencies and perhaps involuntary (e.g. frictional) unemployment (Lindbeck and Snower 1986). Optimum decisions are taken if calculation of utility is based on objective probabilities of different final states; even imperfect information, if used at its best, leads to individually and socially rational behaviour, in the conventional view. Yet, inasmuch as subjective probabilities do not coincide with objective ones, individually rational behaviour may lead to socially unpleasant results. For instance, suppose that skilled individuals reckon that they can find another job, suited to their qualification, in a short time; consequently, they may assign to temporary unemployment a greater value than low-skilled employment. If their conjectures are incorrect, involuntary unemployment occurs.

Distorted incentives occur because of Keynesian-style unemployment benefits[6]. Subsidies increase the reservation wage of the unemployed who will be less willing to accept existing jobs at the current market wage rate. Long-lasting reliance on unemployment benefits reduces motivation to work and impairs self-confidence[7]. As a consequence, duration of unemployment will certainly increase, and probability to exit unemployment will be reduced (Layard, *et al*. 1996; Blanchard 1999). Moreover, benefits are also seized by those who would have exited unemployment: some people may voluntarily postpone exit in order

to hold benefits. This produces inefficiencies and raises, in times of resource shortage, moral issues[8].

The solution to these problems lies, according to most observers, in making the labour market function properly. More specifically, the goal of welfare policies ought to shift from protection of living standards to increased responsibility of all parties. Unemployment has thereby increasingly been tackled by resorting to supply-side schemes (Layard 1999). In particular, attention of economists and policy makers has focused on measures to: improve training and enhance human capital; improve information on market conditions to facilitate job search and job matching; incentivate self-employment; etc. All this constitutes the so-called active labour market policy (ALMP). In fact, ALMP approaches invalidate the Keynesian basic idea that market mechanisms may be held responsible for unemployment. On the contrary, it is believed that unemployment occurs because the labour market cannot deploy all its potential.

An implicit assumption is that demand conditions constitute an objective constraint to which supply can only adhere[9]. Accordingly, ALMPs such as retraining programs try to enhance the opportunity set of individuals by making individual choice more suited to demand requirements; on the other hand, subsidy reduction schemes find their rationale in the (social) need to force hesitant individuals to face reality. Welfare must be reshaped to help people gain self-sufficiency through work. Income support cannot be a goal, it must become an instrument. It should be linked to the condition that recipients are ready to accept job opportunities when these become available: hence the term 'workfare' (from 'welfare-to-work'). Individuals should only be helped to enhance their own *employability*, defined as 'the collection of workers characteristics, including attitudes towards work, expectations regarding employment and wages, and behaviours both in the labour market and on the job' (Peck and Theodore 2000b: 731). Such characteristics are seen as the factors that mostly affect the probability of finding a job, therefore it is a sound policy to try to improve them. Other forms of public aid, first of all the unemployment relief system, should be kept only for the hopeless losers.

Behind these ideas – and, more specifically, behind the notion of employability – lies the presumption that any job is better than no job. This is based on two aspects. On the one hand, there is the evidence that the involuntary unemployed are more likely to be affected by poverty, deprivation and abjection. On the other hand, there is a great social interest in the efficient performance of the system. Accordingly, it is important that individual preferences and choices should fit the structure of available jobs. The economic policy consequences are known. For those who end up in the unfortunate conditions associated with unemployment, the solution is to provide them with a wage, whatever the job they would like. Reduction of involuntary unemployment – regardless of the quality of the jobs involved – appears a sensible policy goal. Moreover, ALMPs such as retraining aim at making labour supply characteristics fit the demand for labour. As a last resort, individuals may be driven to comply with the contraints posed by the existing occupational structure – for instance, by reducing incentives to stay unemployed (e.g. subsidies).

Unemployment as a qualitative issue

It is questionable whether ALMP and its welfare-to-work variant are appropriate as a general solution to the unemployment problem[10]. An example may help us to set out the argument. Consider an economic downturn that causes job losses in many sectors, thereby affecting high-skilled as well as low-skilled and unskilled workers. Even if a worker can find a new job shortly after dismissal, this may or may not be comparable to the previous one. The situation may be trickier for skilled workers than for low-skilled or unskilled workers. The latter possess little or no human capital and they may move from one job to another with little consequences if any. Instead, skilled workers who accept low-skill jobs must be ready to lose their investment in education, training etc. Suppose then that a skilled worker wants to preserve her investment in human capital. On the one hand, she may feel obliged to accept lower wages and worse working conditions, and perhaps to have decreasing opportunities to return to her previous occupational standard as her human capital becomes obsolete and her social relations degrade. Indeed, the market mechanism requires her to accept whatever job becomes available. On the other hand, she may want to bet on her future by giving up the unsuited job and continuing her search. The rationale of her behaviour would be that expected differential returns from a high-skill job exceed the loss associated to the unemployment spell. The ensuing unemployment will be voluntary, inasmuch as it reflects optimal intertemporal choice; instead, it will be involuntary, as we noted above, if the decision is biased by imperfect information (the worker would have chosen the low-skill job, had she had the correct information).

Given this framework, conventional economics acknowledges a public concern to watch over involuntarily unemployed workers and help them to increase the probability to get a job. There is no public interest, instead, to take care of the voluntary unemployed, whose preferences and ensuing costs are essentially private ones.

We believe that, in spite of its commonsensical validity, the distinction between voluntary and involuntary unemployment (and of its policy implications) is not as clear-cut as it appears. Suppose that our skilled worker simply does not accept to comply with the emerging economic constraints; she may have developed a strong taste for good employment, for instance, and wants to avoid jobs that have unpleasant features – no matter how well they are paid[11]. In other words, we could even suggest that her reservation wage may be infinite, due to the high value attached to the desired characteristics.

Conventional economics is rather unsympathetic to this behaviour, which it labels as irrational: insofar as the worker's decision to stick to a 'good' job is inconsistent with social requirements, it is bound to fail. Thus, her decision must be based on biased speculation. Furthermore, unemployment occuring for these reasons is considered voluntary. It entails only private costs and it therefore does not deserve the attention of public policy. We think that such a position is simplistic. Voluntary unemployment of this kind is an increasing phenomenon, especially in Europe, where it is characterized by very long spells (especially for formerly skilled workers who cannot find a suitable place in the market), by the young age

of the unemployed and by their relatively high levels of skills and education. The conventional view – do not take care of the voluntarily unemployed – does not identify any specific measure to tackle these problems. We suggest that, to deal with them, we need to go beyond the explanation that standard economics provides. That the reservation wage is infinitely high suggests that there are aspects that cannot apparently be framed in usual monetary terms; or, more accurately, they can hardly be evaluated according to the standard metric. To this issue we now turn our attention.

The embedded labour market and the formation of labour supply

Evaluation, comparability and decision

In standard labour economics, only one metric (i.e. utility) exists that, common to all activities, allows to compare them. Although it is claimed that utility can only be subjectively evaluated, individual utilities are nonetheless homogenized at the system level through the 'willingness to pay' principle: how much something is worth to someone will be reflected in the amount of money he is ready to pay for it. This allows price variation to become the regulating principle of the market process.

Working will in general not provide utility *per se*, but only through consumption of the goods and services that the ensuing income affords. The worker will cease offering his services when the utility he gets from the consumption of goods bought with his marginal earning will equal the disutility of an additional unit of working time (i.e. the marginal utility of leisure). The supply of labour services therefore involves a piecemeal decision: as far as the wage rate is higher than its opportunity cost, the individual offers a further hour of his services. The model may be refined by plotting work earnings against the value assigned to any activity alternative to working (e.g. child care).

This sketch of the simplest textbook model of labour supply may be enhanced by taking into account that working may grant different degrees of (either positive or negative) satisfaction. The quality of working conditions may be expressed in monetary terms: you can always trade off a positive/negative characteristic for some amount of money. Thus, in equilibrium unpleasant working conditions must be worth – other things equal – a monetary bonus, otherwise those vacancies will not be filled. Fatigue, longer working time, dirt or pollution can be borne only if the wage makes up for them: put another way, the wage that is needed to accept an unpleasant job is, *ceteris paribus*, higher. Correspondingly, jobs that are 'nice' – whether for the task involved, the environment where they are carried out, the people engaged in them, the freedom of will they entail or whatever other feature but income – will be paid less for any given skill. Equilibrium in the labour market is ensured by symmetry on the demand side. In particular, a firm can acquire any desired type and quantity of labour provided it is willing to pay the market wage rate. If its demand remains unfulfilled, the wage rate should adjust upwardly to increase the relative attractiveness of the job offered.

This sort of accounting is feasible when the evaluation scale is the same for any economic activity that yields positive or negative utilities to individuals. For comparability to obtain – and, by the same token, for the possibility of market transactions – it is crucial that all aspects can be traded. Suppose that a job provides both a high income and a number of attractive features, while it shows no negative attribute, and the skills required by the job are rather easy to obtain. We expect that supply for that job will be high, exceeding demand. The market mechanism should guarantee that the wage falls to the equilibrium level; at the same time, wages for jobs that encounter supply shortages will rise to attract workers.

Things do not always work as expected, however. First of all, the competitive mechanism may fail to operate. There may be economic reasons, such as when a skill is scarce, whereby the holder may obtain quasi-rents from her indispensability. Truly, the gain cannot be seized indefinitely: someone may find it profitable to obtain that skill and outcompete the insider. However, some institutional obstacle artificially hinders the competitive mechanism. Access to the worthy occupation may be controlled and restricted institutionally by insiders. If this happens – as in professions, for instance – the wage-reduction effect of an increased labour supply is uncertain at best. Under these circumstances, the correct policy is, in conventional terms, to remove institutional obstacles.

A second issue that hampers the competitive mechanism is related to how valuation and comparability of different aspects is carried out. As we hinted in the previous section, a single aspect of a job may be valued so much as to outweigh any other characteristic. Jobs where the valued feature is present dominate the rest. In this case relative wage variations will not suffice to compensate for the loss of the dominant feature. A way to see the problem is to think of the rate of substitution (RS) between characteristics. The RS may be very high if, for instance, a worker wants to stick to certain kinds of jobs; it will nonetheless exist, insofar as the evaluation space is the same for all characteristics. Inasmuch as the criterion used to value the various aspect falls in the economic domain, we may expect that the price-regulating mechanism is hindered but not totally blocked. As an example, think of a skilled worker who wants to preserve job-specific human capital: we may suppose that there will be a threshold above which the money amount compensates for its loss[12]. Unemployment may occur, as we suggested, but only if the relative wage adjustment is not strong enough to induce people to switch from one job to another.

Things are different if the valuation criterion falls outside of the economic domain. In this case, a rate of substitution does not exist. This may happen if a person has a strong preference for some features that only some kinds of job have, or she may value some issues so much, that she may refuse jobs that do not comply with them. A couple of examples may be given:

- Think of a mother who may want to take care of her children; for her, subsidized unemployment might be better than full-time work (see Roemer 1998 on this).

- Suppose someone is only offered a job by a firm that produces mines, the end-product of her effort possibly being that many children will be wounded or killed around the world; if this undermines her moral value whereby other people's lives and dignity must never be threatened, she will prefer to stay unemployed (a specular example may be given by someone who only wants to work in the environmental protection field, even as a volunteer rather than being employed elsewhere).

Both of these cases introduce social aspects, which may conflict with standard economic evaluation. This is the outcome of an economic process that is embedded in society. In the rest of this section we will stress that the price mechanism may fail to regulate an embedded labour market, thereby leading to a number of crucial unemployment-related issues. The whole subject is related to how individuals make their occupational choice. Our suggestion is that the choice process is socially influenced in two ways.

First, the choice process is a cultural process based on social interaction among socialized individuals. Individual behaviour is embedded in the social context. This means above all that (a) the environment we live in defines the criteria to assess the appropriateness of our economic and social action and (b) we tend to conform to these social requirements (Blau 1964/1986)[13].

Second, a social criterion lies at the center of the process of preference formation. Status is such a principle. We define people's status as their position in a social structure regulated by power and prestige; in some way it is a measure of their social standing or social honour in a community.

The crucial point is that in contemporary capitalist economies the main source of status is occupations[14]. People tend to define their own position in society by moving horizontally or vertically through the occupational structure. The legitimization of mobility at the societal level makes it an institutionalized practice (Blau 1964/1986): social entrepreneurship is a most-valued route to self-assertion. This is just one element of a double movement, though: on the other hand we have a drift by the ruling class to react to the incentives to mobility that capitalism has introduced.

The social nature of the process of choice entails the complementarity of the two issues – restriction of competition and social valuation of occupations. On the one hand, if a certain job is much-valued, there will be many aspirants to the corresponding positions; insiders, however, may find it profitable to shelter that segment from competition. As a result, unemployment may occur. The restrictions reinforce the appeal of that occupation, though, thereby increasing its social value. On the other hand, conformity may also act as a social regulator of the labour market: a family background is often a source of opportunities and constraints, usually inducing choice paths that reinforce original status. Restrictions based on cultural barriers will act on this, ensuring that the number of people competing for a high-status job will never be extremely high. The final outcome will depend on the relative strength of positive and negative feedbacks, as we shall show.

The construction of preferences

The inquiry into how individuals choose what job to look for requires an analysis of the process of preference formation, and of the process whereby preferences are enacted (Levine 1997, 1998). As Amartya Sen has pointed out, any action is based on some prior evaluation process. This concerns both the designation of objects of value and their valuation. Utilitarianism makes it very simple: valuable objects are actions that deliver utility and they are valued according to the utility they yield to the acting individual. Standard economics accordingly relegates the issue of preference formation outside its field of study. This stance is instrumental to the 'common metric' view discussed above, which is based on the postulate that individuals follow self-centered utilitarian patterns of action. This assumption is rather restrictive, and it may be misleading (Sen and Williams 1982).

Individuals may be driven by other motivations and feelings, as Sen (1991) points out. Self-respect, participation to community, emotional feelings, search for honour, altruism etc. are remarkable examples. More specific to labour supply, it should then be considered that a person's work is an essential step for self-expression and the costruction of personality. The job one chooses may also reflect passions and inspirations, specific abilities and propensities, or other driving forces. Moreover, a person's job is the major source of his social identity. As noted above, people are socially valued for what they do, their living standards and their economic autonomy (i.e. dependence on needs), the nature and the structure of their rights and their duties (i.e. power and command). All this leads to suggest that individuals assign a high value to occupations that comply with the construction of their self-image and with the image others have of them. Put another way, occupational status provides the economic and social status people are valued for[15]. It is therefore important to stress that occupational choice does not only involve a search for specific job characteristics; it is above all a search for status.

It can easily be seen that, if individual choice is free, higher-status employment will be preferred. Status demand will direct labour supply towards a restricted number of occupations. Excess supply will be likely to occur in higher-status jobs. Conventional economics assumes a disembedded economy where any particular characteristics may be traded off for money. In such an economy, income and power (or, say, fame, prestige, pleasant working conditions) should be negatively correlated in the long run. In the space of occupations, one chooses whether to have higher income and little power or just the reverse, other things being equal. It is instead worth noticing that status-related rewards (e.g. income, fame, deference, power etc.) are not mutually exclusive but are somewhat correlated in bundles (Tumin 1967; Reiss *et al*. 1961). This complementarity is essential to status: should income decrease, the A-job status would also change. As a consequence, status concerns may overwhelm any other benefit by inducing a hierarchy in preferences, that is by actually framing choice so as to inhibit directions of search. The labour supply tends to converge on a few employment positions, those where the higher-ranked jobs are concentrated. The problem is whether disequilibria will be regulated by price variations or not.

The self-regulating power of the labour market may be limited in a number of ways, according to the issues mentioned above. The first case is one where wages and salaries do not adjust downwards. This is possible if the number of available high-status jobs is kept scarce. Suppose job A grants high status while job B does not[16]. A-type jobs are chosen because high status entails high income. The former are going to be scarce compared to B-type ones. Everybody is going to line up for them, but obviously most people will not get the job. Formal and informal mechanisms work to shelter A-jobs from strong competition, protecting the existing quasi-rents. The case of professions is typical in many countries. The positive correlation between favourable characteristics is maintained and renewed through the institutional rules that govern access to the profession.

The second possible case is when revenues decrease, but this does not lower supply. If assessment of occupations is defined according to a specific feature, it is plausible that some choices dominate the rest in preference orderings. A typical description is that of lexicographic preferences, where one characteristic is absolutely preferred to any other (Georgescu-Roegen 1954)[17]. We may think of those occupations where economic concerns lie outside the realm of motivations (e.g. artists); but we may include any employment where passion may be a strong motivator for individual choice.

Finally a third case exists that probably connects the previous ones. Positive social values – such as respectability – are ordering criteria for action. They find their rationale in the conformity to social attitudes or principles that transcend monetary revenue. Examples are often provided by people who are politically involved, old-fashioned soldiers or priests. If we consider status as one of such social principles, we can include in our list all those commercial activities that are highly valued by the community. Thus, social scarcity on the demand side and non-economic driving forces on the supply side may interact. The outcome may be an imbalance which, for any given occupational structure (which depends on demand-side conditions), cannot be removed by the standard self-regulating mechanism.

The imbalance we are talking about is, of course, unemployment. Excess supply for some high-status job will entail supply shortages in other lines of business. Economists usually refer to this type of unemployment as ‘voluntary’: if people choose not to work in the sector with vacancies, it only means that they are better off without a job. This is a private matter. The framework we are developing tells another story. ‘Free’ choice is in fact bound to social tastes; a possible public concern is thereby surfacing.

Relative prices cannot regulate the market. Should A-job revenues decrease, their status would also change. This is not the normal case. Even though this sometimes occurs (think of the changing status of teachers over the past century), it takes decades of cultural change. When price variation does not accomplish its task, equilibrium is only reached if redundant applicants are pushed onto B-jobs. Eventually, equilibrium can only derive from a forced revision of initial choices, and it is also likely to produce frustration, cognitive dissonance and social exclusion. These are definitely public concerns.

No straightforward conclusions can be drawn from the above though. In principle, capitalism fosters social and economic mobility. This is both a condition for and a result of the correct working of the competitive mechanism. Yet the social structure, which the economy is embedded in, provides different criteria for action. The social structure changes slowly, because it requires cultural change and it encompasses reactions from vested interests. The next section deals with two issues: on the one hand, which social dynamics characterizes and leads to the demand for status; on the other hand, the intricacies of negative and positive factors which feed back and forward on the dynamics of labour supply.

Occupational choice

The basic model of a stratified society

The proclivity to follow social standards calls for a more precise characterization of how social relations may influence labour market outcomes. This is a multifaceted question. As an initial clue, consider that many empirical works do find a significant correlation between starting conditions (e.g. parents' education, father's occupational status) and outcome on the labour market (Picketty 2000). This contrasts with the expected efficient outcome of a free-market society, whereby allocation of people to jobs is based on pure individual preferences, abilities and merits. Two types of explanations may account for this result. The first case is the relatively standard suggestion that life paths reflect family economic constraints, thereby affecting the patterns of accumulation of human capital through education. In addition, there may be institutional features (e.g. imperfect capital markets) reinforcing existing constraints on resources. The second explanation, rooted in the bounded rationality tradition, stresses influences on preference formation. The hypothesis is that the family background poses cognitive and social constraints that make the preference set incomplete and biased.

Despite their apparent exclusivity, the two explanations may be the two sides of the same coin, since economic constraints and biased preference structures are likely to have a common origin in one's personal history (e.g. higher-income families are more ready to invest in education because they have higher expectations on children; on the other hand, they maintain privileged expectations because they can afford to invest much in education). How cultural and material resources are strictly coupled was magisterially shown by the French sociologist Pierre Bourdieu (1979). According to his theory of social action, individuals follow a set of internalized dispositions that derive from the sedimented experience of previous generations, which has been socialized. He calls them *habitus*. It defines a lifestyle, determining the choice of disparate practices, from the most trivial to the more complex and meaningful ones. *Habitus* implies that an individual internalizes objective chances associated to the conditions of her family and its closest environment. Subsequently, she transforms that set of opportunities into a path of subjective aspirations and expectations.

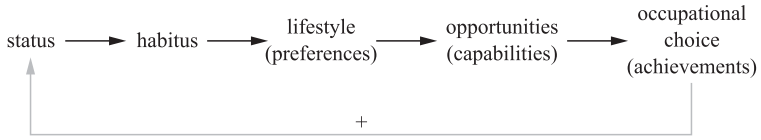

Figure 1 Stratification implies positive feedback between status and achievements

An illustration of Bourdieu's basic idea, adapted to reflect its application to our argument, is provided by fig.1.

Families are the first social setting a person finds. Family status, synthesized by its *habitus*, provides the criteria for action. Comparing one's own situation and ambitions with peers' achievements and aspirations provides further stimuli. Social differentiation is reflected in the related steps of judgement, tastes and their practical consequences: choice of pattern of consumption, occupation and so on. Individual labour market outcomes are therefore predetermined by the *habitus*, which channels the job search path and provides the cultural coordinates for maintaining one's position in the social ranking. Bourdieu's point is that these outcomes follow naturally from the structure of objective classes (although he admits individual deviations). In a stratified society, everyone 'naturally' finds his own place in society. Preferences represent the 'practical assertion of a necessary difference' (Bourdieu 1979: 53 of the italian edition). Truly free choice, in the scheme, is excluded. Moreover, the process is persistent in nature, as is indicated by the positive feedback from outcomes to status. Stratification is consolidated by practices that continuously mark the distance among classes[18].

In the graph, we suggest a correspondence between Bourdieu's concepts and those used in economics (some of which have been introduced by Nobel laureate Amartya Sen). For instance, lifestyle (the unity of practices in the different fields of action) closely reminds us of a preference ordering – still a coherent one, though incomplete and biased, socially structured by conditions that are specific to each class. Sen's capabilities, that is the possibilities among which a person can choose in order to conduct his life, are reflected in Bourdieu's opportunities and expectations, while achievements (i.e. the implementation of capabilities, that is the 'things' actually chosen) can be easily translated into outcomes (Bourdieu's choice). Both capabilities and achievements are defined in the space of functionings, that is the whole range of possibilities concerning what a person can be and what he can do (Sen 1991). According to Sen, inequality should be measured by differences in what people can really do and get. The capability method, he believes, is superior to both measurement of goods possessed, and measurement of the distance from a desired goal: it is more objective than the latter and it allows for more qualitative considerations than the former. Bourdieu seems to follow a similar path, perhaps more radically: the capabilities of distinct social groups

differ sharply, so that the notion of *habitus* suggests that the space of functionings is partitioned.

What we have described here is a rather rigid social system, self-regulated by a positive feedback (achievements-status) that maintains existing positions. Apparently, if the structure of values associated to jobs is consistent with the occupational structure of the economy (the structure of demand), everything is fine (Blau 1994). Valuation of occupations follows from the *habitus;* consequently, expectations of each group are rigid and stringent. There seems to be no sign of possible overcrowding in any of them, because everyone knows which is 'his place'.

The process depicted is somewhat mechanical. We may want to allow changes to occur somewhere in the causal chain above. Social mobility, i.e. the upward or downward movements within a stratification system, is the most important disturbing factor. For Bourdieu, social mobility is a relatively secondary analytical problem; the reason is that he believes that formal and above all informal constraints on access to high-status occupations are at work (and that they work well). This cannot be taken for granted.

Upsetting existing strata: individual social mobility

Social mobility is a key feature of modern society. In old, ceremonial societies status is essentially ascribed (i.e. inherited) status. At the opposite end, pure market economies should eliminate ascribed status. Individual merit is the criterion that provides the major source of capabilities. More specifically, different jobs yield differential income according to the skills and ability they require. Income allows consumption of goods and services; utility from leisure (and/or work) and consumption, as we said, is the only valuable criterion for action. Inasmuch as the system of markets moulds modern industrial societies, we expect that innovative and meritorious people continuously upset the social (economic) ranking. Social mobility is the norm.

Social mobility is introduced in the framework through expectations. As mentioned above, their role is central in driving action (Blau 1964/1986: 145ff.). Mobility reflects a distinction between settled expectations (deriving from the *habitus*) and emerging aspirations (deriving from reference to the external environment and communication). Another way to see the distinction is to consider it as the difference between actual status (status inherited from the family social position) and desired status (the aspirations and expectations one develops over time). At the outset, people find themselves endowed with the cultural and economic resources provided by their families. Suppose that their basic expectation is to replicate family tradition: this is consistent with the conformity principle recalled above (implicit in Bourdieu). Yet social conformity may involve different aspects (Tumin 1967). At a basic level individuals try to adhere to a basic standard of legitimacy and morality. At a higher level, conformity means doing what the community does, particularly what those members in their associational networks do (Bandura 1986). Still upward, conformity may incorporate change, a move upward in the

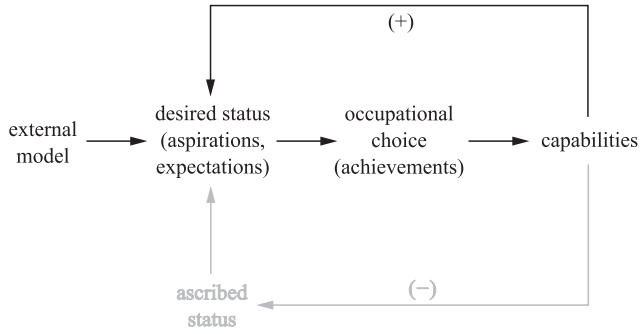

Figure 2 Widespread social mobility means that ascribed status is a weak motivator of action

social ranking. Mobility requires that the existing culture values social-innovative behaviour positively. If climbing the social ladder is a valued behaviour, as it is in capitalist economies, this means that, rather than following traditional trajectories as suggested by Bourdieu, people gain from being social enterpreneurs. The external environment provides stimuli to abandon established paths. At this stage, we suppose that the external model (i.e. the model coming from the environment) may be stronger than the ones arising within the family. This is shown in fig. 2, where the lower influence of ascribed status is revealed by the lighter colour.

Differently from fig. 1, here preferences do not follow automatically from status. They stem from the attempt to compare and make compatible the different aspects of conformity. Upward mobility implies a shift in the individual's goals. Her objective is to perform better than her peers. Social psychology sees the attempt to improve one's status as an innovating behaviour. People with outstanding *forethought capabilities* (and who are also more prone to taking risks) may develop progressive images of a desirable future for themselves (Bandura 1986). They may adopt role models (e.g. people of higher status) to learn from, or refer to symbolic and abstract models (e.g. from television, books or magazines). Since status incentives are extremely powerful (Bandura 1986: 238–9), successful steps in the course of action will also reinforce active learning. In this case achievement in the labour market feeds back positively on aspirations; more importantly, it feeds back negatively on status (that is, mobility weakens the cognitive constraints posed by ascribed status).

By the same token, increased occupational status entails greater opportunities in the future. The search for higher status involves, in fact, a search for increased life opportunities (related to aspects such as health, fertility and mortality) and lifestyle (concerning, for instance, consumption patterns, political and religious attitudes, and personality) (Tumin 1967: Chs 7 and 8). Put in Sen's words, higher status entails improved capabilities. As these are put into practice, an even higher status is likely to be attained in the future. Status may be seen as 'expendable

capital' (Blau 1964/1986: 132ff.) both in interpersonal relations and for personal development. If one uses this capital 'productively' (e.g. by cultivating social relationships or by enhancing his own education) we may anticipate a virtuous circle linking his acquired capabilities to status (see fig. 2)[19].

This cumulative characteristic of status and of capabilities provides a strong motivation for moving away from the initial position one finds himself in. Apart from social comparison issues, in fact, there are substantial reasons related to the level of wellbeing a person will attain. What is more, however, is that increased capabilities means a greater set of open opportunities, that is a greater positive freedom to ensure functionings (Sen 1992). This would necessitate a discussion of what should be valued, from a private as well as a public standpoint. While this task is out of the scope of the paper, a few considerations will be advanced at the end of this section and in the conclusions. It will be shown that (a) the attempt to increase capabilities may be a reasonable strategy for an individual and that (b) social mobility is, under this respect, a socially good thing as it re-shuffles existing strata.

Unfortunately social mobility, while a feature of modern societies, is often constrained and confined to some specific social tiers. We should therefore reflect on how to combine the two extreme cases analysed so far: the stratified and rigid society depicted by Bourdieu and the mobile, ever-changing one.

A possible synthesis

If the framework suggested in fig. 1 has some relevance, we should expect that some of the described mechanisms still hinder upward mobility. A first order of impediments comes from the interrelation among individual behaviours. If individual attempts to move up the social scale are positively rewarded (in terms of achieved status), innovative behaviour may be imitated, leading – mainly through observational learning of behavioural patterns and social emulation in associational networks (Bandura 1986: Chs 2 and 4) – to a process of cultural diffusion of behaviours. But, owing to social scarcity, diffused innovation cannot be rewarded indefinitely (Hirsch 1976; Blau 1994; Schelling 1978)[20]. The point is strictly related to the twofold nature of status. Status not only provides an absolute reward, in that it allows to enhance capabilities; it also provides a relative reward, in that it marks relative position in the social ranking. Diffusion of innovative behaviour is driven by the attempt to capture the advantages of differential status (compared to the position that is still occupied by peers); as the new position is achieved by followers, the amount of differential reward decreases. When the process involves the whole relevant population no differential prize is secured. Eventually, while such a process leaves the relative positions of individuals in the social ranking unchanged, it may still produce a positive effect on the absolute levels of status and capabilities.

It may then be asked how the whole status structure is affected by this collective movement – given that for the highest position no upward mobility is possible.

Bourdieu's defence of his framework is interesting, because it suggests a sort of negative feedback to social mobility. A remarkable feature of cultural capital[61] is that the *habitus* of the upper classes allows members to react promptly to – even anticipate – changes in social behaviour. Two routes may be taken. The first one is that the relevant practices of the upper class change as the lower classes try to emulate them. As a matter of fact, the process of cultural diffusion of behaviours cited above involves new evaluation standards and consequent new standards of behaviour (Bandura 1986: 48), provided the process lasts long enough and is appropriately rewarded[21]. A corresponding change in evaluation standards and practice at the upmost level is therefore likely to occur. The second route is that upward mobility does not involve some aspects on which class distinctiveness is based. In other words, access to some practices is restricted. This is clear for professions, for instance. In this case, *direct* or *pure* social scarcity – whereby satisfaction derives from scarcity itself – is (re)created, rather than being incidental as in the former case (Hirsch 1976: 22).

As we noted above, these adjustments occur mostly in informal ways. On the one hand, since tastes are the 'practical assertion of a necessary difference', distinction need not be sought. By 'being themselves', the elites mark the distinction (this is the very strength of the link between *habitus* and lifestyles in Bourdieu's framework). On the other hand, there often are secondary properties that help to define a position (e.g. a profession) in selective terms: for instance, race, sex, age, place of origin or social background may be used to tacitly discriminate against people, excluding them or allowing them to have access only to marginal positions (e.g. female doctors and black lawyers may essentially have female patients and black clients).

As a consequence, we obtain that the process descibed in fig. 2 is combined with the forces depicted in fig. 1. A possible synthesis is in fig. 3.

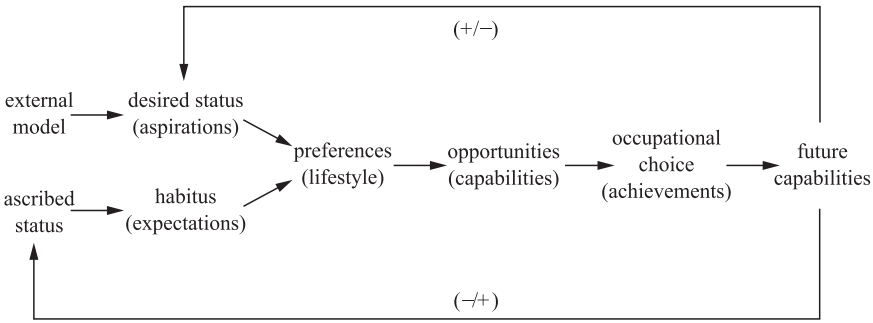

Figure 3 A synthetic representation of how stratification and social mobilization interact

As a general rule, aspirations and expectations must be verified in subsequent social and economic interaction. The path towards the desired status is mediated/constrained by actual status, through the cognitive, social and economic resources that the latter provides. Upward mobility is not blocked, but it may be hampered. Social scarcity may be artificially created, whether through formal norms or informal, tacit practices, but the forces behind mobilization keep on working. The main reason, we suggest, is that status choice contains a substantial element of definition or satisfaction of personal identity – which goes well beyond the ceremonial aim and invidious attitude it may nonetheless include. Occupational choice entails above all the choice of a trajectory for constructing the self.

Propelled by a positive attitude towards upward mobility, the supply of labour may target jobs for which labour demand is (kept) scarce. Independently of high potential requests, low labour demand implies that actual success is limited – relative to candidates, at least. Pure social scarcity mechanisms exacerbate the inconsistency between different levels of rationality: individually, it may be rational to pursue a goal; collectively, it is obvious that a great deal of expectations are disappointed.

If a trial proves unsuccessful a revision process must take place. A disappointed applicant may react in two ways:

1) By remaining unemployed, while still striving to fill one of the few vacancies (perhaps modifying something in the sequence of necessary steps). This requires that a good amount of economic resources are available, as a partial substitute for lacking cultural capital[22].
2) By pushing aspiration levels downward, to match expectations deriving from actual status.

It is not possible to know a priori which revision path will be chosen. The rationality paradox inherent in the social scarcity argument implies that the goal is *difficult but not impossible* to achieve: after all, some succeed. The revision will be enacted by a divergence between the desired state and the achieved outcomes, for which external or contingent factors may be charged; the failure may well be deemed transitory. The way the individual frames the situation is likely to affect his subsequent decisions. He may infer that some steps were wrong and should be revised; or he may conclude that his hopes have been definitely frustrated. The degree to which a drastic revision will occur is likely to be context-dependent and institutionalized. Factors that influence the choice of conduct include personality and motivation, collective perception and evaluation of negative differential status tagged to the achieved job, and awareness of other available opportunities. A crucial role will be played by the relational networks and the resources possessed[23].

Whether it is unemployment or downgraded employment that occurs, however, the outcome entails relevant costs, which stem from a social process but accrue to the individuals involved. The problem is whether these costs should be deemed social costs or, at least, if there is a social interest in reducing them.

A partial conclusion

The above discussion has a range of implications in terms of the ensuing social outcome. Let us begin by discussing the details of the costs involved if the labour supply is forced into lower-status occupations. Suppose a university graduate seeks a job suited to his qualification. When he fails, he unwillingly starts looking for another job (possibly any job, if he is a constrained breadwinner). Although standard theory would include this circumstance under the heading of a 'wrong investment in human capital' (which entails that 'bad investors' have to accept this response and, consequently, change their goal), there are doubts whether the costs involved ought to be ascribed to the individual only.

There are two distinct issues at stake here. The first concerns who actually bears the costs of unemployment that stems from the mismatch between supply aspirations and demand requirements. It may be that social costs are involved. The second regards who ought to be held responsible for those costs. Insofar as there are significant external influences on individually borne costs, the question of who should bear those costs is also relevant.

Let us focus on the revision that follows an unfavourable labour market outcome. If the original search for higher status required investment in human capital, at least part of it will be jeopardized. This cost is mainly private, as underlined by the 'wrong investment' argument recalled above. Yet there is a social counterpart. When a worker is employed in a position that, given his qualification, yields a lower productivity than potentially attainable, this means that social product is lower than optimum; this is also true if the person works less hours than standard working time. This output loss founded the concept of *disguised unemployment*, introduced by Joan Robinson (1937) and recently renewed by Eatwell (1997) and Covick (1998).

Individuals may also face significant psychological costs. Frustration is a possible outcome of the forced revision. On the other hand, an individual may respond to an unfavourable outcome by enacting a cognitive dissonance reduction process, which may entail different routes. A basic one consists in getting convinced that the initial condition was, after all, the only possible one (a typical example that Sen refers to is people who accept to feel happy with what they have). Alternatively, the process may lead the person to adopt the construed belief that what she has reached is *almost* as good as what she aspired to; either way, one may think that what she aspired to was not actually as good as imagined.

Such psychological costs also entail social costs. Apart from consequences accruing to the national health systems, we may suggest that – using a 'human resources' line of reasoning – a frustrated labour supply may also work less enthusiastically, i.e. at lower productivity. This cost is additional to the one observed by Joan Robinson and her followers.

The next issue is whether the above costs are paid by the right subjects, i.e. those responsible for them. Two extreme views are possible. The first one holds that only individuals should be blamed. Costs accruing to individuals are due to their inability to take into the account the more or less obvious demand constraints.

This case is less problematic than that of social costs measured by the lower productivity of the labour force. If the loss of potential output can be traced back to the same inability, then individuals are actually shifting costs to the collective. In conventional terms, had they appropriately assessed the nature of demand constraints, then their optimal choice would have been a shorter and perhaps sharper search.

A second, opposite view is indeed possible; we found it in the different perspective we have adopted on individual choice. We think that the psychological costs described above – while ususally considered an individual concern – are in fact social insofar as they derive from a social process. As we suggested, the job valuation process is not a matter of mere individual will. Jobs are valued for the status they confer, which is socially defined: individual choice heavily reflects a system of collective values. Embeddedness implies that individual choice cannot be reduced to individual valuation. Even if individuals do not assess the constraints correctly, we should keep in mind that scarcity is not objectively given, it is institutionally created. If scarcity is artificial, the central issue turns out to be the distribution of status, which on the other hand depends on the general, public attitude towards mobility as well as on institutional constraints. The problem typically arises when the hightened standard of living induces people to seek jobs that are increasingly higher-ranked than those which were hitherto the accepted standard of worthiness. As we said above, the possibility for individual involvement relates to the individual's possibility to enhance his capabilities, which in turn depends on the extent to which he can improve his status. In this framework, the labour market is the arena where mobility finds richest humus.

The key issue, here, is that there is a collective interest in encouraging social mobility. As we anticipated, social mobility is probably the primary mechanism that produces shifting valuation scales, thereby enhancing living standards. Inasmuch as it is accompanied by a corresponding shift in labour demand, the search for a better status induces a collective move towards better jobs that, for the reasons outlined above, yields an improvement in capabilities. In the long run, jobs characterized by socially unpleasant tasks are displaced and substituted by better jobs. The structure of available opportunities is affected by this mobilization. An evolutionary process of this type has, for instance, marked the passage from agriculture to manufacturing and from manufacturing to the service economy[24]. What counts is that such evolution brings with it increases in overall productivity. Moreover, the process by which capabilities are collectively enhanced may be socially valued, as Sen forcefully and repeatedly argued, for its consequences on the process of economic democratization (see, for instance, Sen 1992; 1999).

If this argument is deemed relevant, then the fact that those objectives (in a nutshell, the evolution toward a more democratic economy) are missed because of counteracting social practices may be valued as a cost for the collective. Accordingly, the fact that individuals do not accept to pay this cost privately and try to stick to their aspiration without internalizing the extant constraints also has a positive flavour. Voluntary unemployment, far from being an inefficient response to an efficient mechanism, is a voice mechanism (Hirschman 1970) against the social inefficiency of the socio-economic structure.

A possible criticism to this argument lies in the difficulty to disentangle individual and social concerns. Truly, although there certainly is a social interest in relieving men of fatigue and alienation, in rendering human life less dependent on unsatisfactory occupations, in elevating human dignity and integrity etc., it may be that the search for better jobs, as described above, implies a ceremonial attitude. Indeed, the status tagged to jobs partly reflects the mechanism of invidious comparison that institutional analysis has stigmatized since Veblen. If we limit ourselves to this level of analysis, we almost inevitably fall back to the *individual cost* type of argument: no social safeguard ought to be warranted to decisions that only serve individual interest, even if they are inspired by (or are entangled in) the social processes of valuation.

The situation is far more complex, yet, thereby requiring further considerations. First, suppose a person undertakes an action that aims at goal A (competition for status), which is prejudicial to some social value, and indirectly pursues goal B (mobilization of society), which is valued positively by the community. The claim that the action is certainly to blame is unfounded[25]. Correlated to that, we note that the ceremonial attitude may be paralleled by an instrumental side-effect: a person who wants to graduate also improves his education and knowledge, thereby sustaining a potential growth process, independently of his possibly ceremonial motives. Second, the pursuit of a better status remains a socially positive issue even if structural – rather than artificial – social scarcity makes the pursuit of better jobs difficult for some and impossible for the rest. Any contrary claim inevitably supports the *status quo*, with its structure of discrimination and inequality. This issue pertains to economic democracy, and real labour markets are apparently unable to deal with it.

Finally, some words are worth spending to introduce the discussion we present in the rest of the paper. Suppose that the tension between the unchanging job structure and the pursuit of a better job have a structural basis. Individuals set their own standards according to the evolution of the social valuation process while labour demand is bounded by technological change and profitability considerations. There is no need that the two should be mutually consistent. Productive knowledge may be such that it is impossible to improve working conditions and the job structure. The job structure that arises from economic and technological evolution may or may not reveal an increasing proportion of *good* (i.e. high-status) jobs as opposed to *bad* (low-status) ones. What is *good* or *bad* is a matter of social valuation, independent of technological requirements. Hence *good jobs* that are not provided in sufficient amount by the productive sector will follow the principle of social scarcity, which individual search runs into. This may produce unemployment, understood as the empty intersection between the *desired* jobs set and the *achievable* jobs set[26]. Under these conditions, if individual status depends on job status, it may be improved either by changing one's job (given the structure of 'status tags') or by changing the 'status tag'. The latter case entails a process of cultural innovation, which may, under some conditions, be socially valuable[27].

We have to deal with two important issues – individual development and freedom, on the one hand, and economic order, on the other – which represent two sides of

the same coin. Understanding their rationales may provide some insights on the policy implications of unemployment. If technology and production are objectively constrained, the crucial issue of the above exposition will presumably be ceremonialism. But if technology and labour demand reflect institutional practices rather than natural conditions, the ceremonial features of labour supply may turn out to be of secondary importance, and the status/capabilities argument fully applies.

The fact that the conditions of labour demand may hinder individual attempts to enhance capabilities calls for an analysis of the objective nature of the knowledge that underlies production-related choice and action. This issue will be discussed in the section that follows.

Labour demand

Technology and efficiency

In conventional economics the demand for labour depends on the marginal product of labour, which, in turn, depends on exogenous technology. Firms aim to maximize real profits. To achieve this goal they hire all workers who are willing to accept the equilibrium wage. As a result, full employment is achieved. To be true, a broad set of conditions must also be satisfied. Even when firms are profit maximizers unemployment may occur because the allocating mechanism based on relative prices does not operate as an efficient coordinating instance. Circumstances that lead to this outcome are known as market failures[28]. Whatever the specific account provided for unemployment – an oligopolistic labour market where unions force firms to pay non-equilibrium wages; incomplete information that leads workers to seek jobs or wages that are not compatible with equilibrium; a monopolistic goods market, which entails an incomplete allocation of resources – it is taken for granted that, so long as firms maximize profit, they are behaving in an efficient way.

Although a range of assumptions underlying this approach have been subject to criticism from different strands of economic thought – they include the notions of maximization and substantive rationality (Simon 1976), equilibrium (Hamilton 1991), ergodicity (Davidson 1994) – the discussion that follows focuses on just two issues: the (efficient) behaviour of firms and the (exogenous) nature of technology. The two issues are strictly interrelated and their implications for a proper understanding of unemployment are, in our view, rather important.

We contend that:

1) Technology generally cannot be conceived of as independent of sectional interests; quite to the contrary, its insurgence is strongly related to business requirements.
2) *A priori*, business goals may or may not be consistent with social goals; the problem is to assess whether, and to what extent, the inconsistency exists.

3) The relation between technology and sectional interests precludes the use of technology as a benchmark to assess the social efficiency of business; an alternative benchmark is therefore called for.
4) Sen's notion of capability is a more appropriate indicator of social efficiency, even though it is subject to similar shortcomings as those of technology.

Our discussion begins with a short overview of four distinct theoretical approaches to the above issues. We begin by referring to Veblen's view of technology and how he pointed out the social inefficiency of business. We then introduce Leijonhufvud's approach to innovation in the factory system. We extend Leijonhufvud's Smithian approach by introducing Dosi and Marengo's evolutionary account of how business innovates in the goods market. Finally we examine Marglin's radical economic claim that business is socially inefficient and that technology reflects this inefficiency.

A major criticism concerning the behaviour of firms comes from Veblen (1921). He distinguishes technological efficiency from business behaviour. Engineers are required to organize production in order to maximize[29] output under given technological and resource constraints. Insofar as they succeed, they achieve technical (and social) efficiency[30]. Businessmen, on the other hand, aim to make money[31]. Insofar as they succeed, they are efficient on economic grounds. The two notions of efficiency do not coincide: output maximization is not the same thing as profit maximization[32]. This divergence may be measured in terms of what we might call an *efficiency gap*: the difference between potential output (social efficiency) and actual output (economic efficiency).

The difference between the conventional and the Veblenian views of efficiency depends on the assumption concerning business behaviour. The conventional view is that profit maximization and output maximization converge whereas this need not be the case according to Veblen. However, the two approaches share the important assumption that technology is independent of business behaviour. It can, therefore, be used as a benchmark to assess the efficiency of the latter[33]. From this perspective, unemployment is the outcome of a technically inefficient use of resources, which may depend on market failures (conventional approach) or on the dominant role that the vested interests of business have in modern society (Veblenian approach).

The use of technology as a benchmark requires that another assumption must hold. The technological process must be defined as 'that behavior which promotes the life process'. This was Veblen's view (Waller 1982: 765) but it is not a convincing one (Swaney 1989) in that it implicitly assumes that the tools associated to a given technology have an instrumental, as opposed to ceremonial, nature. This need not be the case, as even authors close to Veblen's approach have pointed out[34]. In the rest of the paper we therefore assume that technology includes 'tools', (individual) skills, the division of labour and related (firm and industrial) competences, independently of how they arise and of the use they are put to[35]. This allows us to freely discuss both the exogenous and the instrumental nature of technology, and the implications this has for employment.

Following a Smithian approach, Leijonhufvud (1986) provides an interesting account of technological change. He focuses on the relation between the division of labour within firms and the introduction of machinery. The latter occurs every time a single task in the production process is so simple that a machine can substitute the worker who is carrying it out. The reason why tasks turn out to be so simple is that it is convenient to split them up in order to cut production costs. As a result, the introduction of machinery is a consequence – rather than a determinant – of changes in the division of labour that occur on the shop floor. Technology is not entirely exogenous[36].

The division of labour is based on conventional economic efficiency considerations: the aim is to reduce production costs as much as possible. Leijonhufvud's point of departure is that a team may substitute individual craftsmen because in so doing it can achieve a greater output. Cost-cutting is independent of distributional issues. These arise only subsequently, when that output has to be shared. This implies that the technology used is also independent of distributional issues. It depends on considerations of allocative efficiency alone.

Leijonhufvud is concerned with the factory system, thus with what goes on within a firm or an industry. Dosi and Marengo (1994) focus on the firm's final market, thus on the need to meet some customer-related requirement. They follow a competence-based approach whereby the activity of a firm – what it does and how it does it – is to some extent, unique, owing to its idiosyncratic (tacit) knowledge[37]. The authors point to two possible strategies that firms can follow. The first one focuses on competitiveness. It consists in achieving 'a quality as high as possible and costs as low as possible'. The alternative is *strategising*, i.e. changing a firm's market power independently of its competitiveness. The authors argue that strategising is convenient when 'every competitor notionally knows how to make a new car (...) equally well' (Dosi and Marengo 1994: 163) so that there is no scope for improving competitiveness. Dosi and Marengo acknowledge that profit maximization through strategising is possible but they deny that it is a general case. The idiosyncratic nature of knowledge, they argue, implies that competitors hardly conceive of a good in exactly the same way. Thus, as a general rule, it is more convenient for firms to follow a 'competitive' strategy.

Despite the differences between Leijonhufvud's and Dosi and Marengo's approaches, they share the idea that technology is the partial outcome of a firm's activity. Under these circumstances, technology may be used as a benchmark to assess the behaviour of a single firm with respect to what other firms are doing. Leijonhufvud's approach suggests that there is a unique pattern of technological development, whereby some firms may be expected to be at a more advanced stage than others. Dosi and Marengo provide a more complex framework but it may still be possible to judge whether a firm is more innovative than others or – ex post – whether it chose the best technology.

It is less straightforward that technology can be used as a benchmark to assess whether firms (and business as a whole) are socially efficient. The issue is fairly easy to cope with when there is only one way for technology to evolve, as Leijonhufvud suggests. Under these circumstances one might conceive of social

inefficiency in terms that resemble Veblen's approach: a *technology gap* would measure the output difference between the technology actually employed and the technological frontier. The problem is that if business is economically inefficient, it is most likely that it will not seek ways to cut costs – that is to say, it will not innovate – as it might. This implies that the technological frontier will not reach its potential. In other terms, the degree of economic efficiency affects the position of the technological frontier. Consequently, technology cannot be used to assess the efficiency gap: it is not an independent benchmark[38].

Let us extend our discussion to the case where technology does not evolve in a unique way. This means that different patterns of technological evolution are possible, depending on the criteria that underlie innovation. Insofar as business innovates, new technology must meet the requirements set out by business, which may contrast with those of other agents. Should this be the case, business requirements would produce a *technological bias* in favour of business, at the expense of other agents.

A technological bias has a twofold nature. First, the selected technology may not be the one that allows the highest output to be achieved, all other things given[39]. Second, the adopted technology may lead to a distribution – or, more generally, to a quality of life for some sections of a community – that is inconsistent with prevailing social values. Under either of these circumstances, technology is not an appropriate benchmark. It does not provide an external and objective criterion to ascertain whether economic efficiency is consistent with social efficiency.

When Dosi and Marengo refer to competitiveness, they (implicitly) assume that profit-seeking behaviour is consistent with the requirements set forth by workers and consumers (and society as a whole). Distributional conflict is substantially irrelevant and so is any conflict between consumer requirements and profit-seeking activity. Based on these premises, there is no scope for the insurgence of a technological bias at the expense of these agents. On strictly qualitative grounds technology reflects collective wants and may be used as an efficiency benchmark. Insofar as technology is qualitatively neutral – in that it reflects collective requirements rather than some vested interest – technological unemployment cannot be the intentional outcome of innovation, and the conventional claim that business should not be held directly responsible for unemployment remains true. Unemployment is a side-effect of economic activity, much like externalities.

The above assumptions underlying the qualitative nature of technological evolution are open to doubt[40]. It is therefore important to examine what may occur if we relax them. Marglin's (1976) original contribution still provides some interesting insights. He discusses the factory system in much the same way as Leijonhufvud: the division of labour underlies the introduction of specific techniques of production and the consequent development of new technology. What distinguishes his approach, however, is the determinants of the division of labour. Whereas Leijonhufvud focuses on a rather ahistorical notion of cost-minimizing teamwork, Marglin stresses that it is important to know who directs the production process. He argues that the factory system is the outcome of a historical

process where capitalists created the technical conditions that allow them to claim a right to a share of income. Thus, distribution is a key issue from the very outset or, to put it bluntly, it is not the division of labour (and technology) that eventually gives rise to capitalists; it is capitalists who determine the division of labour (and technology).

Much like in Veblen's approach, Marglin argues that economic efficiency consists in profit – as opposed to output – maximization and that the (efficiency) gap between these two outcomes corresponds to the social inefficiency of the capitalist mode of production. His perspective, however, suggests that a technology bias is also possible. In this perspective, so far as technological evolution depends on the intended actions of business, the insurgence of unemployment need not be a mere side-effect of economic activity. It may be a social cost of capitalist accumulation.

Technology, competitiveness and distribution

The above discussion suggests that two views confront each other: the *competitiveness view* and the *distributional view*. The competitiveness view claims that firms try to improve their production process and/or their products and that, insofar as this meets the requirements set by customers (and workers), technology cannot be questioned. Externalities may exist but this has to do with the market, not with the rationale underlying the innovative behaviour of firms.

We believe that this is a rather restrictive view. Leijonhufvud's two-step analysis – whereby teams try to raise the level of output, which is subsequently subject to a distributional conflict – may be heuristically useful but the second step of his analysis is not the end of the story. If we try to reframe it in an evolutionary perspective, a range of unanswered questions arises. When workers resort to unions and claim ever-increasing shares of output, why should capitalists not shift their attention from output-increasing technology to technology that reduces the bargaining power of the workers? In other terms, when teams turn into capitalist firms, why should their behaviour remain unchanged?

As for Dosi and Marengo, can we actually draw a clear-cut distinction between strategizing and the pursuit of competitiveness? Consider, for instance, qualitative changes in a product that raise entry barriers. They may be conceived of as competitive improvements which lead to a temporary monopoly or, alternatively, as a pre-emptive investment that is being window-dressed. What is more, firms may actually pursue both goals at the same time: the peculiarities of a product characterize it and position it in the market; at the same time, they may create a niche that isolates it from competition[41].

Let us suppose, however, that a distinction between the two strategies is possible. Why should a firm neglect strategising if, given the time range it deems important, such a strategy leads to higher net returns than a competitive strategy? Truly, one might question the appropriateness of short-term strategies that neglect long-term survival but short-termism may well be a reasonable reaction to external circumstances, e.g. high interest rates (Perelman 1996).

The distributional view argues that the profit motive depends on distributional conflict. Whether the key issue is distribution between firms and customers or bosses and workers is a minor issue; what is important is that economic efficiency may well conflict with social efficiency. Independently of externalities, business need not be consistent with social goals, thereby determining social costs.

The problem with this approach is that there is no clear-cut way to assess if and when the potential divergence between the outcomes determined by business and socially desired ones actually arises. To clarify this point, let us consider Marglin's approach. Although it is insightful, it tends to disregard that, when firms try to outcompete each other, they may cut costs but they may also improve the quality of the products they sell. This entails that the division of labour, the organization of production and, ultimately, technological change may be directed towards goals other than distribution between 'bosses' and workers and internal bargaining power. The pursuit of qualitative competitiveness, in turn, may produce path-dependent patterns of technological change, which differ from those that distributional goals within the firm would lead to.

The upshot of these remarks is that there is no *a priori* reason why firms should restrict their activity to either 'competitiveness' or distribution: the two motives co-exist. Furthermore, it is reasonable that firms adapt technology to their needs but it is rather difficult to assess whether this determines social inefficiency[42]. This implies that technology does not provide an appropriate account either of the efficiency gap, which depends on the extension of the technological frontier, or of the efficiency bias, which depends on the direction of technological evolution.

An important implication of the above discussion is that, since technology is a key feature in the labour demand function, labour demand cannot be conceived of as an unbiased function that workers merely have to acknowledge and adapt to. Since unemployment is associated to the technology underlying labour demand, long-term employment policies should not disregard the direction, as well as the intensity, of technological evolution.

Research, however, occurs also in other institutions, such as universities. It is therefore important to assess whether scientific and technological research carried out by these institutions may provide the missing benchmark, i.e. whether it may allow the assessment of the efficiency gap and bias determined by business.

The creation of technology: science and vested interests

Technological research is a search process. It ranges from finding the solution to a very specific problem to investigating a field of knowledge that may, some time or other, turn out to be profitable. In business, it may consist in adaptations on the shop floor as well as in the activity of a think tank which is only indirectly related to production. In the latter case, technological research is basically the same thing as any other scientific research. In order to appreciate how it is carried out in practice, it is important to understand its *general goals* and its *procedures*.

Let us first consider general goals. Two extreme situations may occur, which may help to appreciate the above examples. The first one is *skill-centred*

problem solving. It is a goal-related process where the agent involved knows what he/she is looking for and knows how to pursue his/her search. M. Polanyi (1962) provides an example by considering somebody who is looking for her pen. She knows what she is looking for, she will be sure that she ended her search once she finds it, and she (roughly) knows where and how to search.

The other case is *science-oriented problem solving*. This is a self-referential process where the agent does not fully know what she is looking for and may not even know what procedures are more appropriate. A case Polanyi considers is research in mathematics, an open-ended process of inquiry where no specific goal exists and where the rules of the inquiry are established during the process itself.

Let us now consider procedures. They must deal with at least three issues. First, the *boundaries* of the inquiry must be established (Georgescu-Roegen 1976). In a world where everything is – in some way or other – related to everything else, the researcher must decide where to stop searching. The decision relates both to the field of inquiry (spatial, disciplinary, etc.) and to the time assigned to the search process. However he/she sets the boundaries, the agent involved takes a risk: beyond the boundaries he/she sets, something important may be missed and there is no way to know; within the boundaries chosen, a great deal of material may be of no use whatsoever and in some instances, it may even be a 'noisy' impediment to the search process.

Second, an *aspiration level* is required to decide where to draw the boundary (Simon 1979). An aspiration level depends on the subjective characteristics of the agent involved. It also depends on the requirements he/she has to meet. These relate to method: how the community he/she refers to believes things should be done[43]. They also relate to goals: it makes a difference whether the outcome of the search process is expected to be knowledge for its own sake or knowledge that must be put to use in a profitable way[44].

Third, setting a boundary involves some *prior knowledge* of the issue. This may consist in specific knowledge available at the outset. It may also include heuristics that suggest how to go about the search process. Generally these heuristics are the result of past search processes (Egidi 1992; Dosi and Egidi 1991). They are applied to the new one because most often they are the only learning rules available. This is important because it stresses that learning is a path-dependent process, based on past goals and past procedures.

Researchers – or a research community – are never 'unbound'. The procedures they resort to depend on the type of problem they have to solve. Thus, in science-oriented problem solving, the establishment of appropriate procedures is part of an ongoing search process. It is up to the research community to choose what they deem appropriate.

In skill-centred problem solving, as well as in other goal-related search processes, procedures are functional to the specific goal pursued. When the latter is determined by business requirements, the boundaries of the problem and the aspiration level are functional to those requirements. The ensuing technology and knowledge reflect the business perspective that originated them and are consistent

with it. They also feed back on subsequent inquiries in a threefold manner. First, old knowledge provides whatever background information is used for subsequent inquiries. Second, the nature and extension of that background information depends on the procedures adopted, thus on the priorities deemed appropriate for the past inquiry. Third, since old knowledge is used to act upon reality, it generally determines an irreversible change in the existing opportunity sets. Based on these circumstances, learning builds upon previous knowledge according to what we may refer to as a learning path dependence.

Let us now consider independent research, i.e. research where specific goals are not associated to business requirements. Usually skill-centred problem solving takes existing technology for granted. It is, therefore, influenced in a strong way by learning path dependence and by the business priorities implicitly accepted in most inquiries. Science-oriented problem solving is a more intriguing case. In fact, as we move from skill-centred problem solving to science-oriented problem solving, the boundaries of the system under inquiry, as well as the required procedures, are ever more difficult to define in a unique way. Heuristics become a crucial tool. But heuristics basically are past learning experience: there is no way to claim a priori that what was appropriate for one aspect of reality (before) is going to be appropriate for another aspect of reality (now). Furthermore, heuristics are not fully under control because they involve tacit, as well as non-tacit, knowledge. The tacit dimension comprises views that are absorbed indirectly, basically through the normative and cognitive features of institutions (Scott 1995). So the issue is where that tacit dimension comes from, or better still, how it is originated. Given the central role that economic activity has in capitalist societies (K. Polanyi 1957), it should be of no surprise that it is originated – or at least affected – by business requirements.

There are two ways that a business perspective may affect the tacit dimension of technology and (scientific) knowledge. The first way regards goals. It consists in the conflation of business goals and social goals. Independently of any purposive action that business might take to influence the learning process, research focuses on issues that arise in our society. It therefore tends to take it for granted that economic requirements must be met. Machine manufacturers, for instance, appreciate the goals of their customers – other businesses – thereby introducing equipment that meets the latter's needs[45]. Similarly, a common policy prescription is to foster strong cooperative links between business and universities, or other research centres. The expected beneficial outcomes are grounded on the assumption that what business does is socially efficient. This introjection of business goals assigns them an absolute value, so that they appear to be natural. The second way regards procedures. It occurs when the procedures that are required to solve business-related problems are conflated with procedures that are required to solve social problems. A typical instance where this situation comes to the fore is when issues concerning the provision of public goods (health, education, etc) are dealt with in a business-like way.

We discussed the creation of technology as a cumulative learning process based on general and specific goals, as well as on procedures. We also stressed

that business-related goals and procedures may be tacitly accepted within independent learning processes. Finally, we argued that this influence may persist over time owing to learning path dependence. The non-neutrality of technology implies that labour demand cannot be conceived of as a merely technical phenomenon. Insofar as unemployment depends on technological factors, it may be conceived of as a social cost of business activity.

This discussion suggests that when business-related knowledge is not consistent with other perspectives – thereby leading to social costs such as unemployment – it may be fairly difficult for the latter to prevail unless completely independent learning processes are enacted, i.e. processes where priorities are properly assessed rather than taken for granted. Independent research may provide an alternative to the prevailing business-induced criteria in science and technology provided that it has a benchmark to identify its goals and procedures. The next section turns to this issue.

A social efficiency benchmark

Following Sen (1984; 1992; 1999), our point of departure is that economic activity must ensure that people are free to choose how to conduct their lives. This involves that they must have access to a wide range of possibilities concerning what they can be and what they can do (functionings) and that they must be able to choose how to combine them in order to achieve their ends (capabilities). Unemployment stands out as an impediment to this notion of substantive freedom and it is fairly reasonable that action should be taken to prevent it. How it should be prevented, however, is a fairly complex matter.

Choice of how to conduct one's life is possible when the alternatives are appreciated: it makes a difference whether someone chooses a job out of a wide range of opportunities or takes it because he deems it the only one available. There is an information problem, here: for any given amount of vacancies, the agent must know what the choice set is, i.e. what job opportunities he has. A choice set includes the characteristics of the agent – his skills, his wealth, his preferences – and what is made available to him: jobs – with their skills requirements and their wage – and other circumstances that affect the quality of life: income, a range of 'external' facilities concerning housing, health assistance, baby sitting, etc., social relations, and social status.

Information, however, is not the key problem. What is more important is that the choice set may be narrow – i.e. there may be little to choose from – so that the agent is only formally free to choose. The extension of a choice set therefore is a crucial issue. It is particularly so because there is no reason to believe that it is given once and for all: it is itself a matter of choice. The circumstances that define the extension of the choice set are embedded in a *choice context* which frames the set. For instance, both a worker's skills and a job's requirements depend on available technology and on schooling and training facilities; 'external' facilities depend on some agency that provides them; social relations and social status depend on a, broadly defined, cultural context. Finally, as we shall discuss in

greater detail, preferences are associated with the cultural and knowledge context.

When unemployment arises – when the choice sets available to all agents are too narrow to allow labour supply and demand to match – this is not the natural outcome of exogenous circumstances; it is the consequence of past decisions which determined the choice context. Those decisions were the first step in a path-dependent process which eventually led to unemployment. When unemployment arises, measures to contrast it depend on the peculiarities of that process: decisions taken in the past widened or restricted the range of the present choice context, thus the range of identifiable solutions to present problems.

Random factors may, obviously, affect these processes and the ensuing choice contexts. It is, however, important to assess how sectional interests may have a role. From the perspective of the distributional view discussed above, 'bosses' may find it convenient to restrict the choice contexts of workers. This would prevent the latter from interfering with business decisions, including those that have to do with distribution.

Consider how choice contexts may be affected by the division of labour[46]. Given a set of tools and skills, different divisions of labour are possible. Whichever criterion is used to choose, the selected division of labour determines the routines of the firm and the tasks each worker has to carry out. This affects what workers must focus on and what they *need to learn* in order to carry out their tasks. It also affects what they *can learn*, because it determines the information they have access to[47].

The division of labour determines how creative or passive a task is, thus the extent to which it involves decision-taking. It ultimately determines the degree of centralization of decisions within the firm. Since learning both precedes and follows a decision[48], the less one has to decide, the less one learns. Those who suffer this deprivation of knowledge cannot appreciate what is going on. They need to rely on whoever has that knowledge for any interpretation of current events.

What this leads to is that the division of labour influences *economic hierarchies*: who has the power to decide how to carry out economic activities (Marglin 1976; Braverman 1974). It influences *learning hierarchies*: what goals are deemed important in the skill-centred problem solving activities that people carry out when they work, thus what competences and knowledge they have to acquire. Finally, since those goals reflect a value judgement, the division of labour influences *value hierarchies*: what is important and what is relevant when judging what is going on. What all this leads to is that the organization of economic activity according to business priorities affects competences, knowledge and individual (and collective) preferences in a way that is most likely to restrict the extension of choice sets and choice contexts, thereby reinforcing the process depicted in fig. 3[49]. From the viewpoint of our discussion of unemployment, this not only determines a social cost – here and now – to the unemployed, in that it reduces the capabilities of the agents involved. It may cause a further social cost by restricting future choice contexts, that is to say, the command people will eventually have over their lives.

Not all circumstances favour business, however. As far as the division of labour is concerned, rivalry between firms may change priorities so that its main goal may shift from ensuring social control within the firm to increasing the firm's standing on the goods market. Furthermore, as a result of the process that the division of labour initiates, the skills of the workers and the competences of the firms change. It is also most likely that the tools become obsolete and have to be substituted by new ones. At the very least a new division of labour has to adjust to these changes[50].

As for scientific research, although it is influenced by business, it is also subject to patterns of evolution which depend on circumstances fairly external to the business environment: public finance, career dynamics within academia, and the open-ended nature of the search process in fields of inquiry where immediate technological applications may not be intuitive. This makes scientific evolution partly exogenous: changes in technology and in overall knowledge may therefore result from unexpected breakthroughs in research.

Finally, business-related values may clash with extra-economic – e.g. religious – values. Thus, divergence over what should account for a given wage – e.g. productivity (in a broad sense) or a standard of living – may foster distributional conflict within firms. Ethical values – e.g. concerning what should be the appropriate behaviour towards child labour or political dictatorships – may lead to conflicts between firms and their customers[51]. These reactions may force business to change its strategies, thereby affecting the latter's priorities and the patterns of evolution it would desire.

Far from being a mechanistic outcome of business behaviour, the evolution of choice contexts is a typical case of what Kapp (1976b) termed 'sequential interaction within a cumulative process'. The openness of the process need not preclude predictability but it does make it most unlikely. What is crucial from our point of view, however, is that while a range of circumstances reinforces the dominant role of business requirements, there is also scope for 'protective countermoves' and for a 'double movement'[52], so that an alternative to the social costs of business – including unemployment – is possible.

Policy implications

Innovation and value judgements

In the labour market described by conventional theory a new technology may affect the marginal productivity of labour, thereby changing the shape of the labour demand curve. It may also affect the marginal disutility of labour, thereby changing the shape of the supply curve. The new equilibrium may involve a higher or lower level of full employment. The adjustment process is going to be possible only if relative prices reflect these changes. When this does not occur, or if it occurs 'slowly', institutional changes may be required to prevent unemployment.

When learning is taken into account, an effect of technological change is that agents must learn how to do new things or how to do old things in a new way.

Adjustment implies that firms have to change their competences. Although each firm has its own core competences (Teece 1988), which require internal learning and adaptation, the horizontal – or social – division of labour in modern economies enhances the complementarity of economic activities (Richardson 1972), so that what a single firm knows is strictly dependent on what others know. Thus, at the level of a single firm, it is reasonable to consider technology as a basically exogenous phenomenon that the firm simply has to adapt to[53]. At the level of business as a whole, however, technology appears to be much less exogenous, as we argued above.

The effects of technology depend on the requirements assigned to it. We argued that they tend to be business requirements but that alternative – or, possibly, complementary – requirements can also be taken into account. In this perspective, technology can be a function of the socially desired level of employment. Research goals and procedures would then have to reflect this social goal. This is basically what Kapp argued in 1976b:

> In the future the social and natural sciences and applied technology will have to be open to a much greater degree than in the past to environmental and ecological constraints and objectives, that is, to explicit societal needs and human requirements.
>
> (Kapp 1977: 539–40)

We merely extend his conclusion – and we believe he would agree – by arguing that employment is a 'societal need and a human requirement' just as the environment.

Which requirements – business or social – should be given priority to is a matter of value judgement[54]. From this point of view it is worth noting that employment issues raise more problems than environmental ones. Environmental concern has increased over the years. The widespread claim is that the future of humanity is at stake: the rate of depletion of natural resources, on the one hand, and the degree of waste associated to production and consumption, on the other, are changing the environment to the point that not only are our standards of living unsustainable but even the survival of the human race may be undermined. Faced with the 'doomsday argument' (Mayhew 1981), finding a solution to this problem would be in the common interest, even though such a solution would be harmful for some vested interest.

The common interest of finding a solution to unemployment is less intuitive. Unemployment is of little concern for many who have a job. Up to a point, it does not disrupt accumulation. Furthermore, even those who acknowledge that it is the cause of economic waste – in that unemployed resources imply a lower output – and of social grievance, may fear that action in favour of the unemployed will prevent the efficient functioning of the market[55]. Indeed, as we argued above, when labour supply and labour demand do not match, a most likely reason is that different terms of reference clash[56]. The meta-preferences of the workers and the technical requirements of business reflect different goals, and these goals are not

comparable in monetary terms. Thus, unemployment hardly implies doomsday; quite to the contrary, it may be deemed good.

Apart from value judgements, two objections are likely to be raised against the claim that technological change could be constrained by the socially desired level of employment. The first one has to do with the intrinsic nature of innovation. It argues that innovators' concern about unemployment is pointless because they cannot predict what consequences their activity is going to have, thus what effect it is going to have for people's jobs. Furthermore, even though technology is endogenously created, the psychological intricacies underlying innovation – the creation of novelty – preclude whatever action aims to direct or even influence the mindsets of innovators. These objections are noteworthy but they require some qualification. It is true that the effect that technology has on employment is often unintentional. But this unintentional effect is more likely to occur when research is based on goals and procedures that simply disregard unemployment. Furthermore, innovative activity occurs within paradigms and trajectories that reflect, among other elements, learning path dependence, thus previously established goals and procedures. Even though single innovators cannot be forced to achieve a specific goal, the guidelines of their innovative activity may be based on different social priorities.

The second objection is that any constraint on technological change would feed back on profitability, thereby leading to a drop in output and/or employment. We argued above, however, that reliance on a direct relation between profitability on the one hand and output and employment on the other may be misleading. We stressed that unemployment is not (only) the result of some market failure, i.e. the undesired outcome of individual behaviour. It is (mainly) a social cost, determined by the way business operates, that is to say, by its goals and its strategies.

Truly, technological change that does not reflect business priorities may negatively affect key economic variables such as profitability, accumulation, competitiveness, etc., and this might be the cause of a different range of social costs. The identification of an appropriate policy requires a valuation of all the social costs involved. It does not warrant disregard for the dependence of technological change on sectional interests.

Employment policy and technology

Let us now focus on the implications this has from a policy perspective. Much like other agents, policy makers take their decisions in the light of their 'ends in view' – for any given choice set – and of their 'ends', the social goals pursued which define the choice context[57]. Leaving aside the former, recall that the 'end' we are concerned with is the command individuals have over their lives. As far as employment is concerned, this implies an appropriate technological policy. In order to favour social – as opposed to business – priorities in the creation of scientific and technological knowledge, action must focus on learning strategies.

It is beyond the scope of this paper to provide an exhaustive list of possible measures. We only wish to outline some key issues.

The first field of action concerns scientific cooperation between business and public (or publicly funded) research organizations such as universities. There may be instances where such a cooperation is useful but, since business priorities generally do not coincide with social priorities, it may be appropriate that public organizations be equipped to identify how this diversity affects science and technology[58]. This public interference in the directions of scientific and technological inquiry could lead to what Bush and Tool (2001) call a Lisenko effect[59]. However, the claim that scientific inquiry should be subject to collective priorities does not imply either that it should comply with the ends in view, as opposed to the ends, of a government or, for that matter, of a community. It does suggest that scientists should not feel absolutely free to do as they like. There is no doubt that, given the open-ended nature of a great deal of scientific research, it is fairly difficult for anyone but the scientific community to claim that a field of inquiry is more promising than another. Nonetheless, it is important to keep in mind that what 'promising' means ultimately depends on the goals pursued, thus on how the scientific community interacts and dialogues with different sectors of society.

The second field is education. A rather frequently held claim is that education should allow students to find a job. This cannot be denied but it should not prevent the achievement of a different goal: providing students with an understanding of the world we live in, so that they may become responsible citizens[60]. Training and general education need not coincide. This may be one of the reasons why a problem arises with labour supply: potential workers may wish to act as responsible citizens, thereby seeking jobs where they can use their knowledge to the advantage of the community, but this may turn out to be at the expense of the firm that hires them.

These considerations on education suggest that a different view may be required of educational entry barriers. Free access to schools and universities raises costs and sometimes precludes teaching efficiency. Furthermore, since students are not selected in advance, some give up studying when they realize it is too demanding a task for them. Drop-outs appear to be a net loss in that they do not achieve their degree despite all the money the community spent for them. These issues should not hide the fact that, for the reasons outlined above, education is a social – as well as an individual – goal in itself, quite independently of whether a degree is achieved or not.

A final field of action concerns firms and their research activity. As we recalled in section two, unemployment relief systems and organizational rigidity within firms are claimed to prevent workers from adapting to the requirements of labour demand. Relief systems and rigidity may be viewed as socially appropriate, however, precisely because they shift the adaptation requirement to firms: the latter would have to adapt their division of labour, their skill requirements and, eventually, their machinery to what workers look for.

Conclusions

The paper focused on a specific but crucial feature of unemployment: the mismatch between labour supply and labour demand. It stressed that the motives underlying labour supply may lead workers to pursue unavailable jobs rather than fill existing vacancies and that workfare programmes attempt to make it less convenient for jobless workers not to accept the jobs they are offered.

The rationale of workfare is that mismatches occur because of market failures and that incentives must be conceived of so that, when mismatches do appear, supply may adjust to demand as soon as possible. We investigated the motives underlying labour supply and stressed that there is more to mismatches than mere market failures. While the above motives of jobless workers may be 'bad' – in that they prevent adaptation to demand – they may be 'good' in terms of existing social values. In other terms, it is reasonable that individuals attempt to upgrade their social status despite labour demand requirements: it is reasonable – as well as socially appropriate – that they pretend to find jobs that put to use the qualifications they have acquired; it is reasonable – as well as socially appropriate – that they pretend to design life trajectories that enhance their present and future capabilities.

We also investigated labour demand as a function of existing technology. Leaving aside possible market failures, conventional economics has it that when labour supply adapts to labour demand, it is merely adapting to existing technological constraints. This view, however, neglects that a great deal of research has pointed to the – at least partial – endogeneity of technology. It also neglects that business goals – which need not coincide with social values – strongly influence how technology evolves.

Our discussion therefore pointed to the absence of an objective (i.e. generally acceptable) benchmark to assess whether it is labour supply that should adapt to the requirements of labour demand or the latter that should adapt to the former. We contended that the identification of a benchmark implies a value judgement. We consequently chose to base our discussion on Sen's notion of capabilities.

According to the capabilities view, we argued, there is no reason why the first envisaged alternative – supply adapts to demand – should be preferred. Insofar as technology is endogenous, economic action occurs both at the level of prices and quantities and at the level of learning, knowledge, and technology. The latter is not a mere constraint.

Technology may affect capabilities in terms of the freedom of choice available to individuals at any given moment. This is precisely what occurs on the labour market: when a mismatch occurs, individuals cannot satisfy their aspirations, and their freedom of choice is constrained.

Unemployment certainly causes social costs, at the psychological, social and strictly economic level, but what we wish to stress is that it is in itself a social cost: insofar as it is associated to a mismatch between business goals and social goals – where the latter aim to improve the quality of life – unemployment reflects the gap between potential and actual capabilities.

Economic policy is part of general public action. It co-determines what Kapp referred to as sequential interaction within a cumulative process. Single ends-in-view, such as providing everybody with a job, should be pursued with clear ends, such as increasing the capabilities of the individuals involved. This involves acting on how science and technology are created, bearing in mind that any public or private action that copes with today's unemployment problems also affects future technological development, with its opportunities and constraints. It is the contrary of workfare, which restricts the available choice set, reduces present and future choice contexts, and exerts no pressure on technology to meet the requirements set forth by unemployed workers.

The above discussion was restricted to unemployment as a mismatch. It did not deal with aggregate demand as a determinant of the level of employment. Our discussion suggests that the two issues are in no way inconsistent. Quite to the contrary, we believe they are complementary and that policies pursuing full employment should take into account the issues related to both.

Notes

1 A previous version of this paper was presented at the 2003 EAEPE Conference in Maastricht. We thank all participants for their comments. We also wish to thank Benedetto Gui, Angelo Reati and Stefano Solari for their comments. The usual disclaimer applies.
The authors are jointly responsable for the paper as a whole. For purely administrative reasons, however, M. Rangone is responsible for sections 1–4 while P. Ramazzotti is responsible for sections 5–8.

2 '(T)he fact that private entrepreneurs are able to shift part of their total costs of production to other persons or to the community as a whole, points to one of the most important limitations of the scope of neoclassical value theory.' (Kapp 1963: 11). See also Swaney and Evers (1989).

3 'The principle of cumulative or circular causation stresses the fact that social processes are marked by the interaction of several variables both 'economic' and 'noneconomic' which in their combined effects move the system away from a position of balance or equilibrium' (Kapp 1963: 25). A more in-depth discussion of the systemic openness of the economy is in Kapp (1976a).

4 'As long as it continues to confine itself to market value, neoclassical economics will fail to assimilate to its reasoning and to its conceptual system many of the costs (and returns) which cannot be expressed in dollars and cents.' (Kapp 1963b: 11).

5 '(T)o ignore social costs because they require an evaluation by society (...) and to leave social losses out of account because they are 'external' and 'noneconomic' in character, would be equivalent to attributing no or 'zero' value to all social damages which is no less arbitrary and subjective a judgement than any positive or negative evaluation of social costs.' (Kapp 1963: 23).

6 Employment protection schemes are less questioned, as their net effect on unemployment rates is not clear: they restructure labour market pools, by reducing both entry and exit flows, so that employment becomes more stable but, at the same time, unemployment shifts from short-term to long-term. See Blanchard (1999); Layard *et al.* (1996).

7 As a result, unemployment may turn from voluntary to involuntary in the long run.

8 Solow (1998) stresses that altruism has become scarce with regard to these issues. It is increasingly considered unacceptable that people earn subsidies even though they are in principle able to work. The thrust of the argument is that, in a world where wealth is so

widespread, the responsibility of deprivation should fall back on individuals, rather than on society. For a different view on subsidies, see Prasch 2002.

9 By devising the organization of production – thus, skills, tasks, recruitment practices etc. – employers define the characteristics required to participate in the production process. By the same token, they define the characteristics that training programmes must provide to employees in order to ensure job-readiness.

10 It is well out of the scope of this paper to assess whether such an approach to policy perfoms better than the Keynesian one in inducing employment growth, or to assess whether it is succesful or not in terms of its own goal. Critical essays of the new orthodoxy along these lines are Gardiner (2000); Peck and Theodore (2000b); Kitson *et al.* (2000); Darity (1999). Under critical scrutiny, here, is the goal itself.

11 The case can be easily extended. For example, a graduate student may have a strong preference for working in his field of study. If he does not get the job, he may decide to remain unemployed and keep on searching, rather than accept any unrelated job – even if wages elsewhere are high.

12 We allow that wages in the undesired job rise up to a level that would *in principle* compensate for the loss of the valued features; we do not discuss whether this level makes the activity with undesired jobs unprofitable.

13 As has been shown, conformity was a primary concern for Adam Smith, who believed that an efficient self-regulated economy needs a firm moral basis (Loasby, 1996: 315; see also Hirsch 1976: 137). Honourability is a key standard, in this respect. It is achieved by seeking esteem in the eyes of others, by doing what the community thinks is worth doing. Gains based on self-interest coincide with collective improvement insofar as they do not affect standards of behaviour negatively.

14 Achieved status through occupation represents a remarkable innovation compared to pre-industrial societies, where status was essentially ascribed (e.g. derived from non-modifiable traits such as race, ethnicity, sex, physical features, or transmitted to a person through inheritance); see Tumin (1967).

15 According to Tumin (1967), occupational status is valued along three dimensions: honour, that is the respect and deference that usually derive from the social consequences of the role (possibly we may also add power, following Blau 1964/1986); popularity, which depends on the attractiveness of roles (possibly based on scarce personal characteristics); and desirability – so similar a concept to the economic notion of preference – which reflects tastes as well as dominant cultural values in the community. These dimensions accrue to occupations in different amounts and reflect different allocations of job characteristics and rewards (e.g. income and other non-monetary factors).

16 We may think of the primary and secondary labour markets in labour market segmentation theory (Doeringer and Piore 1971).

17 Alternatively, we may refer to the concept of metapreferences (Sen 1982) as guiding principles in the solution of preference conflict. The two representations do not coincide: the former brings in discontinuity in the preference set (Georgescu-Roegen 1954), while the latter may determine its incompleteness (Sen and Williams 1981). Nonetheless, they may easily lead to similar practical consequences, e.g. that some choices will dominate others.

18 Bourdieu calls it 'hysteresis of habitus'. It is worth noting that the very nature of the habitus is its informality. It does not require that people take deliberate or even conscious actions to perpetuate differences. The 'art of living' is connaturate to the habitus of the upper classes. A neat difference can be found here with respect to conspicuous consumption in Veblen (1899), who held that the leisure class made an effort to denote its distinctiveness.

19 There probably is also a vicious circle of deprivation (Lipset-Bendix 1959), although this is more questionable (Heath 1981).

20 Social scarcity occurs when satisfaction from using something is reduced as the number of users increases. Hirsch explicitly refers to job opportunities as a salient domain of application of the idea of social scarcity.
21 It is worth stressing that what is *valuable,* no matter whether for a person or a collectivity, is not something immutable: it comprises relatively stable moral attitudes as well as material displays, whose definitions and relative weights in the evaluation standard change over time.
22 The rate of conversion of the two forms of capital is perhaps the most important battlefield in social dynamics, according to Bourdieu.
23 These factors interact to determine the process of revision, which will reasonably take longer the larger the divergence, the higher the motivation to stick to the original choice and the larger the amount of resources available to carry on the search process.
24 The pattern appears to be well-rooted: 'economic growth by itself, with no redistribution, could remove the servants from the homes of the middle class, because a less hard-pressed population would no longer choose to send its daughter to clean other families' houses' (Wicksteed 1910/1933) quoted in Hirsch 1976: 23).
25 Take the motivation of profit and the invisible hand as an example of how individual interest may trigger social benefits.
26 Technological mismatch follows the same course of action, if emphasis is on skill requirements.
27 For instance, manual work has become stigmatized, especially in agricolture. On the contrary, even manual works in services are deemed more honourable (think of McDonald's or of call-centres). It is likely that this process may be reversed with some social advantage.
28 See Acocella (1998) for an outline.
29 In the discussion that follows, maximization is used in a loose sense. It is not meant to imply substantively rational behaviour (Simon 1976).
30 'The common good, so far as it is a question of material welfare, is evidently best served by an unhampered working of the industrial system at its full capacity, without interruption or dislocation.' (Veblen 1919, Ch. 5).
31 'The business man's place in the economy of nature is to 'make money', not to produce goods.' (ibid. 1919, Ch. 5).
32 For instance, insofar as prices are set at too high a level to absorb potential output, a conflict ensues between what businesses pursue and what a community needs.
33 The centrality to Veblen's research programme of the themes developed in *The Engineers and the price system* is stressed by Knoedler and Mayhew (1999). See also Knoedler (1997).
34 Bush and Tool (2001: 211), for instance, refer to the possibility of an 'encapsulation of technology within the ceremonial value structure of the community'.
35 Some of the authors we will be discussing disregard some of the above elements but the overall picture we provide should not be misleading.
36 Whether a machine actually exists that can substitute a worker, or how it comes into existence, lie beyond the field of inquiry set out by Leijonhufvud. In this sense, there presumably exists a dimension of technological change that he deems exogenous.
37 Competences are also referred to as capabilities. Slight differences exist in the meanings assigned to the two terms – Dosi *et al.* (2000) provide an exhaustive survey of the issue – but they are not particularly important from our perspective. What we wish to avoid is confusion between the notion of 'capability' in the theory of the firm and organization theory and the notion of 'capability' as it is used by Sen. We will therefore use 'competence' when referring to the former.
38 Thus, actual output may be compared to potential output measured in terms of the existing technology (efficiency gap) or of potential technology (technology gap).

39 This gap between potential and actual output resembles a technology gap. The latter, however, is associated with a single technology whereas we are referring, here, to a range of potential technologies.
40 In a subsequent work, one of the cited authors acknowledged that while this approach stresses 'the *coordination* and *problem-solving* nature of organizational routines.' (Coriat and Dosi 1998; 104; emphasis in the original), it 'neglects the second major role of organization and organizational routines, namely their being a *locus of conflict, governance, and a way of codifying microeconomic incentives and constraints'* (ibid.; emphasis in the original).
41 There is no reason to believe that the market can prevent such a behaviour. Customers can assess the quality of goods only within the context of the existing technology. They can judge whether a pharmaceutical product will relieve them of their headache. However, they will hardly be able to find out whether a technically (and economically, from their point of view) viable alternative would be possible beyond that technology. This issue is discussed in greater detail by Ramazzotti and Rangone (2002).
42 The difficulty becomes even greater when strategies are not the deterministic outcome of economic relations, i.e. 'single-exit solutions' (Latsis 1976). See Groenewegen and Vromen (1997) for a discussion of this issue with special regard to the theory of the firm.
43 A Popperian scientist, for instance, would have to formulate falsifiable hypotheses.
44 Loasby's (1991) extremely insightful discourse is based on the assumption that businessmen and economists learn in the same way. A cautionary note, however, would be required with regard not only to different degrees of rigour but to the profitability constraint that businessmen must absolutely meet.
45 This implies that, although Leijonhufvud and Marglin provide a causal account whereby the introduction of machinery follows changes in the division of labour, the temporal sequence may be different: the introduction of new machinery may be the means to carry out a division of labour. A *post hoc ergo propter hoc* account of reality would, however, claim that 'exogenous' technological innovation is the cause of an objective division of labour.
46 The division of labour is only one instance of how choice contexts may be affected. For instance, economic, learning and value hierarchies are also influenced by how advertisements match consumption patterns to the composition of output (Galbraith 1958).
47 Since the competencies of the firms depend on the skills of the workers, as well as on the routines, they too are affected (Ramazzotti 2004).
48 A decision involves a preliminary assessment of the issue but also a subsequent appreciation of the assessment in the light of the outcome.
49 An example of the restriction of choice contexts is provided by emulation, as opposed to revolt. It is in the interest of business to foster emulation in the workplace, through career management, as well as in consumption, through advertisements. 'Emulation replaces the urge to revolt against the top people with the desire to climb into their ranks. It is the strongest of all social control mechanisms. When people are harmed by vested interest, emulation leads them to want their own vested interest. It is not the vested interest to which they object, but their own lack thereof.' (Dugger and Sherman 2000: 71).
50 Should this not occur, 'learning path dependence' might constrain or even preclude subsequent choices of technology, thereby determining a 'competence trap' (Levitt and March 1988).
51 The consensus pursued by business by diverting dissatisfaction towards emulation can hardly be achieved when aggregate demand falls, thereby reducing both consumption levels and opportunities to improve one's job. The ensuing frustration may well expand the scope for social conflict. Thus, even though emulation is functional to business, it may nonetheless clash with business requirements.

52 These are key concepts in K. Polanyi (1944); the relation between these concepts and Kapp's notion of social cost is discussed in Swaney and Evers (1989).
53 Owing to the above-mentioned technological interdependence among firms, however, adaptation is not a firm-specific process. Each firm can adapt only in relation to how other firms adapt.
54 Since the nature of technology depends on which priorities prevail, when we refer to a technological bias we do not imply that an 'unbiased' technology is possible. Rather, we believe that, depending on which priorities are chosen, some bias may be better than others.
55 An interesting case, in this respect, is that unemployment is likely to favour discipline. This issue is stressed by scholars who follow quite distinct approaches. See, for instance efficiency wage theory (Shapiro and Stiglitz 1984) on the one hand and Kalecki's (1943) political limits to full employment. The point, here, is to ascertain whether the negative consequences of unemployment are offset by the outcome of economic activity. In most instances, it is those who are not unemployed that enjoy the beneficial outcome of unemployment. A distributional issue therefore emerges which is in stark contrast with the common interest envisaged when discussing environmental issues.
56 'It is needless to say that the fact that problems of social costs raise issues of income redistribution makes them matters of political controversy and political power.' (Kapp 1963: 15).
57 The notions of 'ends in view' and 'ends' are discussed in Tool (2001).
58 For instance, rather than focusing on how the division of labour may best comply with the profitability goal, research could try to draw inspiration from the social goals underlying the labour supply function in order to devise alternative ways to organise economic activities.
59 A Lisenko effect 'increases the degree of ceremonial dominance through the displacement of instrumental values by ceremonial values. It is this process that is defined *regressive institutional change*.' (Bush and Tool 2001: 215).
60 It is no coincidence that this was a favourite expression of Federico Caffè, one of Italy's major estimators of Kapp (see Caffè 1978), who used it to point out his main objective when lecturing and supervising.
61 For a synthetic definition, see Aschaffenburg and Ineke (1987).

References

Acocella, N. (1998) *Foundations of Economic Policy: Values and Techniques*, Cambridge: Cambridge University Press.

Aschaffenburg, K. and Ineke, M. (1987) 'Cultural and Educational Careers', *American Sociological Review* 62, 4: 573–87.

Bandura, A. (1986) *Social Foundations of Thought and Action. A Social Cognitive Theory*, Englewood Cliffs: Prentice Hall.

Blanchard, O. (1999) *European Unemployment: The Role of Shocks and Institutions*, Roma: Edizioni dell'Elefante.

Blau, P. M. (1964/1986) *Exchange and Power in Social Life*, New Brunswick: Transaction.

——. (1994) *Structural Contexts of Opportunities*, Chicago: Chicago University Press.

Bourdieu, P. (1979) *La distinction*, Paris: Les éditions des minuits.

Braverman, H. (1974) *Labor and Monopoly Capital*, New York: Monthly Review Press.

Bush, P. D. and Tool, M. R. (2001) 'The Evolutionary Principles of American Neoinstitutional Economics', in K. Dopfer (ed.) *Evolutionary Economics. Program and Scope*, Boston: Kluwer.

Caffè, F. (1978) *Lezioni di Politica Economica*, Torino, Boringhieri

Coriat, B. and Dosi, G. (1998) 'Learning how to Govern and Learning how to Solve Problems: On the Co-Evolution of Competences, Conflicts and Organizational Routines', in A. Chandler, P. Hagstrom and O. Solvell (eds) *The Dynamic Firm: The Role of Technology, Strategy, Organization, and Regions*, Oxford: Oxford University Press.

Covick, O. (1998) 'Self-employment as disguised unemployment', in T. Lange (ed.) *Unemployment in theory and practice*, Cheltenham: Elgar.

Darity, W. Jr. (1999) 'Who loses from unemployment', *Journal of Economic Issues* 33: 491–96.

Davidson, P. (1994) *Post Keynesian Macroeconomic Theory*, Aldershot: Edward Elgar.

Dosi, G. and Egidi, M. (1991), 'Substantial and Procedural Uncertainty', *Journal of Evolutionary Economics* 2: 145–168.

Dosi, G. and Marengo, L. (1994) 'Some Elements of an Evolutionary Theory of Organizational Competences', in R.W. England (ed.) *Evolutionary Concepts in Contemporary Economics*, Ann Arbor: University of Michigan Press.

Dosi, G., Nelson, R. R. and Winter, S. G. (2000) 'Introduction: The Nature and Dynamics of organizational Capabilities', in G. Dosi, R. R. Nelson and S. G. Winter (eds) *The Nature and Dynamics of organizational Capabilities*, Oxford: Oxford University Press.

Dugger, W. H. and Sherman, H. J. (2000) *Reclaiming Evolution. A dialogue between Marxism and Institutionalism on Social Change*, London: Routledge.

Eatwell, J. (1997) 'Effective Demand and Disguised Unemployment', in J. Michie and J. Grieve Smith (eds) *Employment and Economic Performance. Jobs, Inflation and Growth*, 76–94.

Egidi, M. (1992) 'Organizational Learning, Problem Solving and the Division of Labour' in M. Egidi and R. Marris (eds) *Economics, Bounded Rationality and the Cognitive Revolution*, Aldershot: Elgar.

Feldstein, M. (1978) 'The Private and Social Costs of Unemployment', *American Economic Review* 68: 155–58.

Galbraith, J. K. (1958) *The Affluent Society*, London: Hamish Hamilton.

Gardiner, J. (2000) 'Rethinking Self-sufficiency: Employment, Families and Welfare', *Cambridge Journal of Economics* 24: 671–89.

Georgescu-Roegen, N. (1954) 'Choice, Expectations and Measurability', *Quarterly Journal of Economics*; reprinted in N. Georgescu-Roegen (1966), *Analytical Economics*, Cambridge Mass.: Harvard University Press.

——. (1976) 'Process in Farming Versus Process in Manufacturing: A Problem of Balanced Development', in *Energy and Economic Myths. Institutional and Analytical Economic Essays*, New York: Pergamon Press.

Groenewegen, J. and Vromen, J. (1997) 'The Theory of the Firm Revisited', in L. Magnusson and J. Ottson (eds) *Path Dependency and Evolution*, Aldershot: Elgar.

Hamilton, D. (1991) *Evolutionary Economics. A Study of Change in Economic Thought*, New Brunswick: Transaction Publishers.

Heath, A. (1981) *Social Mobility*, London: Fontana.

Henry, S. G. B. and Snower, D. J. (eds) (1996) *Economic Policies and Unemployment Dynamics in Europe*, International Monetary Fund.

Hirsch, F. (1976) *Social Limits to Growth*, Cambridge, Mass.: Harvard University Press.

Hirschman, A. O. (1970) *Exit, voice and loyalty: Responses to decline in firms, organizations, and states*, Mass.: Harvard University Press.

Kalecki, M. (1943) 'Political Aspects of Full Employment', *Political Quarterly* 14: 322–31.
——. (1963) *Social Costs of Business Enterprise*, Bombay: Asia Publishing House.
Kapp, K.W. (1976a) 'The Open-System Character of the Economy and its Implications', in K. Dopfer (ed.) *Economics in the Future: Towards a New Paradigm*, London: Macmillan.
——. (1976b) 'The Nature and Significance of Institutional Economics', in *Kyklos* 29: 209–232.
——. (1977) 'Environment and Technology: New Frontiers for the Social and Natural Sciences', *Journal of Economic Issues* 3: 527–40.
Kitson, M., Martin, R. and Wilkinson, F. (2000) 'Labour markets, social justice and economic efficiency', *Cambridge Journal of Economics* 24: 631–41.
Knoedler, J. T. (1997) 'Veblen and technical efficiency', *Journal of Economic Issues* 4: 1011–26.
Knoedler, J. T. and Mayhew, A. (1999) 'Thorstein Veblen and the Engineers: A reinterpretation', *History of Political Economy* 2: 255–72.
Latsis, S. J. (1976) 'A Research Program in Economics', in S. J. Latsis (ed.) *Method and Appraisal in Economics*, Cambridge: Cambridge University Press.
Layard, R. (1999) *Tackling Unemployment*, London, Macmillan.
Layard, R., Jackman, R. and Nickell, S. (1996) 'Combating Unemployment; Is Flexibility Enough?' in *OECD Proceedings: Macroeconomics Policies and Structural Reform*, OECD; reprinted in Layard (1999), Ch. 10.
Leijonhufvud, A. (1986) 'Capitalism and the Factory System', in R. N. Langlois (ed.) *Economics as a Process*, Cambridge: Cambridge University Press.
Levine, D. P. (1997) 'Knowing and acting', *Review of Political Economy* 9, 1: 5–17.
——. (1998) *Subjectivity in Political Economy*, London, Routledge.
Levitt, B. and March, J. G. (1988) 'Organizational Learning', *Annual Review of Sociology* 14.
Lindbeck, A. and Snower, D. J. (1986) 'Involuntary Unemployment as an Insider-Outsider Dilemma', in W. Beckerman (ed.) *Wage Rigidity and Unemployment*, London: Duckworth.
Lipset, S. M. and Bendix, R. (1959) *Social Mobility in Industrial Societies*, California: University of California Press.
Loasby, B. J. (1991) *Equilibrium and Evolution. An Exploration of Connecting Principles in Economics*, Manchester: Manchester University Press.
——. (1996) 'The division of labour', *History of Economic Ideas* 4, 1–2: 299–323.
Marglin, S. A. (1976) 'What Do Bosses Do? The Origins and Functions of Hierarchy in Capitalist Production', in A. Gorz (ed.) *The Division of Labour: The Labour Process and Class Struggle in Modern Capitalism*, Brighton: Harvester Press.
Mayhew, A. (1981) 'Ayresian Technology, Technological Reasoning, and Doomsday', in *Journal of Economic Issues* 2: 513–20.
Peck, J. and Theodore, N. (2000) 'Beyond employability', *Cambridge Journal of Economics* 24: 729–49.
Perelman, M. (1996) *The Pathology of the US Economy. The Costs of a Low-Wage System*, London: Macmillan.
Picketty, T. (2000) 'Theories of persistent inequality and intergenerational mobility', in A. B. Atkinson and F. Bourguignon (eds) *Handbook of income distribution*, Amsterdam: Elsevier.
Polanyi, K. (1944) *The Great Transformation*, New York: Holt, Rinehart & Winston.

——. (1957) 'The economy as instituted process', in K. Polanyi, C. H. Arensberg and H. W. Pearson (eds) *Trade and market in early empires: economies in history and theory*, New York: The Free Press.

Polanyi, M. (1962) *Personal Knowledge. Towards a Post-Critical Philosophy*, London: Routledge.

Prasch, R. E. (2002) 'What is wrong with wage subsidies?', *Journal of Economic Issues* 36: 357–64.

Ramazzotti, P. (2004) 'What Do Firms Learn? Capabilities, Distribution and the Division of Labour' in J. Foster and S. Metcalfe (eds) *Evolution and Economic Complexity*, Cheltenham: Elgar.

Ramazzotti, P. and Rangone, M. (2002) 'Firms and Selection: Possible Features of Innovation Processes', *Rivista Internazionale di Scienze Economiche e Commerciali* 1: 55–79.

Reiss, A. J. (ed.) (1961) *Occupations and Social Status*, New York: The Free Press.

Richardson, G. B. (1972) 'The Organization of Industry', *The Economic Journal* September; reprinted in G. B. Richardson (1990) *Information and Investment*, 2nd edn, Oxford: Clarendon Press.

Robinson, J. (1937) 'Disguised Unemployment' in *Essays in the Theory of Employment*, London: Macmillan.

Roemer, J. (1998) 'Comment', in R. M. Solow *Work and welfare*, Princeton, New Jersey: Princeton University Press.

Schelling, T.C. (1978) *Micromotives and Macrobehavior*, New York: Norton.

Scott, R. W. (1995) *Institutions and Organizations*, Thousand Oaks: Sage Publications.

Sen, A. (1982) 'Rational Fools: A Critique of the Behavioural Foundations of Economic Theory', in *Choice, Welfare and Measurement*, Oxford: Basil Blackwell.

——. (1984) *Resources, Values and Development*, Oxford: Basil Blackwell.

——. (1991) *Capability and Well-being*: United Nations University Press.

——. (1992) 'The standard of living', in *Tanner Lectures on Human Values*, vol. I, Salt Lake City: University of Utah Press.

——. (1999) *Development as Freedom*, New York: Alfred Knopf.

Sen, A. and Williams, B. (1982) *Utilitarianism and Beyond*, Cambridge: Cambridge University Press.

Shapiro, C. and Stiglitz, J. (1984) 'Equilibrium Unemployment as a Worker Discipline Device', *American Economic Review* 74: 433–444.

Simon, H. A. (1976) 'From Substantive to Procedural Rationality', in S. J. Latsis (ed.) *Method and Appraisal in Economics*, Cambridge: Cambridge University Press.

——. (1979) 'Rational Decision Making in Business Organizations', in *American Economic Review*.

Solow, R. M. (1998) *Work and welfare*, Princeton, NJ: Princeton University Press.

Swaney, J. A. (1989) 'Our Obsolete Technology Mentality', *Journal of Economic Issues* 2: 569–77.

Swaney, J. A. and Evers, M. A. (1989) 'The social cost concept of K. William Kapp and Karl Polanyi', *Journal of Economic Issues* 23: 7–33.

Teece, D. J. (1988) 'Technological Change and the Nature of the Firm', in G. Dosi, C. Freeman, R. R. Nelson, G. Silverberg and L. Soete (eds) *Technical Change and Economic Theory*, London: Pinter.

Tool, M. R. (2001) *The Discretionary Economy. A Normative Theory of Political Economy*, New Brunswick: SA Transaction.

Tumin, M. M. (1967) *Social stratification. The forms and functions of inequality*, Englewood Cliffs: Prentice-Hall.

Veblen, T. (1919/1964) *The Vested Interest and the Common Man,* New York: Kelley.

——. (1921/2001) *The Engineers and the Price System*, Kitchener, Canada: Batoche Books.

Waller, W. T. Jr. (1982) 'The Evolution of the Veblenian Dichotomy: Veblen, Hamilton, Ayres, and Foster', *Journal of Economic Issues* 3: 757–771.

Wicksteed, P. H. (1910/1933) *The Common Sense of Political Economy*, edited by L. Robbins, London, Routledge and Kegan Paul, 1933.

8 Social costs and human health: Kapp's approach and its growing relevance today

*Pietro Frigato**

Research in two recently developed 'hybrid disciplines' (Kapp 1961), socio-economic epidemiology and occupational safety and health economics, is of great interest and importance for an evaluation of the substantial performance of public policies and social institutions[1]. We now dispose of a growing body of empirical evidence attesting the avoidable health impairments of socio-economic inequality, of unemployment and of unsound living and working conditions under the conventional contractual obligations and legal institutions of historical capitalism and under the spreading, more precarious, work arrangements of the last two decades. The degree of penetration of economic de-regulation and privatization policies and their repercussions on population health could thereby become the object of systematic analysis and evidence-based criticism. In fact, the monitoring capacity provided by both these fields of inquiry helps to define the public health aspects of social value and is of fundamental importance for the measurement of the overall social efficiency of different regimes of mixed-market economies.

In the present paper I shall present a brief reconstruction of the existing state of the art within these two fields of inquiry in public health, with two major restrictions:

- The discussion of the (socio-economic) causal mechanisms underlying health disparities at regional or global levels will be restricted to the materialistic approach, with special reference to the recent Wilkinson/Coburn debate[2].
- Only limited theoretical and methodological aspects of occupational safety and health economics will be introduced; recent work conducted by Quinlan, Mayhew and Dorman allows for an exploratory discussion both of the prevailing empirical and institutional scientific literature and of the growing precarization and psycho-physical impairment of the human resource due to the massive introduction of non-standard work arrangements which has been visible over the last twenty years in most capitalist countries[3].

In both cases the shortcomings and the demonstrable incapacity of the price system to deal with some kinds of social choice and allocation problems are reflected in widespread and avoidable health impairments.

Kapp adopts an open-system approach for the study of the health and human costs of capitalism. Within his general critique of the typical logic and functioning of business enterprise and of the competitive mechanism, Kapp carefully investigates the causal institutional processes at the origin of unhealthy life and work conditions, unequally distributed among different social groups and classes. Kapp's lesson on this fundamental type of social costs has been totally ignored in socio-economic epidemiology and in occupational risk economics. I shall try to show that this omission leads to severe analytical shortcomings in socio-economic epidemiology. Occupational safety and health economics is more compatible with Kapp's analysis of the impairment of the human factor of production but its approach tends to be rigidly circumscribed and separated from other analyses of social losses. It may therefore fail to elaborate coherent proposals for public policy.

The debate on the social determinants of health

Especially since the mid-Nineties, the interdisciplinary study of the social, psychological, behavioural and biological mechanisms and pathways by which socio-economic status influences health has become the pivotal issue of most research in the field of socio-economic epidemiology (Adler and Ostrove 1999).

Richard Wilkinson's original explanation of the causal processes responsible for the strong association between socio-economic status and health (Wilkinson 1996) has exerted a major influence in the materialistic debate on causation in socio-economic epidemiology (Borgers and Abholz 2001). Wilkinson recognizes that 'there is a missing social economy of well-being' (Wilkinson 1996: 109). He tries to overcome this scientific weakness by arguing, on the basis of a large mass of statistical evidence, that, once certain absolute levels of GNP per capita are attained by the large majority of the population (about US$ 5,000 per capita in 1990), the major determinant of health differences between and within nations becomes the degree of income inequality, i.e. distributive justice: a thesis which seems to echo Hirsch's more speculative distinction between the 'positional' and 'material' spheres of the economy and which shares fundamental elements with Hirsch's analysis of the cumulative depletion of moral standards in advanced market settings (Hirsch 1976). Wilkinson seems to ignore this theoretically refined doctrinal antecedent and pays no attention to the debate that accompanied Hirsch's approach to market failures or to the social limits to economic growth (Offer 2000)

> In the main, however, the political contribution of *Social Limits to Growth* lies in its analysis of the costs of reliance on the free market as the main means of pursuing human ends, that is the cost of the policies of the so-called new right, that political force currently dominant in Washington, Bonn, London, and several major international economic agencies.
>
> (Crouch 1983: 186)

In Wilkinson's *Unhealthy Societies*, we find the clear assumption that an exacerbation of health disparities between social classes and groups is observable

wherever privatization and de-regulation policies, coupled with cutbacks in social security and insurance schemes, were introduced by government coalitions of the New Right[4]. Wilkinson tries to show that a higher degree of market penetration reinforces existing patterns of socio-economic inequality and, through psycho-social and social cohesion mechanisms, exacerbates morbidity and mortality.

Wilkinson explains the example of Japan and of its excellent health performance among advanced capitalist nations as the result of its 'groupish' social structure and of the moral institutions that exist beyond the private sphere of the family (Wilkinson 1996: 133).

Curiously enough, Hirsch's fundamental hypothesis, that moral standards, different from the individualistic ones promoted by the market, are required (while being unavoidably depleted in the course of time) even for the ordinary functioning of a market economy, has been deemed particularly appropriate to understand the success of the Japanese economy (Sen 1983).

Leaving aside theoretical difficulties bound to the meaning of concepts like privatization, de-regulation and liberalization in particular (Starr 1987, 1988), for the scope of the present paper we need only shortly remember that there are two forms of privatization: the privatization of individual involvements (an aspect emphasized, among others, by Hirsch and Hirschman) and the privatization of social functions and organizations. Both are based on the pursuit of private gain and while the first stimulates individualistic behaviours and attiudes, the second leads to the ceremonial celebration of private business enterprise (Starr 1987).

Wilkinson pays no specific attention to the role that private business enterprise and the degree of market penetration in existing mixed public-private economies have on the overall health performance of different institutional variants of advanced market economies. He starts with the assumption that, within and between advanced market economies, health is related to relative rather than to absolute income: this implies that 'health may not be strongly related to economic growth' (Wilkinson 1999: 257; Wilkinson 1996: 36). To him this means that what should count most for an understanding of the health effects of relative income are the 'direct psycho-social effects of low social status' and the 'poorer qualities of social relations found in more hierarchichal societies' (Wilkinson 1999: 260).

So conceived, health disparities are almost exclusively the result of individual perceptions of relative income: negative emotions (Wilkinson lists in particular, among other possible distressing feelings: shame, distrust, fears of incompetence, sense of inferiority, recurrent anxiety) and self-destructive practices tend to impinge negatively on individual health. At a social level, these negative self-perceptions are translated into socially adverse behaviours (higher levels of hostility, increased homicide and violent crime rates), attenuated social involvement and institutionalized patterns of behaviour which can negatively affect social cohesion.

Wilkinson's approach could be convincing and helps to consider the importance of social relativities in order to explain the social gradient in health. It is nonetheless doubtful that it may provide an exhaustive account for the health performance of specific national and regional socio-economic contexts.

Wilkinson believes that

> There is of course a long history of legal safeguards and reforms dealing with hazards that could be avoided by a quick technological fix: such as limiting exposure to toxic materials or putting guards on machines. While there are always complaints about the costs, and last-ditch criticism of the evidence from those whose financial interests are threatened, technical solutions have made major contributions to occupational and environmental health. But what is now coming out of health research, and particularly health inequalities research, is knowledge of the social causes of disease. Instead of exposures to toxic materials and mechanical dangers, we are discovering the toxicity of social circumstances and patterns of social organization.
>
> (Wilkinson 1996: 23)

This view seems to imply that we need not continue investigating the objective risks associated with the specific constellation of environmental factors that influence health

> If the health risks to which people are differentially exposed arose merely from the material circumstances in which we live, that would tell us very little about our society which we could not see with our own eyes.
>
> (Wilkinson 1996: 22-23)

Wilkinson's undervaluation of the objective environmental and material circumstances in which people live and work, as well as of the opacity of informations over their nature and potential dangers, leads to Coburn's criticism.

Within his widely debated paper on the neglected role of 'neo-liberal capitalism' and globalization in influencing both income inequality, social cohesion and, consequently, human health (Coburn 2000: 137; Borgers and Abholz 2001: 371–373), Coburn deliberately accepts the validity of Wilkinson's psycho-social explanation. He only contends that the causal analysis provided is insufficient, since a whole series of interdependent institutional factors and processes have been either ignored or considered outside the proper scope of socio-epidemiological analysis.

Coburn argues in favour of a broader perspective for the study of health inequalities in developed countries: he laments the 'striking' fact that, despite the diffused interest in the causal processes at the origin of health disparities, 'the possible causes of inequalities are seldom, if ever, mentioned' (Coburn 2000: 13). Coburn traces the underlying institutional factors of inequality back to 'obvious international and national political and economic trends', including 'the globalization of the world economy as well as the rise of New Right political regimes and the concomitant "decline of the welfare state"' (Coburn 2000: 137).

His argument can be summarised as follows: market dominance, resulting from the decline, since the Seventies, of the Keynesian welfare state and of the social-democratic and corporatist variants within which its distinctive institutional

arrangements have historically evolved, exacerbates economic inequality and worsens the health performance within and between industrialized nations

> The most recent evidence from the United States, Britain, Australia, Canada, New Zeland and the OECD countries generally, indicates that neo-liberalism in action, while obviously a far from perfect neo-liberalism, is associated with (more or less) rapidly increasing inequality. The US and the UK, but also Canada and Australia, show much higher inequality than do such countries as Switzerland, Germany or the Netherlands, who, in turn, show higher inequality than do Scandinavian countries.
>
> (Coburn 2000: 140)

Coburn's view is coherent with the so-called 'neo-material interpretations of health inequalities', which emphasize the role of 'social connectedness' for the explanation of health disparities (Lynch 2000: 1003). This concept is

> (...) deliberately broad and can be understood both as interpersonal, informal-horizontal links, and as formal-vertical links that characterize how individuals and groups are 'connected' to institutional, legal, political and economic structures.
>
> (Lynch 2000: 1003)

By putting this wide range of institutional factors at the core of a theory of population health

> (t)he neo-material interpretation of health inequalities is an explicit recognition that political-economic processes, discussed by Coburn in terms of neo-liberalism, generate income inequality, influence individual economic resources, and also impact community resources such as schooling, health care, social welfare, and working conditions.
>
> (Lynch 2000: 1004)

With its insistance on objective living and working conditions, the neo-material approach can help to explain the basic feature of current health trends within and between capitalist countries (an aspect to which the psycho-social approach seems ill-adapted)

> The fact that health differences can be observed between the top of the social hierarchy and the next level down is surely important, but may be an incomplete basis for understanding why the largest burden of excess deaths occur at the bottom of the social hierarchy.
>
> (Lynch 2000: 1004)

Both Wilkinson and Coburn argue in favour of redistributive policies by public powers to reduce health inequalities within industrialized nations. This is,

however, too restricted a view of the possible role of public intervention in the economy for the reduction of health disparities. A broader investigation of the role and functioning of policy-making in mixed-market economies would at least require the acknowledgement of Richard Musgrave's classical distinction between: state interference in private allocation for the protection and provision of 'social goods'; the public necessity of adjusting and correcting the unequal market pattern of distribution (through taxes and transfer payments); and macro-economic stabilization (the public pursuit of stable and full employment through monetary and fiscal policy) (Musgrave 2000).

Surprisingly enough, Wilkinson himself admits in his reply to Coburn that a broader spectrum of policy measures and public interventions that can influence the health performance of different economies should be considered besides income redistribution policies (Wilkinson 2000: 997–998).

Both Wilkinson and Coburn fail to offer a general framework of analysis for the study of health disparities in industrialized nations. They cannot help analysing such fundamental problems as: differential mortality and morbidity trends in rural and urban areas; the role of industrial location and urban congestion; organized pressures and influences by different vested interests on public regulators and administrators; 'strategic misunderstanding' activities by business corporations in the direction of the maintainance of high levels of profitable health risks for workers and human communities; delays, biases, duplications, obstructions and insufficient investment in scientific and technological research activities due to the specific mode of allocation of research funds by business concerns. These and other important institutional aspects and factors are left to contextual studies by applied epidemiologists, dealing with cases of industrial pollution associated with the introduction or treatment of specific substances by specific firms or industries at regional or global levels (Castleman 2001; Ludwig, *et al.* 2001).

Occupational safety and health economics and the precarization of the human factor of production

In a recent paper Peter Dorman has isolated three basic issues of occupational risk economics: the definition and measurement of the pecuniary costs of unhealthy working conditions; the investigation of the relationship between the behaviour of private businesses in the different branches of the economy, public regulatory regimes and their effects on workers' health in times of economic globalization; and the compatibilization and harmonization of policy measures devoted to minimizing the impairment of the human resource with other social priorities (Dorman 2000: 1–2).

There can be no doubt about the rapidly growing interest in the institutional analysis of occupational safety and health economics, especially since the middle of the Nineties. This is a consequence of the widespread and radical changes which have occurred over the past two decades in the contractual position of labourers in the large majority of advanced capitalist nations (Quinlan and Mayhew 1999: 4–5).

There is now sufficient evidence that capitalism has remained a system of unpaid costs, partially reflected in the pecuniary losses suffered by substantially impaired workers, their families and the general public. Growing empirical evidence confirms that in general the most dangerous jobs are those at the lowest level of the occupational scale: Dorman has recently defined 'precarious employment, informal employment, work in small and medium enterprises, and jobs performed by groups subject to discrimination and marginalization'[5] as the categories which are most exposed to risks of the generally undervalued or ignored costs of unhealthy working conditions (Dorman 2000: 5). It should be added that, as a rule, these disparities cannot be rigidly separated from other types of aggression: '(...) those who suffer the most from poor working conditions are also the most likely to bear other social and economic costs', thereby becoming victims of a process of cumulative vulnerablity (Dorman 2000: 7).

This distinctive feature of the distribution of health losses seems to be exacerbated by the global trend toward the flexibilization and active reduction of the portion of private business costs of the labour factor (changing employment and management practices of large corporations: outsourcing, franchising, downsizing) in order to favour a massive introduction of non-standard work contracts (i.e. precarious/contingent or informal) (Quinlan and Mayhew 1999).

Robust and mounting evidence shows that 'work which is precarious in employment terms is likely to be physically precarious as well' (Dorman 2000: 4): increased risk is associated with all forms of flexible contractual obligations (Quinlan and Mayhew 1999, Quinlan 2003). The causal mechanisms at the origin of these negative health effects are being investigated: outsourcing and contracting out pose pressing problems concerning inadequate training of workers and inadequate awareness of labour rights: 'in some instances they do not even know who their employer actually is' (Dorman 2000: 4). Private business calculus, especially in those sectors where time minimization and the employment of unskilled and precarious workers are financially more attractive, tends to induce the diffusion of *elusion* practices regarding health and safety measures. Leased or outsourced workers cause serious information problems to public statistical agencies: accident data may be biased by the difficulty in the precise attribution to one specific economic branch or business unit. In addition, workers employed precariously tend to have more reported injuries and lost workdays than regular workers in the same industries (Dorman 2000: 4).

While the diffusion of new contractual forms of employment poses new, complicated questions of coverage and liability to conventional workers' compensation systems, which will not concern us in the present paper, existing data also confirm the more fundamental trend toward increasing losses expressed in real terms (total/partial-temporary/permanent disabilities and fatal injuries due to work accidents; professional diseases and premature deaths induced by specific working conditions)[6].

In his analysis of the institutional context of the USA, Peter Dorman has pertinently observed that the changing composition of the economy, with an expanding service and administrative sector, will make it seem like work has

become safer (from toxic chemical exposures, radiation, noise, mechanical dangers and the like). Controlling for this effect through 'fixed weight indices', Dorman describes

> (a) long-term increase in the frequency of serious injuries: compared to 1980, we left the decade with nearly an eight per cent greater chance that a worker in a given industry would experience a significant health or safety problem.
>
> (Dorman 1996: 15)

According to Dorman, 'on the balance, the statistical evidence points towards worsening conditions in U.S. workplaces' (Dorman 1996: 18). While this is hardly surprising, it should be kept in mind that this happens in a situation in which we still lack information concerning the possible health repercussions of various levels of exposure to different, cumulatively interacting environmental conditions and stressors (Dorman 1996: 195; Ludwig *et al*. 2001).

Recent evidence shows that increased workforce volatility induces a rapidly growing obsolescence of conventional protective legislation in general. Michael Quinlan's remarks concerning the Australian context can be extended to most industrialized nations

> First, inconsistencies in key definitions ('worker', 'employee', 'subcontractor', 'volunteer', and 'family helper') and coverage provisions under OHS, workers compensation, industrial relations and even taxation legislation have contributed to confusion of the responsibilities and entitlements under the former. This links to an earlier point made in relation to business law, the corporate veil and the 'finessing' of legal categories to shift costs or avoid regulatory requirements. Second, the growth of contingent work arrangements has reduced the coverage (both formal and effective) of workers compensation regimes. (...) Aside from the cost-shifting implications (to social security and Medicare) it means that workers compensation claims – the major source of occupational injury and illness data in Australia including the National Data Set (NDS) – is becoming an increasingly partial and potentially unreliable source upon which to base preventative activities. This applies not only to industries where subcontracting is widespread (such as trucking, taxis, home based work and contract cleaning) but also where extensive use is made of leased workers or temporary workers. Leaving aside attempts to expand workers compensation coverage, OHS agencies have used a range of additional data sources including targeted industry/workplace audits and hospital admission records (specially adapted).
>
> Quinlan (2003: 12)

This corrosion of existing legal and institutional arrangements to minimize health risks on the job is also reflected in pressing problems concerning the adequacy of legally defined standard safety limits

> By altering the array of hazards or level of risk, changing work arrangements may be weakening the value of existing guidance material, codes and regulations. For example, a number of flexible employment arrangements are associated with changes to working time that may have critical relevance for existing exposure limits.
>
> (Quinlan 2003: 8)

Quinlan shows that changes to work timing would require a redefinition of legally defined maximum tolerable levels of a large variety of hazardous conditions. In addition, multiple jobholding exposes labourers to multirisk environmental conditions and tends to intensify injuries at work and to further obfuscate the traceability of professional diseases. Both of these circumstances offer concrete allocation incentives to private business initiative to further reduce the 'variable' costs of labour and lead to the institutionalization of various forms of unregulated exploitation of the human factor of production (Quinlan 2003: 8).

Normally, union contractual power and influence are reduced in all those economic sectors in which recourse to precarious employment is more profitable. Inadequate training, distorted perception of risks and insufficient knowledge of protective legal arrangements by contingent workers add substantially to the puzzling problems of the causal relationships between exposures and occupational diseases, and of the consequent obsolescence of inherited legal arrangements and administrative forms of social control of business.

These trends towards a growing social inefficiency of inherited legal and administrative arrangements for the protection and security of labour tend to reinforce cumulative pressures towards the institutionalization of new types of cost-shifting methods and practices, and/or the refinement of already-existing ones. The socially adverse effects of the rapid reorganization of production and of labour markets need to be analysed through detailed, contextual and institutional studies. Efforts to measure the monetary and real losses (expressed in terms of avoidable morbidity and mortality) associated with the shortcomings inherent to economic guidance under private business initiative must become elements of a more general strategy of inquiry: its main task must be to reduce the discretion of business entreprise in the selection of the new channels through which it can shift relevant parts of the total costs of production to the single worker, his family and the general public.

Social costs of business enterprise and public health in K. W. Kapp

Although Kapp's analysis of the health and human costs of business enterprise is based on different methodological premises, it certainly has much in common with current materialistic interpretations of health trends and disparities in industrialized nations: both consider the eco-social environment where people live and work (or search for jobs) as the main determinant of health.

Kapp insists on the role that the institutional structure of production and distribution has on ecological and social disruptions in modern capitalistic contexts.

His theory of the social costs of business enterprise is interdisciplinary and institutional: the causal chain at the origin of social costs is, in his view, the result of a cumulative process of complex interdependencies among physical and institutional factors.

For this reason an adequate inquiry in the nature and significance of social costs requires

> a much greater collaboration between the social scientists and technically competent experts in such fields as agriculture, engineering, public health, city planning and public administration than the traditional economist had been prepared to admit. (...) The time may not be far off when the economist will have to establish the closest possible cooperation with technical experts in engineering, public health and other fields of technology in order to identify and define social costs and social benefits which the market system tends to ignore.
>
> (Kapp 1963: 283, 289)

Climatic, geographic, demographic factors, the degree of urbanization of local communities, the location of industries, and technical aspects of production engineering are all interconnected elements which play a fundamental role in the explanation of environmental disruptions. This comprehensive attitude does not lead to any confusion: social costs are extremely heterogeneous waste phenomena associated with the ordinary functioning of the institution of business enterprise and with its distinctive mode of allocation of factor inputs in a defined context of customary arrangements and public regulations and controls. Social costs can be adequately studied and reasonably mitigated only on the basis of comprehensive, rigorous and systematic scientific investigations of the wide range of possible harmful consequences of the competitive process.

In his treatment of the health costs of business enterprise, Kapp is well aware that while in some cases it is possible to define, with some degree of scientific precision, the relationships between environmental quality and public health, in other circumstances cumulative interdependencies are more tortuous, entangled and less obvious (Kapp 1970: 836–838). Consider, for instance, the social costs of industrial location and of the urban congestion it induces. Under the discipline of the competitive process, business enterprises invest in new plants regardless of their additional costs in terms of air, water and soil pollution, and of their adverse effects on health and on public and private property values. For this reason, 'mislocation and overconcentration in urban centers' tend to be self-reinforcing and to give origin to

> such social costs as air and water pollution, as well as unavoidable expenditures for large scale sanitation, sewage disposal, hospitals, education, transportation, police, and other community services, without which social existence in urban centres is bound to give rise to social disorganization, demoralization and dehumanization of life in a large scale.
>
> (Kapp 1963: 262-263)

Kapp assigned a central role, in his theory of social costs, to the peculiarly weak position of the human factor of production in the allocation and investment choices of competing business units. Occupational accidents and pathologies are two major sources of avoidable morbidity and mortality, at the origin of severe equity and income distribution problems. This is, in fact, a classical line of analysis in the study of social costs by other eminent institutionalists such as Veblen, Commons, Clark, Perroux and Polanyi[7].

Kapp included among the social costs of the impairment of the human factor of production the social costs of unemployment and those induced by industrial accidents and occupational diseases. To him they were both concrete expressions of the basic difference existing in business between privately owned factors of production and the 'free' human durable agent of production: while both are valuable factors of production, in the absence of a comprehensive and effective system of protective labour legislation and social insurance (as, to some degree, always tends to be the case in the real world), the individual worker and his family, or the larger community, will have to bear the burden of severe (monetary and real) losses: 'These losses will be borne either by the injured worker and his family, or by the taxpayer in the form of greater public expenditures for medical care, hospitals and relief' (Kapp 1963: 158).

Quoting Clark's lesson on overhead costs, Kapp explains that

> (t)he analysis of the social costs of unemployment is an elaboration of the dicussion of the social costs of production by occupational diseases and industrial accidents. (...) Just as the human factor of production is subject to deterioration and impairment in the course of the productive process, so does its 'production' and 'upkeep' entail certain fixed costs. Not only are monetary outlays involved in raising the new generation of workers, but, young persons have to be educated and trained. In addition to this fixed investment of time and money involved in training the younger generation of workers there are costs of maintaining the worker himself in 'proper working conditions'. These costs are relatively fixed; indeed, a minimum of these costs must be borne regardless of 'whether the laborer works or not: that is, if it is not borne, if the maintenance is not forthcoming, the community suffers a loss through deterioration of its working power.'
>
> (Kapp 1963: 195; 196)

Technological change and cyclical instability are at the origin of the social costs of unemployment, which

> (i)nclude certain less tangible consequences such as the deterioration of the state of public health, higher mortality, greater incidence of crime, increased alcoholism and lower marriage and birth rates. They are further reflected in the psychological effects of general insecurity and frustration of the hopes of millions of individuals.
>
> (Kapp 1963: 202)

In this passage, Kapp anticipates more recent evidence on the association between the growth of economic inequality and unemployment and the exacerbation of health disparities and the parallel increase in crime rates and anti-social behaviors (Wilkinson 1996, 1999). Unemployment, like occupational accidents and pathologies, determines a secondary redistribution of real income within society. In both cases, overhead costs are shifted either to the individual worker and his family or to the community in terms of increased welfare expenditures. Investment and location decisions by business can induce similar effects at regional level when, with their shift of important industries from one place to another, business concerns throw considerable numbers of workers into forced idleness in the older section, with the accompanying evils of reduced public revenues and emerging local backwardness. The 'psychosomatic effects and diseases' of unemployment and the 'pervasive feeling of fear and general anxiety with all the psycho-cultural consequences which such a situation entails' (Kapp 1975: 65) are types of social costs which can induce dramatic societal consequences, especially in times of economic stagnation, including the diffusion of irrational behaviour models and authoritarian regimes. Among the social costs of unemployment, Kapp also mentions 'psychosomatic effects and diseases of various kinds which are not confined to the period of unemployment (...)' (Kapp 1975: 65) and quotes Chombart de Lauwe's 1973 epidemiological study on 367 qualified French workers to confirm his initial hypothesis on the long-term adverse health effects of recurrent episodes of unemployment on the individual worker (Kapp 1975: 65–66).

Both high and low employment, under capitalist institutional arrangements, entail social costs and are of decisive importance in determining the specific constellations of social costs of industrial accidents and diseases. Kapp's contextual analysis considers this relationship fundamental for an evolutionary analysis of the social waste induced by the weak and exploitable position of the human factor of production in private business allocation decisions. Present investigations, often sponsored by governmental agencies, make use of a series of empirical tests which Kapp adopted in order to analyse the social costs of the impairment of the human factor of production. They help to calculate the monetary losses induced by inappropriate coverage and elusion practices of workmen's compensation acts and by the inadequacy of conventional protective legislation for workers, their families or the general public (the taxpayer). They analyse the constellation of groups and forces that, in conditions of strong information and power asymmetries, reinforce trends toward the multiplication of dangerous jobs. They help to formulate concrete policy targets to correct these dramatic developments.

More generally, Kapp's view of social costs emphasizes that social costs ordinarily imply relevant distributional effects. Substantial social losses, measured in terms of objective degradation of the social and physical environment and of the consequential impairments of human health, fall differentially upon existing socio-economic groups. Their redistributive effects are widespread and complex: constellations of social costs usually lead to disparities between different regional contexts, generations, social classes and groups. The unequal distribution of the

repercussions of environmental disruption and a hypothesis of cumulative vulnerability of socio-economic groups positioned at the lowest levels of the social stratification system are crucial to Kapp's approach to social costs

> Just as the industrial worker carried the brunt of the social costs of the Industrial Revolution in the form of low wages, long hours of work, high accident rates, occupational diseases and unhealthy living conditions, the social costs of the current disruption of the social and physical environment are also borne unequally. It is true, middle and high income groups are not exempted from the effects of air and water pollution, noise, polluted city centres and chaotic traffic conditions. However, these income groups are able to avoid the full impact of the disruption of the environment by moving to less polluted suburbs and by spending their holidays in less affected areas, while poorer income groups are forced to live and work in polluted city centres. This applies particularly to black people (in the U.S.) and foreign laborers in Europe who, due to inadequate education and training are severely handicapped in their job opportunities and hence in their social mobility. These income groups are exposed to noise of factories and cities as well as more pollution and less adequate opportunities for recreation; in short, they carry the brunt of the disruption of the environment to a greater extent than middle and higher income groups.
>
> (Kapp 1974: 140-141)

These equity problems occur in situations ordinarily characterized by pervasive asymmetries in access to information and coercion resources between different private and public interests and socio-economic groups. They tend to give origin to intricate goal and interest conflicts, which usually take the form of insidious tradeoffs between different social goals and priorities.

Kapp was convinced of the urgent necessity of introducing a new type of social control of business: the correction of conventional private and public accounting procedures, the politico-technical definition of social welfare minima of various kinds, and the public promotion of alternative technologies were the concrete policy prescriptions upon which constructive solutions to problems of social disruptions could be reasonably and democratically built (Kapp 1976).

Concluding remarks

According to Kapp, the major socio-economic factor affecting public health in modern economies is the allocation of economically relevant resources by the private business sector. This perspective tends to be disregarded by socio-economic epidemiology and occupational safety and health economics, albeit in different modes and degrees. The hybrid nature of the two disciplines discussed here restricts the search for measures to contrast avoidable health losses for workers and communities. Typically, the strategic role of public research on alternative technologies and of incisively reformed, more informed and participatory, public decision-making processes (Kapp 1976) go unrecognized or are considered

elements of minor importance. Kapp's open systems analysis of the health and human costs of capitalism, formulated in terms of interdependent and circular cumulative causation, may help to enrich these disciplines by providing theoretical insights into how economic and political institutions influence health and into interest conflicts and power dynamics which are directly or indirectly relevant for health matters. It may therefore provide a framework to identify socio-political measures and reforms to correct distortions and to introduce a higher degree of substantive rationality by decisively influencing private investment and allocation decisions.

Kapp's approach may also establish a link between these two strands of hybrid social research. Kapp's analysis of market institutions and processes, of their inherent social inefficiency, and his reformist views on how to correct and redirect the economies we live in could decisively contribute to the discussion on the institutional factors and dynamics at the origin of health disparities in industrialized nations on a common scientific ground. Kapp's perspective is comprehensive enough to transcend an exclusive focus on health inequalities. It focuses on the adverse effects on human health of the various lines of social waste associated to the ordinary functioning of business enterprise and of the market mechanism. It also identifies a broader range of social losses associated with the heterogeneous and ubiquitous nature of social costs. This comprehensive attitude and the deriving systematic and interdisciplinary approach to the social bads of capitalism are fundamental for an adequate analysis of health problems, of the institutional armistices that tend to originate them and of the rigidities that may emerge as a result of the institutional maladjustments between forces and powers with diverging views on social priorities.

For corrective policy prescriptions to be effective under these circumstances, they have to be built on a detailed reconstruction of the many institutional and ecological factors that make social costs a permanent problem.

Notes

* For his incisive and constructive criticism, which allowed me to present my arguments better, I am indebted to Paolo Ramazzotti. I remain responsible for any shortcomings that remain.

1 Social or socio-economic epidemiology and occupational safety and health economics (in its institutional variant) can be defined as 'hybrid fields of investigation' (Kapp 1961: 5) in that they are the result of the existence of intervening factors and elements which cannot be studied within the partial perspectives of conventional scientific disciplines: 'What happens to the truly significant problems and disturbaces which cannot be made intelligible within the narrow confines of any of the specialized disciplines because both their causes and their repercussions transcend any compartmentalization? The inevitable outcome seems to be either that no attention is paid to these problems or that new disciplines are established to explore these issues which otherwise would be lost in the no-man's land separating the older disciplines. Hand in hand with the emergence and multiplication of new hybrid disciplines goes the development of new conceptualizations and new specialized terminologies which compound the already existing difficulties of communication between the disciplines' (Kapp 1961: 9).

2 The problem of health inequalities has indeed been the object of different antagonist approaches: those insisting upon social selective mobility (a bad economic position is

the inevitable consequence of an individual's poor health); genetic selection; statistical artifacts; behavioural risk factors and individual lifestyles; medical care and the role of medicine; objective conditions bound to the socio-economic environment – the so-called 'materialist-structuralist' approach (Abel-Smith 1990; Davey Smith, *et al.* 1994, Wilkinson 1996: 53-71; Elkeles and Mielck 1997). After a careful discussion of these different explanations in the prevailing socio-epidemiological literature, Wilkinson concludes as follows: 'We have now seen that health inequalities cannot be understood in terms of biased measurements, selective social mobility, genetic differences, inequalities in medical care or health related behaviour. Nor should it be thought that each contributes a small part which when added together explains a large part of the picture' (Wilkinson 1996: 69-70).

For the scope of the present work we need only mention that the most potent approach, the materialist-structuralist hypothesis for the explanation of health disparities, tends to consider such factors as living and working conditions and income distribution as the principal determinants of health. A recent effort to articulate these broad causal elements has led to the selection of the following ten items, resting on a large mass of empirical evidence: social organization, social stress and health; early life; the life course and the social gradient; unemployment; the psychosocial environment at work; transport; social support and social cohesion; food; poverty, social exclusion and minorities; cigarette smoking (Marmot and Wilkinson 1999).

3 For a definition of the articulation of the debate in this hybrid discipline, see Dorman 1996.

4 On this specific aspect see also Benzeval's meta-analysis of the epidemiological evidence collected in the Great Britain and referred to the health performance of the great capitalist restoration, actively promoted during the Thatcher-Major neoliberal governments (Benzeval 1997).

5 We shall not discuss in detail the risks associated with working in small/medium businesses and by discriminated groups of workers. This omission is justified by the fact that they do not represent a new frontier in occupational health studies: they simply tend to confirm, in evolving forms and modes, conditions which have for long been recognized as particularly insidious by sectoral studies. Kapp considers the social costs of unproper working conditions in small businesses as a major argument to show that also tendentially atomistic competition is inherently socially destructive (Kapp 1963).

6 For a detailed analysis of the repercussions of growing contractual flexibility on the institution of worker's compensation systems see Quinlan and Mayhew 1999.

7 John Maurice Clark's economics of social overheads (for a general appraisal see Shute 1997) was of fundamental importance to Kapp's definition of the strategy of inquiry of the avoidable and enduring impairment of the human resource in the process of allocation of factor inputs by business enterprise. Kapp explicitly recognizes this intellectual debt (Kapp 1971, xxvii-xxviii). See Stabile (1993) for a comparison between Clark's and Kapp's approaches to the social costs of labour.

References

Abel-Smith, B. (1990) 'Soziale Ungleichheit und Gesundheit: Die britische Diskussion', *Zeitschrift für Sozialreform* 36: 169-78.

Adler, N. E. and Ostrove, J. M. (1999) 'Socioeconomic Status and Health: What We Know and What We Don't', *Annals of the New York Academy of Sciences*, in N. E. Adler, M. Marmot, B. S. McEwen and J. Stewart (eds) *Socioeconomic Status and Health in Industrial Nations,* The New York Academy of Sciences, New York.

Benzeval, M. (1997) 'Health', in A. Walker and C. Walker (eds) (1997) *Britain Divided. The Growth of Social Exclusion in the 1980s and 1990s*, London: CPAG Ltd, London.

Castleman, B. (2001) 'Controversies at International Organizations over Asbestos Industry Influence', *International Journal of Health Services* 31: 193-202.

Borgers, D. and Abholz, H.H. (2001) 'Welches Kapital ist gut für die Gesundheit? Entfremdung und materielle Ressourcen als Determinanten von Gesundheit', in A. Mielck and K. Bloomfield (eds) *Sozial-Epidemiologie*, Weinheim und München: Juventa.

Coburn, D. (2000) 'Income Inequality, Social Cohesion and the Health Status of Populations: the Role of Neoliberalism', *Social Science & Medicine* 51: 135-46.

Crouch, C. (1983) 'Market Failure: Fred Hirsch and the Case for Social Democracy', in K. Kumar and A. Ellis (eds) *Dilemmas of Liberal Democracies: Studies in Fred Hirsch's Social Limits to Growth*, London: Tavistock Publications.

Davey Smith, G., Blane, D. and Bartley, M. (1994) 'Soziale Klasse und Mortaltätsunterschiede', in A. Mielck (ed.) *Krankheit und Soziale Ungleichheit*, Opladen: Leske + Budrich.

Dorman, P. (1996) *Markets and Mortality*, Cambridge: Cambridge University Press.

Dorman, P. (2000) *The Economics of Safety, Health and Well-Being at Work: an Overview*, in Focus Programme on Safe Work, ILO, The Evergreen State College, 1-42. Online. Available <http://www.ilo.org/public/english/protection/safework/papers/ecoanal/ecoview.htm> (accessed 10 July 2003).

Elkeles, T. and Mielck, A. (1997) 'Entwicklung eines Modells zur Erklärung gesundheitlicher Ungleichheit', *Gesundheitswesen* 59: 137-43.

Hirsch, F. (1976) *Social Limits to Growth*, Cambridge, Mass.: Harvard University Press.

Kapp, K. W. (1961) *Toward a Science of Man in Society*, The Hague: Martinus Nijhoff.

——. (1963) *Social Costs of Business Enterprise*, Bombay: Asia Publishing House.

——. (1970) 'Environmental Disruption and Social Costs: A Challenge to Economics', *Kyklos* 4: 833-48.

——. (1971) 'Preface', in K. W. Kapp (ed.) *The Social Costs of Private Enterprise*, New York: Schocken Books.

——. (1974) 'The Disruption and Protection of the Environment: Economics and Politics', in K. W. Kapp *Environmental Policy and Development Planning in Contemporary China and Other Essays*, Paris: Maison des Sciences de l'Homme.

——. (1975) 'Socio-Economic Effects of Low and High Employment', *Annals of the American Academy of Political and Economic Sciences* 418: 60-71.

——. (Bauman H. and Wachtl P.) (1976) *Staatliche Förderung 'umweltfreundlicher' Technologien*, Schriften der Kommission für wirtschaftlichen und sozialen Wandel, Göttingen: Otto Schwarz & Co..

Ludwig, E. R., Madeksho L. and Egilman, D. (2001) 'Letter to the Editor', *American Journal of Industrial Medicine* 39: 524-27.

Lynch, J. (2000) 'Income inequality and health: expanding the debate', *Social Science & Medicine* 51: 1001-5.

Marmot, M. and Wilkinson, R. G. (eds) (1999) *Social Determinants of Health*, Oxford: Oxford University Press.

Musgrave, R. A. (2000) 'Social Science, Ethics, and the Role of the Public Sector', in R. A. Musgrave (ed.) *Public Finance in a Democratic Society*, Northampton: Edward Elgar.

Offer, A. (2000) 'Economic Welfare Measurements and Human Well-Being', *Discussion Papers in Economic and Social History* 34: 1-35.

Quinlan, M. and Mayhew, C. (1999) 'Precarious Employment and Workers' Compensation', *International Journal of Law and Psychiatry* 22.

Quinlan, M. (2003) 'Flexible Work and Organizational Arrangements – Regulatory Problems and Responses', *working paper 16*, National Research Centre for OHS Regulation: The Australian National University.

Sen, A. K. (1983) 'The Profit Motive', *Lloyds Bank Review* 147: 1-20.

Stabile, D. R. (1993) 'Accountants and the Price System', *Journal of Economic Issues* 171-88.

Shute, L. (1997) *John Maurice Clark. A Social Economics for the Twenty-First Century*, London, New York: Macmillan Press Ltd & St. Martin's Press.

Starr, P. (1987) 'The Meaning of Privatization', *Yale Law and Policy Review* 6: 6-41.

——. (1988) *The Limits of Privatization*, Washington, D.C.: Economic Policy Institute.

Wilkinson, R. G. (1996) *Unhealthy Societies. The Afflictions of Inequality*, London: Routledge.

——. (1999) 'Putting the Picture Together: Prosperity, Redistribution and Welfare', in M. Marmot, R. G. Wilkinson (eds) (1999), Social Determinants of Health, Oxford: Oxford University Press..

——. (2000) 'Deeper than "Neoliberalism": A Reply to David Coburn', *Social Science & Medicine* 51: 997-1000.

9 Impact of economic and labour market policy on health: health costs of the 'transition process' in Central and Eastern Europe

Richard Peter and
Johannes Siegrist

Introduction

When K. W. Kapp's landmark book on *The Social Costs of Private Enterprise* was published (Kapp 1950), post-war economic growth was in its take-off stage in Western countries. Under these circumstances, limited attention was paid to his message. Neither ecological nor social and psychological costs of an expanding market driven by private enterprises were discussed, and the axiomatic assumption of economic theories that people act according to the norm of 'homo oeconomicus' remained unchallenged. Within the past five decades far-reaching changes took place at the level of economic and political development, and these changes had, and continue to have, a major impact on economic and labour market policy. Among these changes the following ones are particularly noteworthy:

1) The transformation of labour markets in western societies from full employment into labour shortage, in combination with a shift from employment in industry to the service sector.
2) The rapid technological change, particularly the substantial progress in information technology and automation and, as a consequence, the advent of flexible work.
3) The breakdown of state economies in the communist countries and the accelerated process of economic globalization.

These major changes produce both benefits and costs. At every stage, there are winners and losers of this process. Probably the most important costs associated with the changes mentioned relate to a substantial rise of unemployment and precarious work in economically developed countries.

In this contribution, health costs of precarious work are analysed with particular reference to the third type of change mentioned above, the transformation of state economies into market economies in post-communist societies in Central and

Eastern Europe. We first define precarious work and describe scientific approaches towards analysing its direct and indirect effects on human health. We then discuss the potential contribution of health costs of precarious work to the gap in life expectancy which evolved between Western and Central/Eastern European countries in recent past. In this frame, we present and discuss selected results from socio-epidemiological studies on work and health that were conducted in different working populations, both in Eastern and Western Europe. Finally, some policy implications are derived from the current state of evidence.

Precarious work and health: two theoretical models

It should be stated, first, that work is important for human health and well-being. Beneficial effects on health and well-being can be attributed to a high quality of work that meets basic human needs. On the other hand, lack of need fulfilment induced by poor quality of work may trigger ill health and disease. Need fulfilment is not restricted to the protection against chemical and physical hazards or risk of injury and accident, but includes psychological and social motivations of the working person, such as experiencing self-efficacy, self-esteem, skill utilization and personal growth, or experiencing security and satisfaction as well as receiving financial rewards, social approval and social support. The centrality of work and occupation is further underlined by the fact that having a job is normally a prerequisite for a regular income, and level of income determines a wide range of life chances. Moreover, occupation is an important criterion of the social stratification of a society, thus defining a person's social status and associated social identity. The importance of these conditions for well-being and health becomes evident from consecutive investigations of people who lose their job and become unemployed (see e.g. Bartley *et al.* 1999). Yet, health-adverse effects are not confined to job loss and unemployment, but are equally observed in precarious work. Precarious work is characterized by a low level of job security, low control of working conditions including health and safety standards, and poor earnings (Rodgers and Rodgers 1989). People who are forced into precarious jobs have little decision latitude over their work tasks, their work hours and the location of their work. Their employment contracts are poorly defined, leaving room for unfair treatment. In a globalized economy the prevalence of workers exposed to precarious jobs is increasing. This is due to the transnational activities of large private corporations that undermine or weaken national regulations of occupational life. Moreover, an increasingly competitive international labour market contributes to the growth of precarious work (Pappas *et al.* 2003).

Precarious work has also been frequent – and continues to be frequent – in the transition process from state-driven to private market economies in post-communist Central and Eastern European countries. In order to analyse the links between precarious work and health a theoretical approach is needed that provides an explanation of health effects produced by adverse working conditions. This holds particularly true for the psychosocial conditions of precarious work. Whereas chemical and physical stressors are easily identified by direct measurement of

respective stressors (such as noise, pollution, specific particles etc.), the definition of 'toxic' components of complex and highly diverse psychosocial work environments calls for a theoretical approach. The application of a theoretical model is needed to delineate particular stressful job characteristics such that they can be identified at a level of generalization that allows for their identification in a wide range of different occupations. These models are then translated into measures with the help of specific psychological and sociological research methods (questionnaires, observation techniques, etc.). While several concepts of psychosocial work-related stress have been developed (Cooper 1998; Dunham 2001; Perrewe and Ganster 2002), two models received special attention in the recent past: the demand-control model and the effort-reward imbalance model. Both models are supported by a body of evidence demonstrating their role in explaining health variations in working populations.

The first model, termed demand-control or job strain model (see fig. 1 below), was introduced by Robert Karasek and elaborated by Karasek and Theorell in a comprehensive publication (Karasek and Theorell 1990). Its basic idea can be summarized as follows: two relevant dimensions of the structure of tasks interact to produce strain reactions: high demands on the job and low control over one's tasks as indicated by a low degree of decision latitude and skill discretion. Thus, the model postulates an interaction of demand and control in the prediction of health-damaging outcomes. More recent developments of this model concern the introduction of a third dimension, social support at work (Johnson *et al.* 1989) and of measures of total lifetime exposure to job strain (Johnson *et al.* 1996).

The model of effort-reward imbalance (see fig. 2) was developed more recently in the context of medical sociological studies of chronic disease development (Siegrist 1996). It maintains that availability of a work role is associated with recurrent options of contributing and of performing, of being rewarded or esteemed, and of belonging to some significant group (e.g. work colleagues). Yet, these potentially beneficial effects are contingent on a basic prerequisite of exchange in social life, that is, reciprocity. Effort at work is spent as part of a socially organized

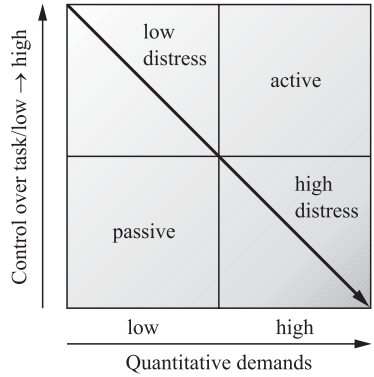

Figure 1 The demand-control model (Karasek and Theorell 1990)

exchange process to which society at large contributes in terms of rewards. Rewards are distributed by three transmitter systems as scarce resources: money, esteem and career opportunities. The model of effort-reward imbalance claims that lack of reciprocity between the costs and gains (i.e. high-cost/low-gain conditions) elicits negative emotions with special propensity to sustained autonomic and neuro-endocrine activation.

In structural terms, this imbalance results from the fact that the social exchange between employee and employer is based on an incomplete contract. An incomplete contract does not specify the full range of detailed obligations and benefits (Fehr and Gächter 2000). In incomplete contracts, assumptions of trust in mutual commitment are made. However, under certain conditions it is likely that incomplete contracts result in high-cost/low-gain conditions in employees. The risk of non-reciprocity in exchange is particularly high if employees have no alternative choice in the labour market, if their skills are poor or if they subscribe to short-term contracts. Less frequently non-reciprocity at work is experienced by workers as a negative life event, as contract violation or failed contract.

Employees themselves may also contribute to high-cost/low-gain conditions at work either intentionally or unintentionally. For instance, they may accept job arrangements that are considered unfair for a certain time for strategic reasons as they tend to improve their chances for career promotion and related rewards at a later stage. This pattern is often observed in early stages of professional careers, among others. Failed success after long-lasting investment is particularly harmful to a person's self-regulation.

Finally, there are psychological reasons of a continued mismatch between efforts and rewards at work. People characterized by a motivational pattern of excessive work-related overcommitment and a high need for approval may suffer from inappropriate perceptions of demands and own coping resources more often than their less-involved colleagues (Siegrist 1996). Perceptual distortion prevents them from accurately assessing cost-gain relations. As a consequence, they underestimate the demands, and overestimate their own coping resources while not being aware of their own contribution to non-reciprocal exchange.

In summary, the effort-reward imbalance model is based on the sociological hypothesis that structured social exchange, as mediated through core social roles (the work role), is rooted in contracts of reciprocity of cost and gain. In addition to its significance for social functioning this reciprocal contractual exchange is assumed to produce beneficial effects on individual health and well-being. The model specifies the conditions under which contractual reciprocity is not maintained and, thus, may reduce individual health and well-being.

These conditions are in part structural (or extrinsic) and in part personal (or intrinsic). Structural conditions of recurrent experience of high effort and low reward at work include lack of alternative choice in the labour market, lack of mobility, low level of skills and confinement to short-term contracts. Personal conditions include strategic choices of the workers and characteristics of individual coping with the demands and rewards at work (overcommitment).

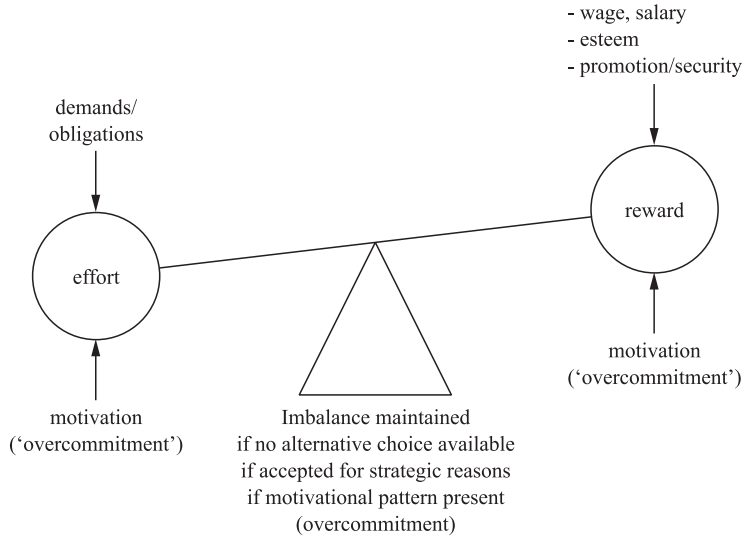

Figure 2 The model of effort-reward imbalance at work (Siegrist 1996)

The two complementary models of demand-control and effort-reward imbalance are assumed to identify specific 'toxic' components within a variety of occupations, and particularly within precarious jobs. Two main mechanisms are responsible for adverse effects on health produced by these conditions. First, a 'direct' pathway operates through continued excessive activation of the body's own stress axes, especially the hypothalamic-pituitary-adrenocortical axis and the sympatho-adrenomedullary axis. Excessive activation of these and related neurohormonal regulatory mechanisms in the long run reduces the adaptive capacities of distinct functional systems within the organism, such as the cardiovascular or the gastrointestinal system, and increases the risk of organ damage, such as atherosclerotic developments of coronary arteries and coronary heart disease (McEwen 1998; Weiner 1992).

The second, 'indirect' pathway acts through reinforcement of health-damaging behaviours, such as cigarette smoking, consumption of alcohol or drugs, and unhealthy food (Bobak and Marmot 1996; Jarvis and Wardle 1999). Under adverse material and psychosocial conditions, unhealthy behaviour may be elicited more often as a 'stress-relieving' compensatory mechanism that counteracts an emotionally distressing state of social reward deficiency (Siegrist 2002).

The question, then, arises to what extent precarious work, as assessed by measures of the two models, affects workers' health and to what extent this association contributes to an explanation of the morbidity and mortality gap between post-communist Central and Eastern European countries and Western Europe.

Morbidity and mortality differences between Central/Eastern and Western Europe

In Central/Eastern European (CEE) states, over the last thirty years, infant mortality has fallen, but, unlike in Western Europe, life expectancy has not risen. In some Eastern countries life expectancy has even declined. As a result, the gap in life expectancy between West and East has increased (see fig. 3). This is illustrated, for instance, by the fact that, in 1994, there was a six-year gap in life expectancy at birth between CEE and Western Europe (Bobak and Marmot 1996). Taking the average life expectancy of fifteen-year-old boys between the poorest Eastern and the richest Western country as another example, a difference of ten years becomes evident. It is important to note that this huge difference in life expectancy is largely due to excess mortality in midlife, and particularly among men (Weidner *et al.* 2002). More than half of this excess mortality is attributed to cardiovascular diseases, and another twenty per cent is attributed to external causes of death. However, excess mortality from these causes is unequally distributed across society. Available data indicate that there exists a social gradient in mortality: with each step one moves up on the social hierarchy, the better one's chances of survival. Conversely, the lower one's socioeconomic status the higher the risk of premature mortality, and especially of cardiovascular mortality. A social gradient of cardiovascular morbidity and mortality during midlife has been documented in Western European countries (Drever *et al.* 1996; Marmot and Wilkinson 1999; Mackenbach *et al.* 1997) as well as in CEE countries (Marmot and Bobak 2000). Further on, preliminary evidence suggests that the socioeconomic

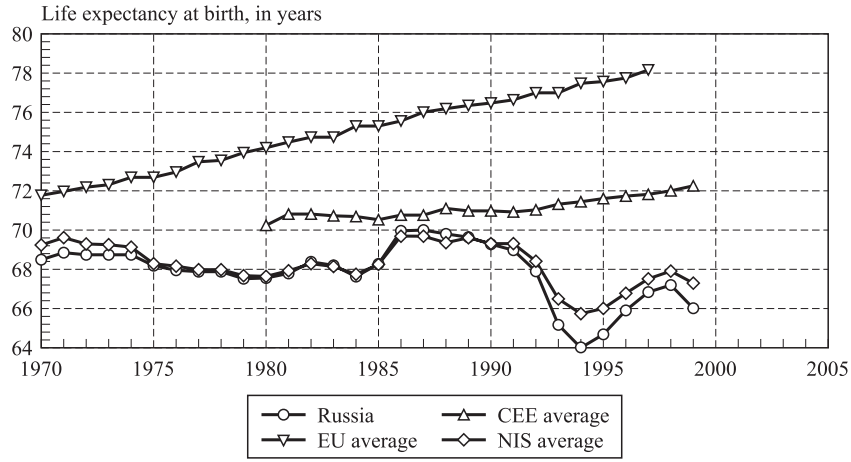

Figure 3 The gap in life expectancy between Western (EU) and Central Eastern Europe (CEE), and newly independent states (NIS) 1970-1999

differences in cardiovascular mortality and in mortality from external causes increased considerably after the breakdown of communist rule in the CEE states (Dzurova 1999; Shkolnikov *et al.* 2002). From a public health perspective it is therefore of primary importance to explain social differences in premature mortality within and across CEE countries and, by exploring the contribution of available evidence from research on social inequalities in health in Western countries, to assess the role of health-adverse working conditions in this context.

A first, most plausible explanation of excess cardiovascular mortality in midlife concerns the so-called classical risk factors of coronary heart disease (CHD), the major contributor to cardiovascular morbidity and mortality. Classical risk factors of CHD include high level of cholesterol, high blood pressure, smoking, lack of physical exercise, overweight and excessive alcohol consumption. These risk factors seem to be as important in explaining CHD in the CEE countries as is the case in Western countries (Bobak 2002; Weidner *et al.* 2002). However, perhaps with the exception of alcohol and diet (see below), these risk factors do not account for the observed socioeconomic differences in CHD morbidity and mortality within and between countries to any substantial degree (The WHO MONICA Project 1994). It is therefore necessary to explore additional risk factors that can explain the social pattern of CHD morbidity and mortality and that can be reconciled with a pathophysiological model of CHD development which accounts for the observed speed of changes in disease rates. Three such risk factors are suggested: alcohol, diet and psychosocial stress-related factors, specifically the ones associated with precarious work.

A strong temporal association between cardiovascular mortality and alcohol consumption supports the hypothesis of a causal explanation. More specifically, it was shown that a specific pattern of excessive alcohol consumption, binge drinking, is widely prevalent in CEE states and in particular among lower socioeconomic status groups (Bobak *et al.* 1999) and that binge drinking is associated with sudden cardiac death (Britton and McKee 2000). Despite this evidence it is unlikely that differences in alcohol consumption explain a major part of the observed large mortality differentials between East and West.

Unhealthy diet is a further important risk factor. It is now well-known that a diet rich in fats increases the risk of CHD whereas a high dietary intake of fresh fruits and vegetables protects against the development of CHD. A recent report covering a large number of Western and Eastern European states concludes that dietary intake of fruits and vegetables is much lower in the CEE countries, compared to Western Europe, and that lack of nutritional antioxidants may contribute to CHD morbidity and mortality differentials to a significant degree (Ginter 1996). At least one study conducted in a CEE country has documented an association of level of antioxidants in nutrition with risk of CHD: men in the lower half of the distribution of beta-carotene plasma concentration had a 3 times higher risk of heart attack, compared to men in the upper half of the distribution (Bobak 2002). It seems therefore that diet is an important factor, but again, diet alone will not be able to account for the observed variations in midlife mortality differentials within and across CEE countries.

Chronic psychosocial stress and health: empirical evidence

When discussing chronic psychosocial stress in midlife as a further potential contributor to elevated CHD risks and associated social inequalities in premature mortality one needs to keep in mind that there exists a substantial gender difference. The difference between male and female life expectancy in CEE countries is probably larger than anywhere else (Shkolnikov *et al.* 2002). More than one third of this difference is due to male excess cardiovascular mortality, and this proportion is even larger if mortality below age 65 is considered. Although an increase in cardiovascular mortality during the first years of the transition process in CEE countries was observed in both sexes, standardized mortality rates were much higher in men than in women. It is possible to further define population subgroups with a disproportional high contribution to male excess mortality. These are characterized by low socio-economic status (especially low educational level) and/or by social isolation (non-married, divorced or widowed men) (Watson 1995).

Importantly, one can conclude that threats to the acquisition of, and agency through, core social roles in early and middle adulthood define a high-risk constellation of premature cardiovascular mortality among men in CEE countries. Men who live as singles and/or who run a high risk of being unemployed or holding a precarious job, due to their low educational level, are most likely to become victims of the epidemic of cardiovascular disease morbidity and mortality. In conjunction with a loss of core social roles and associated societal rewards these men are equally deprived of protective measures from a welfare state or from supporting networks of a civic society. With the collapse of communist rule, economic shock, and social anomie emerging from rapid and profound social change these societal buffers were lost, leaving the most deprived, highly stressed population groups in isolation and, often, despair.

It can be assumed that most severe and most visible critical life events like marital separation, family breakdown, job loss and exclusion from membership in formal or informal organizations elicit particularly strong stress reactions in exposed groups. Yet, the question arises whether less dramatic, less visible conditions of everyday life, as the ones defined by precarious work, may have a similar impact on the severity and duration of stressful experience and its adverse long-term effects on health. This question has two components: first, what is the evidence of an association of precarious work, as defined by the models of job demand-control and effort-reward imbalance, with cardiovascular morbidity and mortality in general? Secondly, is there any evidence of this association in working populations in CEE countries? We answer both questions by summarizing the currently available information.

The first question is answered as follows. At present, more than 20 published reports on an association of job stress (demand-control and/or effort-reward imbalance) with cardiovascular (mostly CHD) morbidity or mortality are available, based on information from prospective or prognostic study designs (Marmot *et al.* 2002). 12 out of 17 studies show a positive association between high demand

and low control at work (or single components of this model) and CHD, five studies report no significant positive association. Concerning effort-reward imbalance, six out of seven studies document a positive association of the model (or its main components) with CHD. In both cases, overall, the relative risk of CHD is doubled among workers who are exposed to stressful work, compared to workers who are not exposed. Several studies indicate that conditions defined by the two models contribute to CHD morbidity or mortality with similar strength and independent of one another. In this regard, recent findings from a prospective study in Finland are particularly noteworthy that analysed cardiovascular mortality risks over a twenty-five year period in a group of male and female industrial employees, using the two work stress models as predicting variables (Kivimäki *et al.* 2002). After adjusting for age and sex, employees with high demand and low control (upper tertile) had a 2.2-fold cardiovascular mortality risk compared with their collegues with low or no work-related stress. Similarly, the corresponding risk ratio for employees with effort-reward imbalance (upper tertile) was 2.4. These ratios remained significant after additional adjustment for major classical cardiovascular risk factors.

In addition to prospective evidence several case-control studies were conducted with similar results. One such study showed that the risk of having experienced a myocardial infarction is critically elevated if work-related stress as defined by the two models is experienced simultaneously (Peter *et al.* 2002). However, all studies reported so far were conducted in Western countries. Therefore, the second question needs to be answered: to what extent can we infer available information from Western countries to working populations in CEE countries?

It should be stressed that few studies have been conducted so far to answer this question. Their essential findings are summarized here. Of the four investigations only one explored the effects of precarious work (in terms of the demand-control model) on cardiovascular health (Bobak *et al.* 1998). Two other studies used self-reported health as an outcome criterion (Bobak *et al.* 2000; Pikhart *et al.* 2001), and one recent report analysed associations of precarious work (in terms of the effort-reward imbalance model) on depression (Pikhart *et al.* 2004).

In a case-control study of 179 middle-aged men with a first myocardial infarction and 784 healthy controls of the same age and sex group in the Czech Republic a test of the demand-control model was conducted, using logistic regression analysis (Bobak *et al.* 1998). Results indicate that decision latitude was inversely associated with the risk of myocardial infarction, after adjustment of important confounders. However, high work demands were not associated with an elevated risk. This study therefore contributes to a limited extent to available evidence from studies in Western countries on this topic. It is important to notice that low control at work as an indicator of precarious work was associated with myocardial infarction risk in a dose-response like manner: the lower the degree of control, the higher the risk of disease (p for trend = .005). In a second study based on a larger sample the role of conditions of low control for health was explored more extensively by analysing the notion of 'perceived control' in various life domains and relating it to level of self-reported health (Bobak *et al.* 2000). Again, this study

found a strong inverse association of perceived control with level of self-reported health.

Two investigations analysed the role of precarious work in terms of the effort-reward imbalance model in explaining poor health in Central and Eastern European countries. The first one is a cross-sectional study of about 4,000 workers from Poland, the Czech Republic, Lithuania, and Hungary where strong associations between indicators of the two work stress models and poor perceived health were observed (Pikhart *et al.* 2001). The risk of suffering from poor perceived health was 1.6-fold increased for employees reporting job strain, i.e. the combination of high demands and low control (95 per cent CI 1.1–2.3). The association between effort-reward imbalance and poor self-reported health was even stronger in these populations (odds ratio 1.75 per 1 SD increase in effort-reward imbalance, 95 per cent CI 1.5–2.0). In addition, education, material deprivation and lack of control over life were associated with poor health, but did not reverse the above-mentioned effects of precarious work.

Finally, a cross-sectional study of 645 working men and 523 working women aged 45–64 years from Krakow (Poland), Karvina-Havirov (Czech Republic) and Novosibirsk (Russia) was performed to analyse associations of an adverse psychosocial work environment with depression. In this study, effort-reward imbalance at work was again strongly related to the health criterion, the depression score as measured by a validated scale (Pikhart *et al.* 2004).

Taken together, available evidence indicates that precarious work, as measured by the two models, is associated with poor health, at least in part, in all studies conducted so far in Central and Eastern European countries. It remains to see whether this may also hold true in the context of prospective epidemiological investigations.

Health policy implications

In this chapter, health costs of the 'transition' process in Central and Eastern Europe were discussed, with particular emphasis on the role of precarious work. Adverse health effects produced by precarious work were analysed with the help of two theoretical models, the demand-control and the effort-reward imbalance model. Solid evidence so far is available from epidemiological studies on cardiovascular disease. As cardiovascular disease is the main contributor to the life expectancy gap between CEE and Western Europe it is important to define and quantify new risk factors that contribute to the epidemic of this disease in CEE above and beyond the classical risk factors.

Based on current information we can identify several targets of health policy interventions that aim at reducing morbidity and mortality from cardiovascular disease and other highly prevalent health risks in middle-aged working populations in CEE. Clearly, measures that reduce income inequality are crucial because income determines such a wide range of life chances, including the availability of health-promoting food (a diet rich in fruits and vegetables). More appropriate salaries would also reduce the distressing imbalance between effortful work and

insufficient rewards. Reducing widely prevalent states of social reward deficiency by providing more appropriate salaries might be more effective in fighting against the threat of binge drinking than traditional health campaigns.

A different set of targets concerns investments into improved quality of working life. Respective measures include a better control of occupational hazards, an intensified redesigning of work task profiles (job enlargement, job enrichment, autonomous work teams) and improved leadership skills and measures. Opportunities of training, skill development and requalification need to be strengthened, and the same holds true for promotion prospects and job security. Large companies should be encouraged by adequate incentives to take the lead in a concerted action of improving the quality of work because the small- and medium-sized enterprises face serious obstacles in this regard. Professional groups and decision-making bodies with responsibility for occupational health in CEE countries are asked to learn from respective experience and from models of good practice available in Western Europe. These and other measures need to be complemented by investments into occupational health care services and preventive medicine in order to monitor risk groups and to improve their therapy.

In conclusion, the social, psychological and health-related costs of private enterprise that were emphasized by Kapp (1950) half a century ago are now visible in the 'transformation' process in CEE in a rather unique way. These costs are not an uncontrollable fate but call for policy interventions. As was argued those who are responsible for policy interventions could benefit from newly available scientific information on the social determinants of unequal health.

References

Bartley, M., Ferrie, J. and Montgomery, S. M. (1999) 'Living in a high-unemployment economy: understanding the health consequences', in M. Marmot and R. Wilkinson (eds) *Social Determinants of Health,* Oxford: Oxford University Press.

Bobak, M. (2002) 'New dietary risk factors: differences between Western and Eastern Europe', in G. Weidner, M. Kopp and M. Kristenson (eds) *Heart Disease: Environment, Stress and Gender*, Amsterdam: IOS Press.

—— and Marmot, M. (1996) 'East-west mortality divide and its potential explanations: proposed research agenda', *British Medical Journal* 312: 421–25.

——, Hertzman, C., Skodova, Z. and Marmot, M. (1998) 'Association between psychosocial factors at work and non-fatal myocardial infarction in a population-based case-control study in Czech men', *Epidemiology* 9: 43–7.

——, McKee, M., Rose, R. and Marmot, M. (1999) 'Alcohol consumption in a national sample of the Russian population', *Addiction* 94: 857–66.

——, Pikhart, H., Rose, R., Hertzman, C. and Marmot, M. (2000) 'Socioeconomic factors material inequalities, and perceived control in self-rated health: cross-sectional data from seven post-communist countries', *Social Science & Medicine* 51: 1343–350.

Britton, A. and McKee, M. (2000) 'The relation between alcohol and cardiovascular disease in Eastern Europe: explaining the paradox', *Journal of Epidemiology and Community Health* 54: 328–32.

Cooper, C.L. (1998) *Theories of Organizational Stress*, Oxford: Oxford University Press.

Drever, F., Whitehead, M. and Roden, M. (1996) 'Current patterns and trends in male mortality by social class (based on occupation)'. *Population Trends* 86: 15–20.

Dunham, J. (2001) *Stress in the Workplace: Past, Present and Future*, London: Whurr.

Dzurova, D. (1999) 'Mortality differentials in the Czech Republic during the post-1989 sociopolitical transformation', *Health and Place* 6: 351–62.

Fehr, E. and Gächter, S. (2000) 'Fairness and retaliation: the economics of reciprocity', *Journal of Economic Perspectives* 14: 159–81.

Ginter, E. (1996) 'High cardiovascular mortality in postcommunist countries: participation of oxidative stress?', *International Journal of Vitamin and Nutrition Research* 66: 183–89.

Jarvis, M. J. and Wardle, J. (1999) 'Social patterning of individual health behaviours: the case of cigarette smoking', in M. Marmot and R. Wilkinson (eds) *Social Determinants of Health*, Oxford: Oxford University Press.

Johnson, J. V., Hall, E. and Theorell, T. (1989) 'Combined effects of job strain and social isolation on cardiovascular disease morbidity and mortality in a random sample of the Swedish male working population', *Scandinavian Journal of Work, Environment and Health* 15: 271–79.

Johnson, J. V., Stewart, W., Hall, E. M., Fredlund, P. and Theorell, T. (1996) 'Long-term psychosocial work environment and cardiovascular mortality among Swedish men', *American Journal of Public Health* 86: 324–31.

Kapp, K. W. (1950) *Social Costs of Private Enterprise,* Cambridge, Mass.: Harvard University Press.

Karasek, R. A. and Theorell, T. (1990) *Healthy Work: Stress, Productivity, and the Reconstruction of working Life*, New York: Basic Books.

Kivimäki, M., Leino-Arjas, P., Luukkonen, R., Riihimäki, H., Vahtera, J. and Kirjonen, J. (2002) 'Work stress and risk of cardiovascular mortality: prospective cohort study of industrial employees', *British Medical Journal* 325: 857–61.

Mackenbach, J. P., Kunst, A. E., Cavelaars, E. J. M., Groenhof, F., Geurts, J. J. M. and the EU Working Group on Socioeconomic Inequalities in Health (1997) 'Socioeconomic inequalities in morbidity and mortality in western Europe', *Lancet* 349: 1655–59.

Marmot, M. and Bobak, M. (2000) 'International comparators and poverty and health in Europe', *British Medical Journal* 321: 1124–28.

——. Theorell, T. and Siegrist, J. (2002) 'Work and coronary heart disease', in S. A. Stansfeld, M. G. Marmot (eds) *Stress and the Heart*, London: BMJ Books.

—— and Wilkinson, R. (1999) *Social Determinants of Health*, Oxford: Oxford University Press.

McEwen, B. S. (1998) 'Protective and damaging effects of stress mediators', *New England Journal of Medicine* 338: 171–79.

Pappas, G., Hyder, A. A. and Akther, M. (2003) 'Globalization: towards a new framework for public health', *Social Theory and Health* 1: 91–107.

Perrewe, P. L., and Ganster, D. C. (2002) *Historical and Current Perspectives on Stress and Health*, Amsterdam: JAI Press.

Peter, R., Siegrist J., Hallqvist, J., Reuterwall, C. and Theorell, T. (2002) 'Psychosocial work environment and myocardial infarction: improving risk estimation by combining two complementary job stress models in the SHEEP Study', *Journal of Epidemiology and Community Health* 56: 294–300.

Pikhart, H., Bobak, M., Siegrist, J., Pajak, A., Rywik, S., Kyshegi, J., Gostautas, A., Skodova, Z. and Marmot, M. (2001) 'Psychosocial work characteristics and self rated health in four post-communist countries', *Journal of Epidemiology and Community Health* 55: 624–30.

——, Bobak, M., Pajak A., Malyutina, S., Kubinova ,R., Topor, R., Sebakova, H., Nikitin, Y. and Marmot, M. (2004) 'Psychosocial factors at work and depression in three countries of Central and Eastern Europe', *Social Science & Medicine* 58: 1475-82.

Rodgers, G. and Rodgers, J. (1989) *Precarious Jobs in Labour Market Regulation*, Geneva: International Labour Office.

Shkolnikov, V. M., Mesle, F. and Leon, D. A. (2002) 'Premature circulatory disease mortality in Russia: population and individual-level evidence', in G. Weidner, M. Kopp and M. Kristenson (eds) *Heart Disease: Environment, Stress and Gender*, Amsterdam: IOS Press.

Siegrist, J. (1996) 'Adverse health effects of high-effort/low-reward conditions', *Journal of Occupational Health Psychology* 1: 27–41.

——. (2002) 'Adverse health effects of effort-reward imbalance: Applying the model to Eastern Europe', in G. Weidner, M. Kopp and M. Kristenson (eds) *Heart Disease: Environment, Stress and Gender*, Amsterdam: IOS Press.

The World Health Organization MONICA Project (1994) 'Ecological analysis of the association between mortality and major risk factors of cardiovascular disease', *International Journal of Epidemiology* 23: 505–16.

Watson, P. (1995) 'Explaining rising mortality among men in Eastern Europe', *Social Science & Medicine* 41: 923–34.

Weidner, G., Kopp, M. and Kristenson, M. (2002) *Heart Disease: Environment, Stress and Gender*, Amsterdam: IOS Press.

Weiner, H. (1992) *Perturbing the Organism: The Biology of Stressful Experience*, Chicago: Chicago University Press.

World Health Organization (2001) *Health in Europe*, Copenhagen: WHO Regional Office.