Do standard classifications still represent European welfare typologies? Novel evidence from studies on health and social care

Giovanni Bertin\textsuperscript{a}, Ludovico Carrino\textsuperscript{a,b}, Marta Pantalone\textsuperscript{a,*}

\textsuperscript{a} Ca’ Foscari, University of Venice, Department of Economics, Fondamenta S. Giobbe - Cannaregio 873, 30121, Venezia, Italy
\textsuperscript{b} King’s College London, Department of Global Health & Social Medicine, Aikhyech 40, London, WC2R 2LS, UK

\textbf{ARTICLE INFO}

Keywords: Classification Meta-analysis Health policy Social policy Public policy Welfare regime Public economics

\textbf{ABSTRACT}

Due to the profound changes that have characterised welfare systems, the representativeness of standard welfare classifications such as Esping-Andersen’s Three Worlds of Welfare (TWW) have been questioned. In response to concerns that welfare services do not share a common rationale across policy areas, new typologies focused on sub-areas of welfare provision have been introduced. Still, there is little evidence on whether such policy-specific typologies are (i) consistent with the standard TWW classifications; and (ii) consistent across policy areas.

We reviewed 22 recent studies which identified welfare typologies in 12 European countries focusing on economically relevant areas such as healthcare and social care. We build novel indices of “welfare similarity” to measure the extent to which welfare systems have been grouped together in previous studies. Our findings are twofold: first, healthcare and social care policies are characterised by the coexistence and overlap of multiple regimes, i.e., a hybridisation of the original TWW taxonomy. Second, countries classifications are substantially different between healthcare and social care, which highlights the lack of coherence in welfare systems rationales across policy areas. Our findings suggest that comparative analyses of welfare systems should narrow their focus on policy-specific areas, which may prove more informative than general classifications of welfare states.

\textbf{1. Introduction}

The beginning of this century has been characterised by the redefinition of welfare systems in all European countries, involving a phase of stagnation and review that, at least in some cases, has undermined its original logic (Arts and Gelissen, 2002; Bonoli and Natali, 2012; Ellison and Fenger, 2013; Häusermann, 2010; Hemerijck, 2012). Most comparative studies on welfare systems employ, as a benchmark, the Esping-Andersen (EA) Three Worlds of Welfare classification, which was originally conceived in the phase of maximum expansion of welfare policies (Esping-Andersen, 1990, 1999). The relevance of EA’s classification has been questioned due to the profound changes that have characterised, and still characterise, welfare systems. However, there is a limited understanding of how welfare regimes evolve and adapt over time and how particular regimes can be extended from the original core nations to other countries originally belonging to different groups in EA’s typology (Powell and Barrientos, 2015), resulting in a partial hybridisation of the original EA regimes. This paper provides novel evidence on whether the transformations observed in European welfare states might be considered as an evolution within the boundaries of the original EA classification or whether they have led to a reduction in the internal homogeneity and consistency of the EA classification.

The recent literature points to a lack of consensus on how welfare states in OECD countries have been classified during the 2000s. Two recent studies performed a meta-analyses of existing welfare state classifications, and came to opposite conclusions: Powell et al. (2020)’s review resulted in a mixed picture of hybrid EA regimes; conversely, Buhr and Stoy (2015) confirmed the validity of EA’s classification. Similarly, Ferragina and Seeleib-Kaiser (2011) classified countries based on a wide set of welfare-provision indicators, providing evidence in favour of EA’s three worlds of welfare capitalism. This conflicting evidence should be read in light of the ongoing debate on the theoretical and empirical foundations of the analysis of welfare regimes. Some authors have highlighted that the lack of a clear definition for crucial concepts, e.g., “regime” and “commodification”, has led to inconsistent operationalisations across different analyses (Bambra, 2006; Castles and Mitchell, 1993; Powell, 2015; Rice, 2013). Other scholars have highlighted how welfare regimes classifications often lack consistency in the choice of statistical indicators and methods (Barrientos, 2015; Powell and Barrientos, 2015; Yorilik et al., 2019).

\textsuperscript{*} Corresponding author.
E-mail addresses: giovanni.bertin@unive.it (G. Bertin), ludovico.carrino@kcl.ac.uk (L. Carrino), marta.pantalone@unive.it (M. Pantalone).

https://doi.org/10.1016/j.socscimed.2021.114086
Received in revised form 29 April 2021; Accepted 26 May 2021
Available online 29 May 2021
0277-9536/© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
Most meta-analyses of welfare classifications typically include multiple areas of social policy (i.e., the “welfare-as-a-whole”). For example, Powell et al. (2020) reviewed studies that focused on the welfare state as-a-whole rather than on specific services. Ferragina and Seeleib-Kaiser (2011) included papers which focused on cash transfers and a mix of indicators related to the concepts of “decommodification”, “social stratification” and “defamilisation”. Similarly, Buhr and Stoy (2015) jointly reviewed studies on social care, healthcare and education policies. By considering several policy areas at once, the aforementioned studies implicitly assume that, within the same country, the provision of welfare services across different policy areas share a homogeneous and coherent rationale. Alternatively stated, each welfare regime is assumed to reflect a set of values coherently realised in each policy area. However, recent studies have suggested that such assumption is unlikely to hold.

In an influential work, Kasza (2002) highlighted that welfare states exhibit significant inconsistencies across different areas of intervention within and between countries (incoherence hypothesis). As the welfare-as-a-whole taxonomy ignores such variation, Kasza deemed it unable to capture the complex motives that inform each country’s welfare programs and argued in favour of policy-specific typologies. Other scholars have suggested that the welfare-as-a-whole classification is not well suited for studying specific welfare areas such as healthcare, as it lacks focus on social and healthcare services (Bambra, 2005; Wendt, 2009). Similarly, comparative studies have shown incoherence across policy classifications due to different time frames and cultural orientations (Saraceno and Keck, 2010). For example, while healthcare policies in United Kingdom and Italy have been sometimes linked to social democratic regimes (as they are universalistic), their social care and pension policies reflect cultural models akin to the liberal (UK) and corporatist (Italy) typologies (Berrin and Carradore, 2015).

Driven by concerns on the limitations of the welfare-as-a-whole classifications, a large body of literature has identified policy-specific welfare typologies through comparative analyses of, for example, healthcare (Bambra, 2005; Böhm et al., 2013; Jensen, 2008; Jomuad et al., 2010; Reibling, 2010; Wendt, 2009, 2014) and social care policies (Bambra, 2004, 2007; Bettio and Plantenga, 2004; Boje and Eijnes, 2012; Chau and Sam, 2013; Cho, 2014; Jensen, 2008; Kautto, 2002; Kraus et al., 2016; Leiner, 2003; Thévenon, 2011; Verbeek-Oudij et al., 2014). However, to the best of our knowledge, no study has yet analysed: (i) the extent to which the cluster structures resulting from policy-specific studies overlap with the standard EA classification (hybridisation hypothesis); and (ii) whether the typologies emerging from studies focusing on different policy-areas are consistent or not (incoherence hypothesis).

This paper aims to fill this gap by performing a meta-analysis of studies that produced classifications of, respectively, healthcare and social care systems in 12 European countries in the first decade of the 2000s. Both policy areas have substantial economic relevance: in OECD countries in 2005–2017, the public expenditure for healthcare and social care (defined as family support, especially for children and older people, see Jensen, 2008), averaged 5.7% and 2.3% of GDP, respectively (OECD, 2019). We build an index of “welfare similarity” to capture, separately for the social care and healthcare sectors, the extent to which welfare systems have been grouped together in the reviewed papers. In order to build a robust index of similarity, our analysis focuses on countries which have been extensively included in welfare classification studies. Most of the existing literature overwhelmingly focused on mature welfare regimes in Europe, our main country selection includes Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, the Netherlands, Spain, Sweden and United Kingdom. We improve upon previous meta-analyses (Buhr and Stoy, 2015; Powell et al., 2020) in that we introduce a novel methodological approach that accounts for the variation in the number of times a country is analysed in the reviewed literature. We show that failing to account for this variation may lead to biased results.

Our findings are twofold: first, we highlight the coexistence and overlap of multiple regimes in both healthcare and social care policies, which results in a hybridisation of the original EA classification. Second, we find that countries classifications are substantially different between healthcare and social care policies, which provides evidence for the lack of coherence of welfare provision rationales across policy areas.

Our results are relevant for both the academic and policy debate, as they suggest that classifications of welfare systems should narrow their focus on specific policy areas, which are not necessarily in line with standard classifications. Moreover, comparative analysis based on policy-specific welfare typologies may prove more informative to policymakers than welfare-as-a-whole classifications.

The study is structured as follows. Section 2 reviews the existing debates on how transformations in welfare systems can affect welfare classifications. Section 3 outlines the data and methods used to generate the welfare similarity index. Section 4 presents the results of the welfare similarity analysis, while Section 5 concludes by discussing our findings in light of the existing debates.

2. Classifications of evolving welfare systems: background and hypotheses

A large stream of recent literature has discussed the complexity and multidimensionality of the transition process within welfare regimes (Häusermann, 2012; Jensen, 2011). Following Jensen (2011), the development of welfare states can be explained according to three different perspectives.

The first, the ideological perspective, suggests that welfare reforms are the outcome of an interaction and bargaining process between competing welfare-state ideologies (conservative, democratic or liberal). The complexity of such interaction makes it often hard to place such policy changes within the classic EA taxonomy. Moreover, the context within which welfare reforms take place is in constant evolution and characterised by unstable equilibria and non-linear developments, reflecting the instability of recent political processes, including the dissolution of massive ideological blocks. This has led to the overlap and merging of different welfare perspectives and rationales, as in the case of liberal neo-welfarism, which results from an overlap between laissez-faire theories and the theoretical grounds at the base of welfare systems (Ferrera, 2013). Moreover, several authors have advocated the need to move beyond the debate on stereotypical political ideologies regarding social protection (e.g., the left as the advocate of social protection and the right as the driver of social spending cuts; see Häusermann, 2012), in favour of a differentiation of policies that aim (or not) to shift public intervention from old to new social risks.

Second, the neo-institutional perspective links welfare-state changes to the ability of institutional factors (e.g., vested interests and veto powers) to facilitate or delay the expansion of individual policy areas. By underlining the role of institutions in preserving the status quo (Bonoli, 2001), the neo-institutional approach to public policies highlights how institutions often seek their own self-preservation. Such institutions may attempt to influence the development of the system on the basis of their history and well-established features, thus potentially becoming a strong factor of resistance to change. These processes are associated with the dynamics of political consensus and policy development. For example, Weaver (1986) argued that politicians are more focused on blame avoidance (i.e., avoiding criticisms for unpopular choices) than on credit claiming (i.e., being praised for taking popular actions). These dynamics end up consolidating both the resistance to change and the political and cultural matrix of welfare systems.

According to the third, neo-functionalist, perspective, the changes in welfare regimes are related to the evolution (and adaptation ability) of economic systems. Jensen, building on the work of Iversen and Stephens (2008), argued that coordinated market economies in social democratic welfare states will increase the demand for childcare and education policies as well as for active labour market policies. Such policies are
consistent with specialised and knowledge-intensive economies. Indeed, although these characteristics have always been particularly present in social democratic countries, human capital formation has become more central in all advanced economies since the 1990s. Moreover, welfare systems must face new risks that have emerged from ongoing societal transformations (Hemerijck, 2012; Pestieau and Lefebvre, 2018; Saltkjel, 2018; Taylor-Gooby, 2004).

While the analysis by Jensen (2011) was mainly focused on the role of the state, seen as an open system influenced by external environmental factors, we acknowledge that the inherent complexity of welfare systems extends beyond the dynamics between civil society and the state, requiring further consideration of the role of communities and social relationships, namely, the stakeholders involved in the welfare regimes’ dynamics (Häusermann, 2010; Levy, 1999; Vail, 2010). Moreover, third-sector organisations, including volunteering organisations, may affect welfare regimes independently from the role of the state.

All the aforementioned factors influence the development processes of welfare systems and emphasise their complexity and discontinuity. In this paper, we argue that all the complexity factors highlighted in the literature point towards a non-linear transition process in welfare regimes, which is highly affected by the dynamics between stakeholders representing different ideologies and social preferences. As the balance of power often shifts over time, the evolution of welfare regimes is often inconsistent. Moreover, although shifts in bargaining power are likely to impact policy makers’ decisions, they cannot completely overturn the existing systems. Therefore, the result of the evolution processes up to the 1990s decade is the implementation of policies that may be grounded in very different ideologies and social preferences, that is, a hybridisation of welfare regimes (Bertin and Pantalone, 2018; Ciccia, 2017; Yang et al., 2020). Furthermore, specific welfare policies (e.g., health, social care) have often been developed under different time frames and following different cultural and political rationales (hence, regimes), even within the same country (Bertin and Carradore, 2015).

According to Esping-Andersen himself, welfare regimes do not exist in pure form. Rather, they are an approximation of the most prevailing characteristics within a cluster of countries (Esping-Andersen, 1990). While the EA classification outlines a taxonomy of ideal-types that never find full realisation in reality, it constitutes a crucial interpretative framework, where the key element is the similarity of the welfare regimes within clusters. However, the recent debate have highlighted how such similarities may be weakening.

Our analysis naturally stems from this debate and is aimed to investigate whether the body of research carried out at the end of the evolutionary phase of welfare regimes confirms the similarities between states, or whether it highlights the presence of hybridisation processes which weakened the similarities between welfare systems. Specifically, we aim at testing the following hypotheses:

H1: hybridisation hypothesis. National welfare systems which consolidated in the first part of the 21st century and originally belonged to the same welfare regime, show low degree of similarities. In particular, we will consider the seminal EA regimes, i.e., Liberal (United States, Canada, Australia, United Kingdom), Conservative (Deutschland, Austria, France, the Netherlands), and Social democratic (Sweden, Denmark, Norway), and we will test whether welfare systems originally belonging to the same regime are less similar to each other and more or equally similar to systems belonging to other regimes.

H2: incoherence hypothesis. The classifications of countries vary substantially depending on the specific welfare policy concerned. In particular, due to the different time-frames and rationales in which health and social care policies were structured and reformed, we hypothesise that the degree of similarities between countries with respect to their healthcare systems is different than with respect to their social care systems.

3. Methods

3.1. Literature review

3.1.1. Study design and inclusion criteria

We performed a meta-analysis of studies that, since 2000, have proposed a classification of healthcare (HC) and social care (SC) policies. Healthcare policies are one of the largest areas of social welfare and concern the provision of health services to persons in need. Social care policies comprise a set of services aimed at helping families throughout the risks they encounter during their life course, e.g., childbirth, work life balance, the aging of a parent and the need for long-term care.

We followed several selection criteria for the literature review. First, we included original articles, published conference papers and books published in English following a peer-review process; research reports from accredited and internationally recognised organisations were included. Second, to enhance comparability, we selected studies that employed data from the first decade of the 2000s, which can be considered as the end of the expansion phase of welfare systems. We believe this interval allows for an acceptable equilibrium in the trade-off between a narrow time-frame (which enhances comparability) and a large number of studies (which enhances the results’ robustness). Our results are robust to restricting the time-frame to a shorter interval, excluding older studies. Third, we focused on research works that included a classification of healthcare or social care systems.

3.1.2. Search strategy

We employed electronic database searches (in PubMed, SCOPUS, Sage Journal Online and ScienceDirect) with the following keywords: ‘welfare state regimes/typologies’, ‘social policy’, ‘welfare services’, ‘healthcare/social care systems’, ‘cluster analysis’, ‘classification’. We also included papers/reports referred to by the outcomes of the database search, which complied with our inclusion criteria.

3.1.3. Outcomes

We identified 19 research outputs and a total of 22 welfare classification analyses (7 in healthcare and 15 in social care, with 3 papers performing 2 analyses each).

Table 1 summarises the data sources and methods implemented in the reviewed analyses. The data cover both European and non-European countries from 2000 to 2010, with two analyses employing data slightly outside this interval. Most of the reviewed studies produce welfare clusters through hierarchical cluster analysis (HCA), though some included non-hierarchical cluster analysis, principal component analysis or other logical methods. We checked that our results are robust to excluding studies employing non-statistical methodologies (logical methods).

The sets of indicators used to classify welfare states vary widely across papers. In total, 39 indicators were used for the classification of healthcare systems, while 55 indicators were used for social care systems (full details are available in Tables 1 and 2 in the Electronic supplementary material).

The 39 healthcare indicators can be categorised in six areas: expenditure and funding sources – EXP (e.g., social spending measures as % of GDP); governance – GOV (e.g., degree of decentralisation to sub-national government bodies); cost sharing – COST (e.g., visits to GPs and specialists); entitlement to receive care – ENT (e.g., complexity of GP access procedures); care coverage – C.COV (e.g., population covered by the healthcare system); and supply – SUP (e.g., number of GPs/physicians/specialists/nurses per capita).

The 55 social care indicators can be categorised in eight areas: parental leave – LEAVE (e.g., duration of parental leave); expenditure – EXP (e.g., public spending on childcare and elderly care services); gender issues – FEM (e.g., gender–employment gap); service coverage – S.COV (e.g., % of people aged over 65 receiving home-care services); entitlement to receive care – ENT (e.g., implementation of means
The vast majority of indicators are used in only one or two studies, while five indicators were shared by three or more studies. The most recurrent indicators are GP registration and GP remuneration (included in three healthcare studies), maternity leave duration (six social care studies), maternity leave compensation and female labour participation rate (three social care studies). This fragmentation may not be fully explained by the relative broadness of the categories “healthcare” and “social care” (e.g., social care includes both child care and elderly care), as it persists even among studies that focused on similar sub-categories such as “child care” and “elderly care.”

For example, among studies on elderly care, Kraus et al. (2010) considered public expenditure for long-term care (LTC) services and information on and benefits entitlement and quality assurance, while Verbeek-Oudijk et al. (2014) included public expenditure for long-term care services and information on and benefits entitlement and quality assurance, while Verbeek-Oudijk et al. (2014) included public expenditure for long-term care services and information on and benefits entitlement and quality assurance, while Verbeek-Oudijk et al. (2014) included public expenditure for non-institutional LTC services, and an indicator on the balance of public and private expenditure for LTC services. In the latter study, the authors used a non-hierarchical cluster analysis to identify distinct patterns of policy provision, while the former study used a hierarchical cluster analysis to identify distinct patterns of policy provision. The two studies differ in their approach to cluster analysis, the former using a non-hierarchical method and the latter using a hierarchical method. This choice of methodological approach is reflected in the different clustering results. The non-hierarchical method identifies three clusters, while the hierarchical method identifies four clusters. The non-hierarchical method also provides a more detailed breakdown of the clusters, with the first cluster consisting of countries with high levels of government provision, the second cluster consisting of countries with mixed levels of government and private provision, and the third cluster consisting of countries with high levels of private provision. The hierarchical method, on the other hand, identifies a single large cluster and two smaller clusters, with the larger cluster consisting of countries with high levels of government provision and the smaller clusters consisting of countries with mixed levels of government and private provision and countries with high levels of private provision. This difference in clustering results highlights the importance of considering the methodological approach when conducting policy analysis.

### Table 1: Outcomes of the literature review.

<table>
<thead>
<tr>
<th>Policy</th>
<th>Paper</th>
<th>Data sources, year(s)</th>
<th>Area of analysis</th>
<th>Method</th>
<th>Countries included</th>
</tr>
</thead>
<tbody>
<tr>
<td>healthcare</td>
<td>Bambra (2005)</td>
<td>OECD 1998, WHO 2002</td>
<td>EXP, C.COV, SUP</td>
<td>Logical typology</td>
<td>EU1 (- ES) AS1 CA CH NO USA</td>
</tr>
<tr>
<td></td>
<td>Jensen (2008)</td>
<td>OECD 2001</td>
<td>EXP</td>
<td>HCA</td>
<td>EU1 AS1 CA CH NO USA</td>
</tr>
<tr>
<td></td>
<td>Wendt (2009)</td>
<td>OECD 2007</td>
<td>EXP, ENT, GOV</td>
<td>HCA</td>
<td>EU1 GR LU PT</td>
</tr>
<tr>
<td></td>
<td>Relbling (2010)</td>
<td>OECD 2009</td>
<td>COST, ENT, GOV, SUP</td>
<td>HCA</td>
<td>EU1 (-IE) CH CZ GR PL PT</td>
</tr>
<tr>
<td></td>
<td>Jourmard et al. (2010)</td>
<td>OECD 2008</td>
<td>EXP, COST, ENT, C.COV, GOV, SUP</td>
<td>HCA</td>
<td>EU1 AS1 CA CH CZ GR HU IS KR LU NO PL PT SK TR</td>
</tr>
<tr>
<td></td>
<td>Bühn et al. (2013)</td>
<td>OECD 2008</td>
<td>GOV</td>
<td>Logical typology (logic tree)</td>
<td>EU1 AS1 CA CH CZ EE HU IL IS KR LU NO PL PT SI SK</td>
</tr>
<tr>
<td></td>
<td>Wendt (2014)</td>
<td>OECD 2010</td>
<td>EXP, GOV</td>
<td>HCA</td>
<td>EU1 AS1 CA CH CZ EE GR HU IL IS KR LU NO PL PT SI SK TR USA</td>
</tr>
<tr>
<td>social care</td>
<td>Knutte (2002)</td>
<td>Eurostat 2000</td>
<td>EXP</td>
<td>HCA</td>
<td>EU1 GR NO PT</td>
</tr>
<tr>
<td></td>
<td>Boje and Ejrnæs (2012)</td>
<td>ESS 2008</td>
<td>LEAVE, EXP, S.COV</td>
<td>HCA</td>
<td>EU1 BG CZ EE GR HU IS LT LU LV NO PL PT SI SK</td>
</tr>
<tr>
<td></td>
<td>Bambra (2004)</td>
<td>OECD 1998</td>
<td>LEAVE, FEM</td>
<td>Logical typology</td>
<td>EU1 (- ES) AS1 CA CH NO USA</td>
</tr>
<tr>
<td></td>
<td>Bambra (2007)</td>
<td>UN, 2005</td>
<td>LEAVE, FEM</td>
<td>HCA</td>
<td>EU1 AS1 CA CH NO USA</td>
</tr>
<tr>
<td></td>
<td>Jensen (2008)</td>
<td>OECD 2001</td>
<td>EXP</td>
<td>HCA</td>
<td>EU1 AS1 CA CH NO USA</td>
</tr>
<tr>
<td></td>
<td>Kraus et al. (2010)</td>
<td>OECD 2009</td>
<td>EXP, S.COV, ENT, COST, GOV</td>
<td>Non-hierarchical cluster analysis</td>
<td>EU1 (-IE) BG CZ EE HU LT LV PL PT RO SI SK</td>
</tr>
<tr>
<td></td>
<td>Thevenon (2011)</td>
<td>OECD 2005</td>
<td>LEAVE, EXP, S.COV</td>
<td>Principal component analysis</td>
<td>EU1 AS1 CA CH CZ GR HU IS KR LU NO PL PT SK USA</td>
</tr>
<tr>
<td></td>
<td>Chau and Sam (2013)</td>
<td>UN 2010 Taiwan National Statistics 2009, 2010</td>
<td>LEAVE, FEM</td>
<td>HCA</td>
<td>EU1 (- ES) AS1 CA CH HK KR NO SG TW USA</td>
</tr>
<tr>
<td></td>
<td>Cho (2014)</td>
<td>OECD 2012</td>
<td>LEAVE, EXP, FEM</td>
<td>HCA</td>
<td>EU1 AS1 CH GR NO PT USA</td>
</tr>
<tr>
<td></td>
<td>Verbeek-Oudijk et al. (2014)</td>
<td>SHARE (2010–2011)</td>
<td>EXP, GOV</td>
<td>Logical typology</td>
<td>EU1 (-FI-IE-UK) CH CZ EE HU PL PT SI</td>
</tr>
</tbody>
</table>

Legend: OECD = Organisation for Economic Cooperation and Development; WHO = World Health Organisation; ESS = European Social Survey; SHARE = Survey of Health Ageing and Retirement in Europe; EC = European Commission; ECHP = European Community Household Panel; UN = United Nations.

Cost = cost sharing; C.COV = care coverage; S.COV = service coverage; EXP = expenditure and funding sources; ENT = entitlement; FEM = gendered indicators; GOV = governance; LEAVE = parental leave; SUP = care supply.

HCA: hierarchical cluster analysis.
EU1: AT BE DE DK FR IT NL SE.
AS1: AU JP NZ.

### Table 2: Results for healthcare similarity.

<table>
<thead>
<tr>
<th>at</th>
<th>be</th>
<th>De</th>
<th>Dk</th>
<th>es</th>
<th>fi</th>
<th>fr</th>
<th>ie</th>
<th>it</th>
<th>nl</th>
<th>se</th>
<th>uk</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.71</td>
<td>0.57</td>
<td>0.14</td>
<td>0.17</td>
<td>0.14</td>
<td>0.71</td>
<td>0.33</td>
<td>0.29</td>
<td>0.29</td>
<td>0.29</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>0.14</td>
<td>0.17</td>
<td>0.14</td>
<td>1.00</td>
<td>0.33</td>
<td>0.29</td>
<td>0.43</td>
<td>0.29</td>
<td>0.14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.14</td>
<td>0.57</td>
<td>0.33</td>
<td>0.29</td>
<td>0.43</td>
<td>0.57</td>
<td>0.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.67</td>
<td>0.57</td>
<td>0.14</td>
<td>0.50</td>
<td>0.43</td>
<td>0.57</td>
<td>0.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.83</td>
<td>0.17</td>
<td>0.20</td>
<td>0.17</td>
<td>0.33</td>
<td>0.50</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.14</td>
<td>0.14</td>
<td>0.14</td>
<td>0.14</td>
<td>0.33</td>
<td>0.50</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.33</td>
<td>0.29</td>
<td>0.43</td>
<td>0.29</td>
<td>0.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.00</td>
<td>0.50</td>
<td>0.33</td>
<td>0.67</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.43</td>
<td>0.29</td>
<td>0.29</td>
<td>0.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.14</td>
<td>0.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: The table presents the similarity values calculated using the Jaccard index, which measures the similarity of two sets by dividing the size of their intersection by the size of their union. The values range from 0 (no similarity) to 1 (perfect similarity). The table includes only the values above the diagonal, as the values below the diagonal are the transposes of the values above the diagonal. The table also includes a row of 0s for the diagonal, as the similarity of a set with itself is always 1. The table includes missing values represented by dots. The values in bold indicate the highest similarity values for each row.
results section, we discuss how such heterogeneity in the indicators selection might affect our results.

3.2. Relative index of welfare similarity and sample selection

We build an index of “welfare similarity” to capture, separately for the social care and healthcare sectors, the extent to which welfare systems have been grouped together in the papers we reviewed (i.e., found to share substantial common characteristics), as suggested in recent analyses (Buhr and Stoy, 2015; Powell et al., 2020). However, we depart from previous methodologies in that we factor in the country-specific number of observations (i.e., the number of analyses a specific country appears in). We thereby show that failing to account for this information may lead to biased results.

Let $N_i$ be the number of analyses which include country $i$. The maximum number of country-specific observations corresponds to $N_i$; that is, max ($N_i$) = $N^*$, where $N^*$ = 15 for social care and $N^*$ = 7 for healthcare. Furthermore, we define $N_{ij}$ as the overall number of times the dyad made of countries $ij$ is included in a study (e.g., countries $i$ and $j$ jointly appear in the same study), with $N_{ij} = N_{ji}$ and $w_{ij} = w_{ji}$. Only a small subset of countries appear in all studies, as described in Tables 3 and 4 in the Electronic supplementary material. European countries appear more frequently than non-European countries, which translates to a clear geographical selection in terms of the country-dyad observations: European countries are much more likely to be simultaneously present in an analysis than non-European countries.

Following Buhr and Stoy (2015) and Powell et al. (2020), we define a variable $W_{ij} (= W_{ji})$ which counts how often countries $i$ and $j$ were grouped (linked) in the same welfare cluster (for any $i$ and $j$). Such a measure is, however, likely to be affected by the fact that both the number of total appearances $N_i$ and the number of joint appearances $N_{ij}$ largely differ across countries. For example, among social care studies, Canada is included in the same welfare cluster as Austria three times ($W_{CA,AT} = 3$), while France and Austria are clustered together six times ($W_{FR,AT} = 6$); however, Canada and Austria appear together in just five studies ($N_{CA,AT} = 5$), while France and Austria appear in all studies ($N_{FR,AT} = 15$). Hence, the higher number of links (in absolute terms) between Austria and France may be a direct result of the difference in the number of joint observations. However, when comparing the absolute number of links $W_{ij}$ to the number of joint appearances $N_{ij}$, Austria is linked to Canada 60% of the times that they are studied together (3/5), while Austria is only linked to France 40% of the times (6/15). Therefore, adopting $W_{ij}$ as a measure of similarity may bias the results. To avoid this distortion, we therefore introduce a relative measure of similarity $w_{ij}$ that, for any pair of countries $i$ and $j$, compares the number of joint classifications $W_{ij}$ to the number of joint appearances of $i$ and $j$, namely, $N_{ij}$.

We therefore introduce a relative measure of similarity $w_{ij}$ between countries $i$ and $j$, defined as:

$$w_{ij} = \frac{W_{ij}}{N_{ij}}$$

with $N_{ij} \geq 0.666 (N^*)$, $\forall i,j \in \{1, \ldots, N\}$ for social care studies $N = 15$ for social care studies $N = 7$ for healthcare studies

For example, if countries $i$ and $j$ were simultaneously analysed in 10 studies and linked together in two (i.e., 20% of the times they are analysed together), then $w_{ij} = 0.2$.

However, the interpretation of $w_{ij}$ may be misleading if $N_{ij}$ varies substantially across pairs $(ij)$. For example, two countries $a,b$ which appear jointly in just one study and are linked together would have $N_{ab} = 1$, $W_{ab} = 1$ and $w_{ab} = 1$. Similarly, two countries $c,d$ which appear jointly in 15 studies and are always linked together would have $N_{cd} = 15$ and $W_{cd} = 15$ but still have $w_{cd} = 1$. Although the relative similarity is the same, the evidence emerging for countries $c$ and $d$ is arguably more robust than for countries $a$ and $b$, as the former is based on 15 studies, while the latter is based on one study.

To allow for a consistent interpretation of the $w$ index, we restrict our sample to country-dyads with a sufficiently high number of appearances. Specifically, we select our sample of countries in such a way that for any pair of countries $ij$ they are jointly present in at least two-thirds of the total number of studies. That is, $N_{ij} \geq 0.666 (N^*)$, $\forall i,j$. For the social care sector, where $N^* = 15$, we therefore include any country which has at least 10 joint appearances with any other country in the sample. After applying such restriction, a $w_{ij}$ score of 1 implies that countries $i$ and $j$ were simultaneously clustered together 100% of the time they were analysed together, corresponding to a minimum of 10 times, and a maximum of 15 times. For the healthcare sector, our threshold corresponds to five joint appearances (that is, 66.6% of $N^* = 7$).

By applying our inclusion criteria, we are left with a set of 12 European countries, identical for both the social care and healthcare sectors: Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, the Netherlands, Spain, Sweden and United Kingdom. These countries share at least 12 joint appearances for social care (with the exception of Ireland and Spain, which share 10 appearances), and at least 6 joint appearances for healthcare.

Although our inclusion criterion is arbitrary, our results are robust to the adoption of a stricter inclusion threshold, limiting the sample to countries that are jointly present in at least 80% of the studies with any other country (available upon request from the authors); or looser

<table>
<thead>
<tr>
<th>Table 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descriptive statistics for similarity index.</td>
</tr>
<tr>
<td>Similarity index</td>
</tr>
<tr>
<td>Minimum</td>
</tr>
<tr>
<td>25th percentile</td>
</tr>
<tr>
<td>Median</td>
</tr>
<tr>
<td>75th percentile</td>
</tr>
<tr>
<td>Maximum</td>
</tr>
</tbody>
</table>
thresholds, including countries that are present in at least 3 studies with any other country (Supplementary Material 2).

4. Results

We now show the results of the welfare similarity analysis, separately for healthcare and social care policies. We aim to understand (1) the extent to which the typologies identified by the existing literature on healthcare and social care systems overlap with the EA classification, and (2) whether the emerging typologies are consistent between healthcare and social care studies.

4.1. Healthcare

We summarise our results for the healthcare similarity analysis by plotting a similarity network in Fig. 1. The strength of the association between any two countries is represented by the thickness of the segment which links them and expressed explicitly as a ratio coefficient (which can be easily converted to percentage terms) in Table 2, following the definition in equation (1). Table 5 in the Electronic supplementary material reports the absolute number of links between all the countries included in the reviewed papers. We also report some descriptive statistics relating to the distribution of the similarity index in Table 4.

Overall, our results highlight the lack of a “pure” overlap between the EA typology and health-care typologies.

First, the descriptive statistics in Table 4 column (a) show that, among the 12 countries in the sample, no pair of countries have a zero similarity coefficient, as the minimum similarity score is 0.14 (i.e., all countries have at least one link with any other country). Moreover, the median similarity score is 0.33. Given that any country dyad appears in at least six studies, this means that half of the country dyads are clustered in the same healthcare regime at least twice (33%). As such, these results suggest that healthcare systems are more hybrid than the EA classification would suggest.

Second, focusing on the countries traditionally included in the Scandinavian regime (Sweden, Denmark and Finland), the similarity coefficient does not exceed 0.57. Moreover, some of these countries strongly show links with liberal or continental countries. Spain has a similarity index of 0.5 with Sweden and 0.83 with Finland. Denmark and United Kingdom have a similarity index of 0.86.

Third, liberal countries such as United Kingdom and Ireland share a similarity score of 0.67, yet United Kingdom is, as mentioned, more strongly linked to Denmark (0.86) and similarly linked to Italy (0.57) and Spain (0.5). Ireland has a much stronger link to Italy (coefficient of 1, meaning that the two countries are always classified together) than to United Kingdom, while having a 0.5 similarity index with Denmark and the Netherlands. Moreover, the Netherlands is another example of the absence of a clear similarity pattern, as it shows comparable similarity scores with countries traditionally belonging to very different regimes, such as Italy (0.43), Ireland (0.5) and Belgium (0.43).

However, a relatively more stable picture emerges for the continental regime countries (Austria, Belgium, France and Germany). These continental countries share high similarity scores, ranging between 0.57 and 1. Conversely, the similarity between the continental countries and the remaining countries always lies below 0.43.

4.2. Social care

The results for the similarity structures in social care studies, shown in Fig. 2, point to the coexistence of multiple regimes that are not clearly differentiated, yet in a different way than for the healthcare sector (results are summarised in Table 3). Table 6 in the Electronic supplementary material reports the absolute number of links between all the countries included in the reviewed papers.

As reported in Table 4(b), the median similarity index is 0.27, indicating that half of the country dyads are classified in the same welfare regime fewer than 30% of the times (which corresponds to roughly four studies). As the maximum similarity index is 0.87, no dyad can be considered a “pure system”. However, unlike for healthcare policies, we find dyads with zero similarity.

Scandinavian countries exhibit low, though mostly non-zero, similarity scores with non-Scandinavian countries. However, while Denmark and Sweden have a high similarity score of 0.87, they have a weaker link with Finland (0.5), thus highlighting the lack of a “pure” cluster within the EA Nordic regime.

Among the Mediterranean countries, a similar pattern emerges: Italy and Spain have a similarity score of 0.69; however, Italy is also linked to Austria, Ireland and the Netherlands in 40% of the studies, while Spain is linked to Ireland in 40% of the studies.

The similarity score among continental countries ranges between 0.4 and 0.67, suggesting that they are grouped in the same policy cluster in only around half the analyses; moreover, Austria and Belgium have slightly lower similarity scores with Ireland and United Kingdom (0.43).

The results for countries traditionally in the liberal EA welfare group reinforce the aforementioned evidence for the existence of hybrid welfare typologies. The similarity score for United Kingdom and Ireland (0.58) is only slightly higher than that between United Kingdom and France (0.50), United Kingdom and Austria (0.43) and United Kingdom and Finland (0.43); moreover, it is only slightly higher than the score between Ireland and continental (Austria and Belgium) and

Fig. 1. Frequency of country pairs being grouped into the same cluster with respect to healthcare policies (Bastian et al., 2009).

Fig. 2. Frequency of country pairs grouped into the same cluster with respect to social care policies (Bastian et al., 2009).
Mediterranean (Italy) countries, with an index of around 0.4. Similar to what emerged from the healthcare analysis, the Netherlands is paired with countries belonging to different regimes (continental, Mediterranean and liberal), with an index score of around 0.5 (around half the analyses).

4.3. Comparisons across welfare areas

Overall, the comparison between the findings for healthcare policies and social care policies highlights that some traditional EA welfare typologies are particularly unstable.

First, a substantial change emerges in the similarity network of Mediterranean countries. Within healthcare, Italy and Spain are very rarely linked together (similarity score of 0.17), while both countries share substantial similarities with (different) Nordic and Liberal countries. However, within social care policies, such links are almost non-existent. Specifically, stark differences appear in the similarity index between Spain and Denmark (HC 0.67; SC 0), Spain and Finland (HC 0.83; SC 0.17), Spain and Sweden (HC 0.5; SC 0) and Spain and United Kingdom (HC 0.5; SC 0.25); as well as between Italy and Denmark (HC 0.43; SC 0); Italy and Ireland (HC 1; SC 0.42) and Italy and United Kingdom (HC 0.57; SC 0.21). Conversely, Italy and Spain share a higher similarity score in social care policies (0.69), than in healthcare (0.17).

Second, a similar change can be identified for Liberal countries. While United Kingdom and Ireland are similarly linked in both healthcare (0.67) and social care policies (0.58), both are much more strongly linked with the Nordic block and Italy in healthcare than in social care analyses. For United Kingdom, the similarity scores substantially change with respect to Denmark (HC 0.85; SC 0.14), Sweden (HC 0.57; SC 0.14) and Italy (HC 0.57; SC 0.21). Ireland’s similarity index with Denmark (HC 0.5; SC 0) and Italy (HC 1; SC 0.42) exhibits a similar drop.

4.4. Sensitivity checks

4.4.1. Could the choice of indicators affect our findings?

The hybrid clusters emerging from our results may be, in principle, an artifact of the choice of indicators employed in the specific analyses if, for example, studies employing similar indicators lead to similar country-clusters. Hence, we might wonder whether our findings would be robust to an alternative choice of welfare indicators. However, we argue that this concern is not relevant to our analysis, as the studies we reviewed employ very different indicators (Section 3.1). For example, United Kingdom and Denmark are paired together in six out of seven healthcare studies, which employ a total of 20 indicators. Fourteen indicators are uniquely used by single studies (not shared by other studies). Five indicators are shared by two studies; one indicator is shared by three studies on of GPs). Within social care policies, Spain and Ireland are clustered together in four studies, which employ a total of 19 indicators, each of them employed by just one study.

On the other hand, the variability in the choice of the indicators might explain the hybrid classification of countries. Although essentially not empirically testable, we argue that this explanation would be in line with our starting hypotheses. In a context of welfare systems hybridisation, particular sub-sections of the healthcare or social care systems might follow different rationales. Hence, studies employing different indicators to characterise health or social care services could capture such heterogeneity. In other words, should the observed hybrid clusters be due to the variability in the choice of the indicators, this would confirm, rather than contradict, our starting hypotheses.

4.4.2. Could different cluster methodologies affect our findings?

A. In principle, the observed hybridisation of welfare clusters could be affected by the heterogeneity in the methods used by the reviewed studies, rather than depicting an actual heterogeneity across welfare systems. We believe this not to be a concern for our findings, for two main reasons. First, our review only selected studies published in top field journals and books (the vast majority), or in working paper series edited by world renowned organisations with strong quantitative focus. Although the assessment of the quality of a method adopted is ultimately subjective, we believe that the peer-review process that such studies underwent before publication should already be a partial guarantee of their value. Second, we performed a robustness test by arbitrarily excluding from the reviewed classifications those which were not resulting from a statistical algorithm (7 classifications). The resulting similarity indices, available upon request, entirely confirm our main findings, suggesting that they are not an artifact of the variation in methods.

4.4.3. Could the selected time-frame affect our findings?

Our review includes studies using data from the 2000s, with some study also employing data from 1998 or 2012. This time interval is usually referred as the “post-expansion” era of welfare state evolution. However, scholars have noted that since the end of the 1990s, a process of rationalisation has been put in place, which has resulted in a number of welfare state reforms. We might therefore be concerned that the hybridisation might result from mixing studies from the early and the late 2000s. We therefore replicated our analysis on the studies using data from 2005 onward (the majority of the studies), and obtained results entirely in line with our main findings, suggesting that they are not driven by the studies timeframe.

4.4.4. Could a larger country selection affect our findings?

Our main analysis focused on 12 European countries which have been jointly included in at least 66% of the reviewed studies (section 3.2). However, our main findings are robust to broadening the analysis to countries which are less often included in the reviewed studies. Specifically, when enlarging the sample to 21 countries, covering Southern, Eastern and Northern Europe, all of our main findings are confirmed (results and methods are available in the Supplementary Material Section 2). However, as such results are partially based on countries appearing in a small number of studies, we prudently consider them as less robust than our main findings.

5. Discussion and conclusion

While the seminal work by Esping-Andersen (EA) has provided a classification of welfare systems during their developing phase in the second half of the previous century, welfare states have, since then, undergone major transformation processes in most Western countries, which might have weakened the representativeness of Esping-Andersen’s classification. Our study stems from two major results in the existing literature. On the one hand, previous studies have hypothesised a progressive hybridisation of welfare typologies, which results in welfare systems borrowing characteristics from more than one regime. On the other hand, public welfare systems have been shown to follow different rationales (e.g., liberal, corporatist and social democratic) in different areas of service provision, even within the same country.

In this paper, we provide novel evidence that the recent literature focusing on healthcare and social care systems in Europe has identified country-clusters which are, to different extents, not fully overlapping with Esping-Andersen’s classification. Our main findings are twofold: first, we provide evidence for a progressive hybridisation of healthcare and social care systems classifications across European countries (H1: hybridisation). With respect to healthcare systems, our results highlight the absence of clear clusters across European countries. This is mainly due to the inconsistencies in the classifications provided by the reviewed studies. For example, both Sweden and the Netherlands are never unanimously classified together with any other country in the sample. Moreover, most studies group together countries which originally belonged to different EA typologies. For example, Mediterranean-welfare countries such as Spain and Italy are often clustered together with Finland, Denmark and Ireland which belong to a social democratic
and liberal welfare tradition, respectively. While the healthcare systems of central-Europe countries (the corporatist regime) exhibit a stronger joint similarity, they are often clustered together with countries belonging to different EA typologies.

Similar conclusions can be drawn from social care studies. On the one hand, the reviewed studies are never unanimous in classifying any two countries in the same cluster. On the other hand, several countries exhibit a very low degree of similarity with any other country in the sample. This is particularly evident for continental-regime countries, such as Austria and the Netherlands, as well as for social democratic (Finland) and liberal (Ireland and United Kingdom) countries. It is still possible to identify clusters of countries which are consistent with the original EA classifications. Still, even within these clusters, the degree of similarity is far from unanimous across studies, and typical clusters (such as Spain and Italy; Belgium, Germany and France; Sweden and Denmark) are only identified by roughly two-thirds of the studies. All in all, these results seem to confirm the hybridisation hypothesis.

Second, we show that welfare typologies and similarities can substantially differ across policy areas (H2: incoherence). For example, the Spanish healthcare system is clustered together with Finland’s and Denmark’s, suggesting the coexistence of characteristics from both the Mediterranean and social democratic regimes. Another Mediterranean country, Italy, has similarly been linked with Ireland and United Kingdom with respect to its healthcare system. However, within social care policies, Spain and Italy are found to be closely linked. This suggests that the rationales for welfare provision, even within a country, are often inconsistent across policy areas, therefore providing support for the incoherence hypothesis.

There are several reasons for the observed lack of consistency across studies in the classification of welfare systems. First, from a methodological perspective, we note that only a few studies employed a similar set of indicators (Powell et al., 2020; Yörük et al., 2019). While this methodological fragmentation could partially explain the lack of strong links between countries that originally belonged to a well-defined welfare regime, it can hardly explain our findings of hybrid welfare clusters, where countries from different EA regimes are consistently grouped together. Moreover, we have shown that our results are not sensitive to the methodology employed by the reviewed studies. A second reason lies in the complexity of the unfolding developments of welfare systems (Bonoli, 2001; Jensen, 2011), whose transformations have not been continuous nor constant. Such transformations are the outcome of power dynamics that change overtime, including the resistance to changes in welfare institutions, and the reactions to changes in societal needs and preferences. Hence, the hybridisation of welfare characteristics across systems is likely to be an outcome of these dynamics.

Our study provides two relevant contributions for both the academic and the policy debate on welfare policies. First, we highlight that, against a background of changing nature of welfare systems, the comparative analysis and classification of welfare policies should devote more focus to the complexity of the unfolding developments of welfare systems and their transformations, which are not necessarily continuous and constantly in line with standard classifications. The relevance of welfare systems classifications is strongly dependent on the ability of such classification to capture the characteristics of the system itself. In our view, our findings underline the importance of narrowing the focus of welfare regime analyses on specific policy areas, to enhance the relevance of the classifications themselves. With respect to health and social care policies in particular, our findings suggest that the existing studies lack consistency in the choice of dimensions and indicators for classification purposes, and do not come to an accepted taxonomy.

Hence, further research is needed to strengthen the specificity of healthcare and social care systems’ studies, for example, by including more specific indicators which might better capture the transformation processes which are interesting welfare systems in the last decades.

Second, our study suggests that enhancing the specificity of welfare classifications might be relevant for future comparative empirical research, which often have to rely on general classifications of countries welfare systems while studying specific policy areas, due to data limitations (e.g., Carriero et al., 2017; Floridi et al., 2021). Our findings show that broad classifications might underestimate the hybridisation of countries welfare systems in specific policy areas. Hence policy-specific welfare typologies may prove more informative to academics and policymakers than general classification of the welfare state as a whole.

Finally, we note that, due to the country selection in our study, our findings are relevant with respect to the core set of mature European welfare states. Further comparative research is needed to broaden the perspective beyond the Western European regimes, for example, to Eastern Europe, American and Asian countries.

Authors’ contribution

Declaration of competing interest
None declared.

Acknowledgments
Ludovico Carrino is supported by the Economic and Social Research Council (ESRC) (grant number: ES/S01523X/1, IN-CARE project). The views expressed are those of the author(s) and not necessarily those of the ESRC or King’s College London.

Appendix A. Supplementary data
Supplementary data to this article can be found online at https://doi.org/10.1016/j.socscimed.2021.114086.

References