1 Introduction

In this introductory part, we present the basic concepts and some evidence concerning social protection. In particular we define the welfare state, social protection, and social insurance. We also introduce some taxonomies of the welfare state, the most important being based on the redistributive nature and the size of the programs.

1.1 Definition

It is not easy to give a good definition of either the welfare state or social protection.[1] To paraphrase Anthony Atkinson (1991), there is indeed a lot of the proverbial elephant in the room as regards social protection and the welfare state: we may not be able to define this elephant, but we can recognize one when we see it. The usual way is to list the objectives to be pursued and the instruments used to achieve them. Accordingly, social protection is composed of a set of actions financed by the state that (i) support individuals and families in dealing with vulnerabilities during their lifecycle, (ii) help especially the poor and vulnerable groups in having the resilience to respond to crisis and shocks, including social-environmental risks, and (iii) favor social inclusion and support families. More specifically, social protection helps in coping with the various lifetime risks linked to unemployment, disability, sickness, early or late death, retirement, and family. And it strives to alleviate both temporary and permanent poverty and income inequality. The standard instruments are transfers or the provision of services such as education and housing. Both health and long-term care can be directly provided by the state or indirectly through transfers that allow one to pay for those services. Among the transfers, one can distinguish between those of social assistance, generally based on means-testing, and those of social insurance that are in great part contributory. Figure 1 provides a sketch of the main components of the welfare state and of social protection. One sees that the welfare state comprises social protection but also benefits in kind such as education and social housing. From this figure, one sees the distinction made between the welfare state and social protection.

This taxonomy can be discussed. For example, family allowances can be part of social insurance as they are in most European countries, or they can be attached to social assistance. In this Element, benefits in kind are only dealt with in Section 5.3. Furthermore, the objectives mentioned earlier can be achieved not only with budgetary means but also with a number of nonbudgetary tools. Among them are the laws; for example, the law that constrains builders to

[1] For a good overview, see Barr (2020) and Pestieau and Lefebvre (2018).

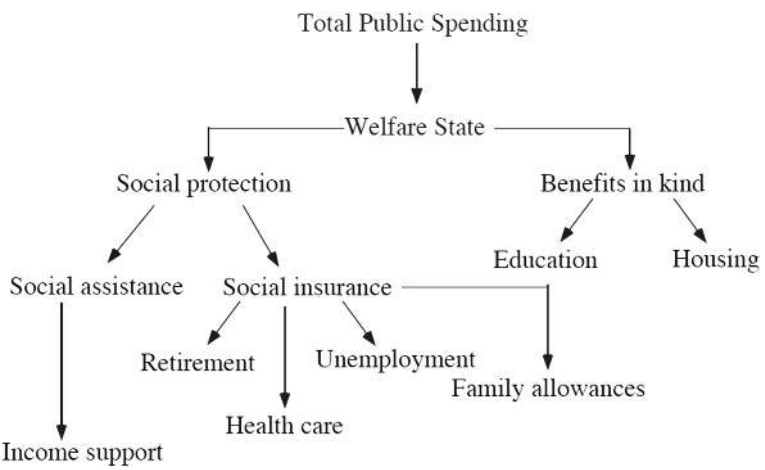

Figure 1 Welfare state and social protection

include in their constructions amenities that are friendly to handicapped people or the law that forces employers to hire a certain number of handicapped workers. There is also the legislation protecting workers against on-the-job accidents. Furthermore, there are several private social protection arrangements that, although mandatory, are not financed through funds that are part of public spending. These include mandatory private pensions and health care insurance. The importance and generosity[2] of the welfare state are usually measured by the share of social spending in GDP. This share has a rather wide range across industrialized countries, from 13.4 percent in Ireland to 31 percent in France for the year 2019. Note, however, that if we take into account mandatory private protection spending and the taxation of social benefits to obtain what is called the generosity of net social spending, this range narrows. For example, whereas the share of gross social spending in 2017 was equal to 18.4 percent in the United States versus 31.5 percent in France, in net terms, it was 31.1 percent in France and 29.7 percent in the United States.[3]

1.2 The Crisis of the Welfare State

Even though the generosity of the welfare state has been increasing almost everywhere and converging across countries, a number of books and plenty of articles and reports have been written over the last decades to indicate that the welfare state is in crisis. There is some discontent that reaches a peak with the emergence of populist movements and the various social divides that have

[2] Throughout this Element, generosity means the level of social benefits.
[3] Adema (2001), Adema et al. (2011).

emerged lately. One major critique is that the welfare state may have focused too much on poverty alleviation and inequality reduction and not enough on social mobility. Plenty of evidence shows that the social elevator does not work anymore, and this generates frustration among people who don't see any prospects for them and their children.[4] A related criticism is that the welfare state is ill-adapted to current problems. This is because it was designed several decades ago in a world where there was little mobility of factors and the ensuing social dumping, the labor market was less precarious, family solidarity was stronger, and there was more compliance. Demographic aging is also making things difficult. Given these assessments, there is a clear need to design social policy that meets those new challenges.

1.3 Types of Welfare State

There does not exist a single model of welfare state in the OECD countries. Each country has its own model that is the result of its political and social culture and of its economic evolution. There exists a number of taxonomies of welfare states which focus on specific features of their functioning. Economists tend to focus on a taxonomy based on two characteristics: the generosity and the redistributiveness of programs. The main interest of distinguishing among types of social protection programs is the different implications they have in terms of efficiency, equity, and political sustainability. To measure the redistributiveness, one can use the progressivity index I_R that for each program looks at $I_R = 1 - G(b)/G(w)$ where $G(b)$ is the Gini of benefits and $G(w)$ is the Gini of income.[5] If the program is contributory, that is, the benefits are closely linked to income or contributions, $I_R = 0$. If, on the contrary, the benefits are flat, unrelated to contributions, then $I_R = 1$. Using those two characteristics, generosity and redistributiveness, one can distinguish among three main types of welfare states:

- Those which are redistributive and generous. Typically, the Nordic countries.
- Those which are contributory and generous. They are also called Bismarckian and comprise Germany, France, and Italy.
- Those which are redistributive but not generous. These are labeled Beveridgean and are the Anglo-Saxon countries.

[4] OECD (2018).
[5] Biggs et al. (2009).

In Section 2, we come back to the distinction between Bismarckian and Beveridgean programs, named after two famous founders of social protection, Otto von Bismarck and William Beveridge.

Sociologist Esping-Andersen (1990) used the concept of "decommodification" of social protection to rank countries. Decommodification means that services are rendered and transfers are made as a matter of right, without reliance on the market. Using a number of indicators, he rates his sample of welfare states according to their decommodification score. This allows him to distinguish among three welfare state regimes: the Anglo-Saxon "new" nations are all concentrated at the bottom of his index; the Scandinavian countries are at the top; in between, we find the continental European countries, some of which, like Belgium and the Netherlands, fall close to the Nordic cluster. Besides the preceding taxonomy, three related concepts are often used to distinguish among welfare states. These are activation, responsibility, and individualization. The first line of separation is the extent to which those benefiting from social benefits are induced to get out from the state of dependence in which they are. The Danish flexicurity program is typical of such a proactive option. Another line of separation is between countries where benefits are only awarded when benefiters are deemed unlucky and not responsible for what they are. Finally, there is the distinction between programs that focus on the individual and others that are targeted to the family unit. As will be seen, those distinctions can have important implications regarding the financial and political sustainability of social protection and they impact the standard trade-off between equity and efficiency.

2 Design and Sustainability

This section is devoted to the idea that a well-designed social protection should be both financially and politically sustainable. And we introduce the key trade-off between the redistributive nature of a program and its political support.

2.1 Bismarck versus Beveridge

The issue at hand is to choose either politically or normatively a type of social protection system that is politically sustainable. We choose the example of pensions, but the issue of political support applies as well to other parts of social protection. We consider a society consisting of N types of individuals i. An individual of type i is characterized by a labor productivity or wage, w_i. There are n_i individuals of type i. Each individual lives two periods. They work during the first period and retire in the second period. They allocate their earnings, net

of a contribution to the pension system, between first-period consumption and saving. Second-period consumption is equal to the pension denoted p_i, plus the returns from saving. The pension system is here assumed to be fully funded, namely behaving like private saving. Individual i's lifetime utility is

$$U_i = u(c_i) + \beta u(d_i) = u(w_i(1-\tau) - s_i) + \beta u((1+r)s_i + p_i),$$

where

$$p_i = (1+r)\tau\left(\alpha w_i + (1-\alpha)\bar{w}\right).$$

We use the following notation:

- $u(.)$ is a strictly concave utility function;
- c and d are respectively first- and second-period consumption;
- s is saving;
- β is the time preference factor;
- τ is the payroll tax;
- w_i and r are respectively the wage level and the interest rate;
- $\overline{w} = \sum_{i=1}^{N} n_i w_i$ is the average wage level.

Finally, the parameter α is the Bismarckian parameter that can be related to the progressivity index defined earlier. Instead of using the Gini coefficient, we can use the variance of either pension benefits or earnings. We thus have $I_R = 1 - V(p)/V(w) = 1 - \alpha^2\tau^2(1+r)^2$, where $V(.)$ stands for variance. A pure Bismarckian system is such that pension contribution and saving bring the same return and are thus equivalent. A pure Beveridgean system is such that everyone gets the same pension irrespective of their contribution.

We now allow for some work in the second period. Let z be the fraction of second period devoted to work. Given that each period has a length of 1, the age of retirement is $1+z$ and the length of life 2. We assume that the individual can choose z and that earnings in the second period are taxed. Finally, the disutility of postponing retirement is $v(z) = z^2/2$, a strictly convex function. We now write the lifetime utility of individual i as

$$U_i = u(w_i(1-\tau) - s_i) + \beta u((1+r)s_i + w_i(1-\tau)z_i + p_i - z_i^2/2),$$

where

$$p_i = \tau\left\{\alpha\left[(1+r)w_i + w_i z_i\right] + (1\text{-}\alpha)\left[(1+r)\bar{w} + \overline{wz}\right]\right\}$$

and

$$\overline{wz} = \sum_i n_i w_i z_i.$$

Maximizing U_i with respect to s_i and z_i, we obtain the following FOCs (first-order conditions):

$$u'(c_i) = \beta(1+r)u'(d_i)$$

and

$$z_i = w_i[1-\tau(1-\alpha)],$$

where we implictly assume that the choice of z_i does not affect $\overline{wz}$.

If $\beta(1+r) = 1$, we have equality between the two consumption levels. When $\tau = 0$, $z_i = w_i$. We clearly see that with the Beveridgean regime, there is an incentive to retire earlier than with a contributory regime. In fact, when $\alpha = 1$, the pension system is neutral toward the retirement decision. We come back to this issue in Section 4.5.1.

2.2 Optimal Design

It is legitimate to look for the pension design that would maximize a utilitarian social welfare function such as

$$SW = \Sigma_i n_i \nu_i \left[u(c_i) + \beta u(d_i)\right],$$

where n_i is the relative number of type i's individuals and ν_i is the social weight given to type i. Those social weights make it possible to encompass the case where the objective would be the Rawlsian maximin, namely the maximization of the utility of the worst-off. One easily checks that, without liquidity constraints (saving can be negative) and with fixed age of retirement (z is a constant), the solution would be $\alpha = 0; \tau = 1$. Otherwise, the problem gets more complicated and the solution depends on key factors:

- liquidity constraints ($s_i \geq 0$)
- tax distortions
- social weights
- the wage distribution.

With plausible assumptions, the solution is $\alpha \leq 0; 1 > \tau > 0$.

Note that $\alpha < 0$ amounts to a means-testing regime. The pension level decreases with the level of wage:

$$\frac{dp_i}{dw_i} = \tau\alpha\left[1+r+z_i\right] < 0.$$

The implication of this result is that normally the optimal scheme implies means testing or at least a flat rate benefit. We illustrate this point with a simple

example of an economy where the interest rate is nil, there is no time preference and no work in the second period. The individual's utility is

$$U_i = u(c_i) + u(d_i) = u(w_i(1-\tau) - s_i) + u((s + \tau\,[\alpha w_i + (1-\alpha)\,\bar{w}]).$$

In case there is no liquidity constraint, namely saving can be negative, we have that $c_i = d_i$ and thus

$$c_i = d_i = \frac{w_i - \tau\,(1-\alpha)\,(w_i - \bar{w})}{2}.$$

It is clear that social welfare is maximized when $\tau = 1$ and $\alpha = 0$. With these parameters, we have

$$c_i = d_i = \bar{w}/2.$$

Note that if the tax rate is limited to a value below 1, the same result can be obtained by having a value of α below 0 such that $\tau\,(1-\alpha) = 1$. The problem is that such an optimal scheme is not politically sustainable. As soon as introduced, it is subject to progressive erosion. The middle class that is needed to gather a majority in favor of the pension system does not find it attractive.

A combination of Beveridge and Bismarck is to be chosen. And in fact, this solution will even be preferable for the poor.

2.3 Political Support

The preceding optimal solution might end up lacking political support according to the idea that a program exclusively targeted to the poor tends to become a poor program. Lack of political majority leads to a slow erosion of such a program. To formalize this idea, we adopt a two-stage collective choice: first, the choice of α, and second, the choice of τ (majority voting). The choice of α could be normative or positive, in which case we would have sequential voting, first for α and then for τ.

In Casamatta et al. (2000), α is chosen according to the Rawlsian maximin criterion on the basis of a relation $\tau(\alpha)$ that is obtained from majority voting on τ, given the parameter α. With well-behaved utility functions, Casamatta et al. obtain a solution with both parameters between zero and one and a positive correlation between them. Empirically, this correlation can be verified.[6]

To make the point that a program for the poor is politically unsustainable, let us use a simple model wherein each individual pays a tax on their earnings

[6] Conde-Ruiz and Profeta (2007).

to finance a uniform benefit to the poor, namely those with income below w_P. The utility of each individual is given by

$$U_i = u(w_i(1-\tau) + b_i),$$

where $b_i = \hat{b}$ for $i < P$ and $b_i = 0$ for $i \geq P$. The median income is w_m that is larger than w_P and smaller than $\overline{w} = \sum n_i w_i$.

If this benefit is chosen by majority voting, it is pretty clear that the median voter will vote against it. To support it, one should award all individuals the benefit $\hat{b} = \tau\bar{w}$. In that case the median voter will back it, given that they pay τw_m and obtain $\tau\bar{w}$.

To conclude, a more redistributive program calls for a more generous program. This has been called the paradox of redistribution (see Section 3.1).

2.4 Notional Accounts

Assume that we have a pension system with a given α. Some economists believe that instead of having a hybrid system, it would be preferable to split the system into two parts: a purely redistributive one providing a flat-rate pension and a purely Bismarckian one that would be purely contributive and mimic a private scheme. Their concern is that when making decisions within the hybrid system, individuals consider that their tax is fully distortive. In other words, their perceived α, denoted α^P, is equal to zero and thus, in the preceding example, they retire at $z = w(1-\tau)$ and not at $z = w(1-\tau(1-\alpha))$. On the contrary, some other economists, close to labor unions, are in favor of the hybrid system, convinced that the perceived α is equal to one and thus that $z = w$. In their view, the pension system is a social compact, a common good, that brings utility to everyone.

The same reasoning applies to the whole social protection. In many countries, about two-thirds of social spending concerns lifetime redistribution.[7] This includes branches such as health, education, pension, long-term care, and unemployment. For these branches, one could have individual notional accounts that would start at birth and end at death and that would be actuarially fair: present value of benefits would be equal to present value of contributions. There would be no distortion in those accounts. Besides those notional accounts, there would be a redistributive program that would unavoidably be distortionary.

There are several problems with such an approach. First, there is the previously discussed issue of political support of the redistributive scheme.

[7] See Sorensen (2003).

Second, it could be used as a first step toward privatization of the notional accounts system along with abandonment of the redistributive scheme. Third, it assumes away the possibility of myopic behavior, which is one of the main reasons for forced saving. Finally, it does not take into account a number of market failures (annuities, long-term care, unemployment insurance) that motivate public action. Let us illustrate the point concerning the perceived contributory factor.

Consider a society comprising two (types of) individuals with wage income $w_2 > w_1 = 0$. The only source of income of individual 1 is a transfer from the government expressed as $P_1 = b$. Individual 2 pays a tax τ on his earnings $w_2 l_2$, where l_2 is labor supply, and receives from the government

$$P_2 = b + \alpha\tau w_2 l_2.$$

The revenue constraint is thus

$$2b = \tau\,(1-\alpha)\,w_2 l_2.$$

With a quadratic disutility of labor, the utilities of both individuals end up to be equal to

$$U_1 = b; U_2 = w_2 l_2\,(1-\tau) - l_2^2/2 + P_2.$$

Individual 2 maximizes his utility with respect to l_2:

$$Max_{l_2}: w_2 l_2\left[1-\tau(1-\alpha^P)\right] + b - l_2^2/2.$$

This implies that $l_2 = w_2\left[1-\tau(1-\alpha^P)\right]$. Assume that $w_2 = 10$, $\tau = 0.2$, $\alpha = 1/2$. We consider three cases of (mis)perception.

1. $\alpha^P = \alpha$. Then $l_2 = 9; b = 4.5$.
2. $\alpha^P = 0$. Then $l_2 = 8; b = 4$.
3. $\alpha^P = 1$. Then $l_2 = 10; b = 5$.

Clearly for individual 1, case 3 is ideal, followed by case 1.

3 Performance of the Welfare State

One of the most widespread critiques leveled against the welfare state is its inefficiency in distributing benefits and in producing services. This section deals with the measurement of performance of the welfare state in providing social services and social transfers. Ideally, the performance of the welfare state as a whole or that of its components can be assessed by the extent to which they fulfill the objectives assigned to them. We will analyze the efficiency

of various transfer programs (social security, unemployment insurance), that of production of services (education, health, transportation), and that of the welfare state as a whole.

3.1 Transfer Programs: Administrative Costs and Redistributive Efficiency

A well-functioning transfer program, such as public pensions or health care, is expected to minimize its administrative cost and to allocate benefits to the targeted population. Private and public insurance schemes incur what is called expense loading, that is, the amount covering both administrative and maintenance costs. Compared to the private sector, administrative costs tend to be quite low in the public sector. Two factors explain such a difference: the scale that is much larger in public programs, most often covering the entire population, and the absence of expensive advertising campaigns.

Another problem with social assistance programs is the distributive inefficiency that occurs when needy households do not exercise their right to benefit from them, while other households not suffering from precarious conditions do benefit from them. The first problem is linked to the issue of take-ups caused by ignorance or fear of stigmatization. The second problem that arises when well-off social groups often benefit from social provisions intended for disadvantaged groups has been studied under the name of the "Matthew effect."[8]. Accordingly, for cultural and institutional reasons, well-off individuals outsmart disadvantaged ones to have access to various social programs. One famous example is the use of Medicaid[9] for long-term care by the American middle class through a process of strategic impoverishment. The strategy is standard: spending down to be entitled to means-tested programs.

The Matthew effect is generally considered as bad and thus should be fought against. At the same time, following Korpi and Palme (1998); Korpi and Palme (2003), we have the "paradox of redistribution" according to which it can be desirable to let the middle class benefit from social programs to ensure their political support. In other words, a program that is restricted to the poor has a smaller redistributive effect than universal systems. We dealt with this issue by showing that a social program would benefit from being partially Bismarckian according to the old saying: A program for the poor is a poor program.

So far, the focus is on the efficiency of social transfer programs and their administrative costs. In this particular activity the index of performance is

[8] Deleeck (1979), Merton (1968).
[9] Medicaid is the American public health insurance program for people with low income. It covers one in five Americans and rests on means testing.