Western&Graduate&PostdoctoralStudies

8-17-2023 10:00 AM

Matters of Measurement: Investigating the Universal Welfare State

Kaitlin Pauline Wannamaker, *The University of Western Ontario*

Follow this and additional works at: https://ir.lib.uwo.ca/etd

Part of the Quantitative, Qualitative, Comparative, and Historical Methodologies Commons

Abstract

This master's thesis investigates the mechanisms leading to the development of robust and generous welfare states, focusing on the ongoing debate around targeting versus universalism in welfare state studies. By leveraging multiple welfare state datasets and expanding the scope of welfare outcomes using cross-national time-series analyses, it uncovers the pivotal role that measurement plays in understanding universalism's effects. This research unveils substantial variations in outcomes based on the universalism measures used, thus highlighting how our perception of the welfare state is deeply entwined with our methodological choices. The findings of this study not only carve new paths in the field of comparative welfare state literature but also set the stage for future research agendas. This work is a critical step towards more fully comprehending the intricacies of the welfare state debate, alongside the importance of conceptual precision and measurement alignment in welfare state studies.

Keywords: Welfare State, Targeting, Universalism, Social Benefits, Redistribution, Social Policy, Methodological Choices.

Summary for Lay Audience

The following thesis grapples with a significant social policy debate: how can we most effectively administer social benefits to reduce poverty and inequality? This debate oscillates between two approaches: (1) targeting—giving aid specifically to the poor, and (2) universalism—dispersing benefits to everyone regardless of income level. Yet, 'universalism' lacks a consistent definition, making it difficult to measure and compare its outcomes. My thesis aims to clarify this targeting-universalism debate by examining how different definitions of universalism affect the outcomes of welfare states, such as generosity, poverty, and inequality levels. I do this using data from 22 developed countries to analyze how changes in defining and measuring universalism influence these outcomes. This work has two key implications. Firstly, it can help policymakers create more effective strategies to tackle poverty and inequality. Often, studies have used a single measure of universalism, which may not capture its full impact. By examining multiple measures, my research project provides a more comprehensive understanding of universalism's effects, and thereby shedding light on the importance of methodological consistency in social policy research. Secondly, the debate between targeted and universal programs carries normative implications. It relates not only to effective policy design but also to the kind of society we want to live in. Do we aim for an inclusive society that treats everyone equally, or do we prioritize aid for those most in need? Furthermore, my research addresses methodological issues in the existing literature, such as differing conceptualizations of universalism and the use of different measures to compare targeting and universalism's effects. In summary, my thesis is a critical contribution to social policy debates, as it not only deepens our understanding of how welfare states can combat poverty and inequality, but also prompts us to reflect on the kind of society we want to create.

Acknowledgments

I am deeply grateful to acknowledge those whose invaluable contributions have shaped my thesis. First and foremost, I express my profound gratitude to my supervisor David Calnitsky. His scholarly insights, unwavering support and belief in me, and dedicated mentorship have been the cornerstone of this endeavor. His guidance was instrumental in navigating the complexities of my research and fostering my growth. I would also like to express my sincere appreciation to my committee member Patrick Denice for his constructive criticism, detailed feedback, and always expressing his interest and excitement in my work. Special thanks to Professors Tracey Adams and Rachel Margolis, whose guidance and continual support have set me on the path towards my PhD at McGill University starting this Fall. Lastly, I extend my heartfelt gratitude to Sandra, Gayle, and Erin, the trio who have a solution to every problem. Their relentless support and problem-solving abilities have undoubtedly lightened the burdens for all Soc grad students. Thank you, everyone.

Table of Contents

List of Tables

List of Figures

List of Appendices

Preface

In engaging with the 'targeting versus universalism' debate that characterizes the ongoing literature emerging from comparative welfare state studies, I address the issues of conceptual and operational variation in the concept of 'universalism'. While some scholars argue the most effective way to administer benefits is to direct funds strictly to the poor, others advocate for the universal allocation of funds to all citizens, as programs with a broad base of recipients cultivate political support in favour of high levels of welfare spending. Despite this rich debate, the concept of 'universalism' itself is often granted less attention. What is meant by universalism is not always clear, and one's chosen definition may very well impact the relationship in question. Using a unique combination of welfare state data, I conduct cross-national time series analyses of 22 OECD countries to evaluate how five different measures of universalism fare in terms of outcomes on social spending, social transfers, poverty, and inequality. My findings indicate that the merit of 'universal' social policy fluctuates substantially according to one's chosen operationalization of universalism, with some supporting the universalist argument more than others. The main contributions of my M.A. thesis are methodological, demonstrating the important role played by measurement, alongside offering a holistic analytical strategy for addressing social policy debates within welfare state scholarship.

Chapter 1: Introduction

1 Scope of Thesis

The question of whether social benefits should be targeted or universal has figured prominently in debates among welfare state scholars for decades. Should welfare benefits be reserved only for the neediest members of society? Or should welfare programs and benefits be universally accessible? The response of some scholars is that the frugality achieved by targeting maximizes the spending directed to the poor (Greenstein 1991). Others propose an alternative road to benefit generosity, claiming that by covering a broader segment of the population, universalism motivates the political will necessary for further social spending reform (Brooks and Manza 2007). However, despite its routine use, *what* is meant by 'universalism' is not always clear. Universalism is rarely explicitly defined within the literature, and measures of universalism vary substantially, which may very well translate into differences in findings determined by measurement choice.

For example, some scholars treat universalism and targeting as bookends of a spectrum (Korpi and Palme 1998), while others are guided by the view that universalism is a distinct concept and should be defined independently of targeting (Brady and Bostic 2015). Amplifying the uncertainty stemming from this is the fact that most works deploy only a single measure of universalism within their analyses, and fewer studies incorporate multiple indicators of universalism (Brady and Burroway 2012; Rasmussen and Knutsen 2021). Given the lack of a clear and universally accepted definition of 'universalism' in the literature, the dispute over whether social benefits should be targeted or universal remains in a state of unresolved tension, leading to a standstill in the management and delivery of social benefits.

I argue that gaining further insight into the debates on universalism's role within the welfare state requires thorough examination *across* a wide range of universalism measures. I first focus on identifying the relationship between the extent of a welfare state's universalism and welfare state outcomes such as social spending, poverty, and inequality levels. Subsequently, I focus specifically on issues of measurement by

investigating how sensitive the potential relationships between universalism and welfare state outcomes are to the way in which universalism is operationalized. Drawing on pooled time series welfare state data for 22 OECD countries merged with universalism data from six different sources, I clarify the circumstances under which universalism is associated with social spending and other important welfare state outcomes. In this process, I utilize four primary dependent variables to scrutinize the potential impacts of universalism: social spending, social transfers, inequality, and poverty. The variable *'social spending'* encapsulates a wide range of government expenditures geared towards addressing societal needs, measuring them as a portion of a state's GDP. *'Social transfers,'* on the other hand, refers to government-provided cash transfers, excluding social transfers in-kind, shedding light on the state's redistributive efforts. *'Inequality'* is gauged by the post-tax and transfer Gini coefficient of household income, offering insights into income distribution after state redistribution. Lastly, *'poverty'* is assessed by the post-tax and transfer relative poverty rate, highlighting the efficacy of state interventions in poverty reduction. Studies on inequality and poverty are increasingly considering underlying dynamics such as income mobility and poverty persistence (Saunders 2021). I will provide a more in-depth discussion and exploration of these variables in Section 3.1.

1.1 Rationale and Literary Contribution

The implications of my proposed work are twofold. First, there is pragmatic value in providing greater detail into the relationship between universalism and welfare state outcomes. Clarity aids policymakers in developing the sharpest tools with which to cut back poverty and inequality. While much literature has been dedicated towards the targeting-universalism debate, an overwhelming majority of these works have done so with the application of a single recycled measure of universalism within their analyses (Korpi and Palme 1998; Kenworthy 2011; Marx, Salanauskaite, and Verbist 2013; Gornick and Smeeding 2018). The development of a customary approach within a research area for operationalizing a core variable would not make for such a compelling call for further investigation if researchers justified their choice of measurement by

arguing either (1) that it is the best, or perhaps only, measure currently available, or (2) that the measure strongly correlates with any alternative measurement choices.

However, for the vast majority of literature concerned with the targeting-universalism debate, the extent of scholarly dedication to a singular customary measure of universalism far surpasses the attention granted to establishing either of these two arguments. This disconnect between the *execution* and *explanation* of methodological choices is evidenced by the amount of journal articles that allocate a disproportionate amount of their word count to providing a standard historical overview detailing the origins of how the measurement of universalism came to be, rather than offering a comprehensive and critical evaluation of said measure. One such example of granting insufficient attention to the assessment of one's universalism measure and quickly moving to the analysis stage can be seen in Marx, Salanauskaite, and Verbist's 2013 paper. With the goal of revisiting the targeting-universalism debate with updated data, the authors briefly note a few potential issues and limitations associated with the universalism measure used in past studies, but ultimately proceed with using the very same measure for their own analyses.

This tendency to gloss over the potential implications associated with the composition, advantages, and limitations of key variables stifles progress within the discipline and discourages experimenting with alternative methods for measuring concepts. In other words, when it comes to the operationalization of central variables used within studies investigating welfare strategies and their subsequent outcomes, welfare state scholarship seems to be held back by an academic variation of a "if it ain't broken, don't fix it" mentality. Consequently, little work has examined the relationship using a variety of measures of universalism. Therefore, I argue that to most effectively demystify the targeting-universalism debate, one must look *across* measures of universalism. After all, the *absence* of broken is an insufficient reason for ruling out efforts to uncover the *presence* of something better.

The second implication of my thesis research is normative, as the institutional characteristics of the welfare state may have an influence not only on generosity and

poverty levels, but also on our assumptions regarding how our communities ought to operate. Through its restrictions or lack thereof, the welfare state defines 'who counts'. One might expect targeted programs to be conducive to an exclusionary environment, while universal programs may foster social cohesion and social solidarity. This makes debates regarding targeting versus universalism of central importance not only in terms of developing effective policy, but also in contemplating the type of world in which we wish to live.

1.2   Objectives and Expected Results

The objective of this thesis is to contribute to longstanding social policy debates by determining the extent to which differences in measurement impacts the role of universalism. A consideration of measurement sheds light on previous research findings and the classic targeting-universalism debate over welfare state outcomes. Doing so will allow me to provide greater insight into how generous and robust welfare states come about.

The first literary contribution of my proposed work is pragmatic in value by providing greater detail into the relationship between universalism and generosity. Clarity aids policymakers in developing the sharpest tools with which to cut back poverty and inequality. The second literary contribution is normative, as the institutional characteristics of the welfare state may have an influence not only on generosity levels, but also on our assumptions regarding how our communities ought to operate. I expect my results to reveal that targeted programs encourage an environment of exclusion, while universal programs foster more social cohesion and social solidarity. These expected findings can largely influence debates on targeting versus universalism in terms of developing effective policy, alongside in contemplating the type of world in which we wish to live.

Chapter 2: Literature Review

2   Targeting vs. Universalism in Comparative Welfare State Studies

Two competing strategies of welfare distribution—targeting and universalism—possess substantial influence over the defining factors comprising overarching welfare systems. Existing in a conceptual limbo, universalism poses a challenge to its empirical operationalization. Defined generally, 'universalism' is an approach to social policy characterized by high levels of social benefit accessibility across demographic groups. These high levels of accessibility are typically premised on the absence of eligibility requirements for receiving benefits, alongside viewing the receipt of benefits as a matter of citizenship over deservingness (Rothstein 1998). This broad perspective on social policies is grounded in the concept of universalism, a principle suggesting that these policies should apply to everyone, regardless of their socioeconomic status (Anttonen and Sipilä 2014).

As Thompson and Hoggett (1996) further elucidate, universalism supports the idea that social policies should be applied equally to all citizens without discrimination. This concept stands in contrast with selectivism, which posits the allocation of resources based on specific criteria, such as income or health status (Thompon and Hoggett 1996). Alternatively, 'targeting' entails an approach to social policymaking centered around strict eligibility requirements and means testing (Nelson 2004). In practice, however, the ways in which universalism and its related concepts are defined fluctuate substantially across welfare state studies. For instance, Sainsbury (1996) posits that universalism involves social benefits and services anchored in compulsory legislation. Here, universalism calls for entitlements designed for the entire population, irrespective of financial needs, thereby ensuring equal access to services and uniform benefits. Crucially, this equal access to benefits and services is understood as a central component of citizens' rights. Thoroughly outlining the defining characteristics of targeted and universal welfare state strategies – and the stark contrast between them – contextualizes the emergence and persistence of both scholarly and political division over issues related to the welfare state. Moreover, it underlines citizens' rights to benefits and services. Given the extensive

discourse around targeting and universalism – and their significant influence on social policy outcomes – the forthcoming Section 2.1 will provide an in-depth overview of the ongoing targeting versus universalism debate by unpacking and contrasting the perspectives of pro-universalism scholars and pro-targeting scholars.

2.1 Overviewing the Targeting-Universalism Debate

Scholarly debate over the merit of targeted and universalist social policies often begins with reference to Korpi and Palme's seminal work, *The Paradox of Redistribution and Strategies of Equality*. KP[1] are concerned with investigating the divergent outcomes in terms of inequality, poverty, and social spending of welfare states with contrasting 'logics' of operation. KP propose the existence of a social policy paradox, wherein the more social benefits are targeted exclusively to the poor, the less successful the welfare state is at reducing poverty and inequality (Korpi and Palme 1998). While the topic of targeted social benefits had already been established within the literature well-before KP's *Paradox of Redistribution*, their work is often deemed responsible for revitalizing scholarly attention over the targeting-universalism question. The inspiration sparked by KP has since snowballed into a large body of literature that aims to investigate both the inner workings of the welfare state, alongside the determinants of robust public institutions characterized by generous social spending and desirable levels of redistribution. However, the successive scholarship following KP (1998) has done more to preserve rather than resolve the divide between those in favour of targeted and universal social policies.

The divide and debate between targeted and universal social policies have not only been influenced by differing ideological standpoints but also by the historical development of universalism, which has been shaped by a myriad of political, societal, and economic changes (Anttonen and Sipilä 2014; Garcia-Fuente 2021). This preservation of the divide is aptly observed by Van Kersbergen and Vis (2015) who argue that welfare state

---

[1] This abbreviation refers to 'Korpi and Palme' and is adopted from *Paradoxes of Social Policy: Welfare Transfers, Relative Poverty, and Redistribution Preferences* by David Brady and Amie Bostic in 2015.

scholarship has fallen into a pattern of normal science, where research activity primarily confirms existing theories instead of discovering new paradigms or fundamentally questioning the established ones. They suggest the need to revitalize the field by focusing on social outcomes produced by different welfare regimes, aiming to break away from this normal science approach. Their recommendations align well with the objectives of this M.A. thesis, which works to address the need for new approaches and a greater emphasis on the study of social outcomes.

2.1.1　　In Favour of Targeting

On the selective side of this divide, these scholars envision a 'targeted road' to building an adept welfare state. Proponents of targeting often validate their position either by: (1) asserting that universal policies divert attention away from the poorest among us (Lee 1987); (2) framing universalism as economically unviable (Gilbert 2002); or (3) attributing a moral evaluation of deservingness to the allocation of social assistance (Mead 2003).

Focusing on the first of these claims, a frequent justification for rejecting universal policies boils down to the belief that such policies fall short in effectively addressing poverty. In line with this view, prioritizing universalism may unintentionally result in the neglect of those in the greatest need. Assuming funding for social benefits is a finite resource, spreading benefits too thin coincides with a net decrease in resources allocated to the poor (Greenstein 1991; Devereaux 2016). Essentially, universal policies are thought to result in a sort of 'leakage', where benefits provided to middle class individuals are viewed as an unnecessary and wasteful use of funding. Lindbeck (1988) provides support for this perspective, noting that targeting aims to specifically remove barriers for certain disadvantaged groups.

A prominent concept that has emerged from this argument is known as an "ordeal mechanism". Alatas et al. (2013) describe this mechanism as a strategy that imposes differential costs on rich and poor, potentially improving the precision of targeting. In their field experiment in Indonesia, an ordeal mechanism was introduced where villagers had to apply for aid and take an asset test at a central application site. This increased the

likelihood of the poor applying for aid, resulting in a substantially poorer group of beneficiaries (Alatas et al. 2013). As a result, when making decisions based on the allocation of scarce resources, the most effective approach is to channel benefits according to a strict needs-basis (Besley 1990; Squire 1993). From this perspective, while the welfare state may not be *universally* generous, higher levels of social spending are made possible for a subset of the population through the savings incurred by excluding other groups from receiving social benefits. However, Alatas et al. (2013) also noted that simply increasing the ordeal did not necessarily improve targeting, as both rich and poor were deterred from applying, suggesting a more complex interplay between the ordeal mechanism and socioeconomic status than previously assumed. This point was also highlighted by Besley and Coate (1992), suggesting that welfare programs should carefully implement "ordeal" mechanisms as small participation costs could potentially dissuade the poor from partaking (Besley and Coate 1992). Put differently, by engaging in strategic targeting the welfare state is thought to achieve the 'biggest bang' for its buck.

Another common pro-targeting argument found within both the literature and public discourse is the sentiment that universally orientated social policies foster a culture of dependency and erode individual work ethic (Murray 1984; Lindbeck 1994). This assumption has led some to advocate for forms of targeting designed to incentivize employment, such as lower income support for the unemployed compared to low-income and part-time workers (Saez 2006). Without the imposition of some form of targeting, proponents argue that overly generous and inclusive social policies inevitably promote undesirable economic conditions that undermine their own sustainability. Embedded in arguments such as this is the assumption that generous social spending and robust welfare programs promise long term feasibility only when designed for a minority, and consequently infeasible when carried out for the majority.

While these two pro-targeting arguments are primarily concerned with economic viability and efficiency, a third justification fixates on normative assumptions and the potential adulteration of societal values brought on by indiscriminate social assistance. Narratives such as this stem from the longstanding tendency to divide individuals into the

'deserving' and 'undeserving' poor (Katz 2013). While those in the former category are viewed as legitimately in need due to circumstances beyond their control, the latter group is deemed personally responsible for their unfortunate circumstances. From this perspective, targeting serves as a device for ensuring those who are impoverished due to their own 'avoidable' life choices are not unjustly rewarded for their individual deficiencies with generous social assistance (Banfield 1974; Mead 2003). Adding on, Busemeyer and Iversen (2020) explore the deservingness debate, drawing a line between those who are deserving, such as those who receive pensions, and the undeserving, like the unemployed. They note that this vantage point assumes that targeted social programs ought to be adopted over universal ones, due to the belief that social assistance – which caters to the 'undeserving' – will be perceived as unfair and ultimately fail to sustain public support.

In sum, *selective* social policies are viewed as a superior strategy for designing successful welfare institutions that effectively fulfill the desired functions of the welfare state, namely by prioritizing a healthy economy and reserving social assistance for 'legitimate' recipients. Moreover, they point out that the rise of private alternatives in many fields of welfare state activity have undermined the supporting coalitions for sustaining a public, universal welfare state model. They posit that the universalist model of the welfare state is challenged by the rise of these private alternatives, leading to a loss of support from higher-income groups (Busemeyer and Iversen 2020). Further, Sumino (2018) raises the issue of how targeted spending can potentially amplify attitudinal polarization between different income groups.

2.1.2   In Favour of Universalism

In contrast to pro-targeting scholars, others assert that universalism *is* in fact compatible with a robust and effective welfare state, and perhaps even a prerequisite to achieving desirable welfare state outcomes. I argue that the case for universalism – alongside the value of undertaking further investigation of the potential positive effects of universal policies – can be advanced with reference to three main themes from the literature. The upper hand of universalist policies can be conveyed by: (1) the ability of universalist institutions to enhance social solidarity and serve as a building block for cumulative

welfare spending gains (Korpi and Palme 1998; Brooks and Manza 2007); (2) the lower administrative costs associated with universal policies (Sefton 2006); and (3) universalism's capacity to satisfy the societal desire for a just perception of social structures (Rothstein 2001; Hannah, Brown, and Gibbons 2020).

The first body of literature supporting universalism as an avenue for achieving a desirable welfare state centres around the ability of universalist institutions to cultivate a political climate that is conducive with achieving generous *and* expandable welfare state efforts. This capacity has been explained by scholars through the combination of two related processes known as 'policy feedback' and 'solidarity effects.' Some studying the effects of 'policy feedback' argue that political institutions shape attitudes and preferences, and not the other way around (Pierson 1993). Accordingly, after the initial establishment of political institutions, the preferences of individuals and groups that operate and make decisions within them become influenced in accordance with the interests embodied by said institutions (Skocpol 1992). Therefore, societies are susceptible to continual processes in which public opinion is refracted by institutional forces, consequently shaping the motivations that inform the further evolution of institutions (Pierson 1993; Garland 2016). The composition of a society's welfare state institutions influences the playing field in which political activities are conducted, and as a result, also plays a decisive role in shaping the nature of alliances and antagonisms that characterize a populace (Pierson 1994; 1996). Building on this, Greve (2022) posits that universal benefits including middle class can lead to larger social budget and therefore greater ability to reduce poverty; the welfare state has potential to act as a shared and cheaper risk solution as opposed to individual insurance. Following this logic, one can begin to imagine how the presence of universally orientated political institutions may work to promote a public interest aligning with fostering generous social programs. Thus, universalist political institutions nurture collective support for a strong welfare state presence, alongside providing leverage for future expansions of the size and scope of welfare (Nelson 2004). Anttonen and Sipilä (2014) add an important consideration that quality of services is crucial to universalism; if lower-income groups are unable to access high-quality services, they may resort to private alternatives.

Complementing the above arguments, literature on 'solidarity effects' finds that universal welfare states are often larger and generate more desirable outcomes, due in part to their broad appeal. By administering social programs that treat all individuals the same, the welfare state emphasizes similarities; effectively translating into *united* rather than divided political sentiment (Korpi and Palme 1998; Rothstein 1998). Universal policies cater to the interests of the majority by design, generating much greater support than policies that only cater to specific groups. Echoing these thoughts, KP (1998) add that the Achilles' heel of targeting as a means of fostering generosity is its inability to hold the middle-class support necessary for the maintenance of the welfare system. Without the largescale support of constituents motivated by their common interest in universally applicable social policies, the political will necessary to secure generous social benefits fails to materialize. Lindbeck (1988) furthers this argument by suggesting that universal policies promote positive freedom – that being, freedom to act – by providing resources to everyone, further consolidating the core principles of equality and solidarity inherent in universalism.

Turning to financial considerations, the operational structure of universal welfare programs is often argued as more cost effective than the requirements inherent to targeted programs (Brady and Bostic 2015; Greve 2020). Universal programs incur a certain amount of savings given that they are much less reliant on monitoring. Targeted programs involve continual oversight by nature and are often associated with strict and complex administrative requirements (Besley and Kanbur 1991; Bradshaw 2012). Universal policies avoid the cumbersome process of verifying an individual's eligibility for a given benefit by establishing inclusivity from the start. As a result, social savings are incurred as additional social spending towards administrative oversight is avoided, thereby leaving more funding available for overall distribution. In this context, Lindbeck, Nyberg, and Weibull (2003) offer an additional perspective, asserting that universal policies, coupled with social norms against free riding, can help moderate the usage of income-insurance systems.

Last, the literature has pointed to universalism as being conducive with conceptions of justice, maximizing the support and potential for welfare state expansion, alongside

enhancing individual flourishing in a variety of ways (Barrientos 2019; Byun 2022). This evidence highlighting universalism's high potential for establishing a just and prosperous society merits further investigation on the effects of universalism on welfare state outcomes.

2.2 A Universally Ambiguous Debate: What Defines an 'Effective' Welfare State?

Proponents for both targeting and universalism claim that theirs is the more effective strategy – which leads to the question, "what does it mean for a welfare state to be effective?" Thus far, the discussed literature has dealt with evaluating differing strategies for *achieving* an optimal welfare state but has not approached the process of determining the *qualities* characterizing the ideal welfare state with the same level of enthusiasm. Should we design our analyses under the assumption that an effective welfare state is one that best alleviates poverty? Or should a welfare states' redistributive capacity serve as the primary outcome guiding our valuation of targeting and universalism? This leads to a bigger question about the impact of economic shifts, such as liberalization and privatization, since the 1990s. These changes have redefined societal structures and thus challenged traditional universal service provision (Anttonen and Sipilä 2014). With these issues in mind, I argue that two discrepancies are generated within the targeting-universalism literature caused by insufficient attention to the 'dependent variable' side of analyses. The first problem involves the lack of diversity on how welfare state strategies are evaluated, and the second deals with directly addressing how the *outcomes* of universalism may often prove just as inconsistent as the *effects* of universalism.

Beginning with the issue of outcome diversity, several works provide evidence of one type of outcome receiving priority over others, with poverty or inequality reduction being the suitable answer to the question of what counts most in defining an effective welfare state (Korpi and Palme 1998; Guillaud, Olckers, and Zemmour 2020; Byun 2022). This is further substantiated by findings from Marx, Salanauskaite, and Verbist (2016), who reveal that some systems that include means-tested benefits have been highly redistributive. This fixation with framing the targeting-universalism debate narrowly on poverty and inequality outcomes may be attributed, at least partially, to the extent of KP's

influence. As previously discussed, much of the recent targeting-universalism literature has maintained the relevancy and development of the debate by drawing upon KP's blueprint (Kenworthy 2011; Marx, Salanauskaite, and Verbist 2013; Brady and Bostic 2015; Jacques and Nöel 2018). Ultimately, this tendency of recent targeting-universalism scholars has resulted in directing substantially less attention towards assessing the respective impact of targeting and universalism against both outcomes that stray from the familiar territory of poverty and inequality. To address this issue, Scruggs and Tafoya (2022) stress the importance of looking at the entire range of policy dimensions, not just expenditures, when evaluating welfare generosity. Despite the undeniable importance of factoring endeavours such as poverty reduction when evaluating an effective welfare state, I argue that the set of standards involved in judging the quality of welfare states ought to be much broader than what is often seen within recent works on welfare state effectiveness. Expanding the outcomes included in studying the role of universalism is one way of answering the call for a reprioritization of the effects and outcomes of welfare state policies that we see made by some welfare state scholars (Hannah, Brown, and Gibbons 2020). In carrying out restrictive analyses, scholars run the risk of minimizing the capabilities and functions the welfare state has to offer, alongside stifling the potential of one's own work for contributing to our understanding of how the welfare state functions (Baldini et al. 2016). Adding to this, Lindh and McCall (2022) suggest future research should delve into public opinion about specific labor market policies and expand to include a wider set of countries to better understand different societal preferences.

The second dependent variable issue is that findings on the outcomes of universalism are as much of a mixed bag as the different ways universalism has been measured across the literature. Put simply, the literature varies both in terms of what measures of universalism are *composed* of, alongside the possible outcomes that result from imposing universalism. This notion is reinforced by Huber and Stephens' (2001) research, which finds varying distributive outcomes and poverty levels associated with different welfare state regimes. For example, findings from Brady and Bostic (2015) are consistent with findings from KP (1998) on social transfers, but inconsistent with KP's findings on poverty. Adding to this complexity, Marchal and van Lancker (2019) emphasize the need for more approaches to be created that measure targeting outcomes rather than just targeting

design. Upon considering the potential for inconsistencies across *both* the independent and dependent variables used in welfare state studies, the more apparent the intricacies that characterize the path to resolving the targeting-universalism debate become. This is underscored by Sánchez and Goda (2018) who demonstrate that corruption can distort the demand for and supply of redistributive policies, affecting the effectiveness of a welfare state. The subsequent sections will further elaborate on these complexities and how they can influence the effectiveness of welfare policies.

2.3   The Dilemma of Methodological Narrowness

The final and perhaps most salient factor contributing to the persistent disagreement over the targeting-universalism question is rooted in two methodological shortcomings: (1) the use of *two* main approaches to measuring universalism, where the potential implications of doing so either go unacknowledged or are simply treated as interchangeable options for capturing the same concept, and (2) the tendency of scholars to test only one measure of universalism when drawing conclusions involving the comparison of targeting and universalism on a given outcome. Before assessing these two methodological issues, it is worth noting how these concerns over measurement begin at the stage of conceptualization. Naturally, how one conceptualizes a phenomenon influences how a given concept is transformed into a measurable variable. Despite the scarcity of a clear and comprehensive definition of universalism, two common approaches to capturing the general nature of the concept can be discerned from the literature. Interestingly, Blomqvist and Palme (2020) have proposed a measure of universalism that addresses these shortcomings by including four different dimensions. Despite being an improvement, it's still a singular measure, defined and tested in a specific way.

2.3.1   Conceptualizing Universalism

The conceptual history of universalism can be organized along two lines: (1) the nature of universalism, and (2) the content subsumed within the interpretation of universalism. The first concern centres on the question: "What is the conceptual terrain in which universalism takes shape?". To answer this, two overarching conceptual frameworks are suggested: (1) a spectrum approach and (2) a distinct concept approach.

I refer to the first framing of universalism as the 'spectrum approach'. This approach treats universalism as an extreme on one end of a spectrum, with targeting assuming the opposite end. KP (1998) popularized the treatment of universalism as a spectrum, where universalism was identified by the extent to which social transfers were concentrated – or targeted – within a population. Following this, the more targeted social benefits are, the further from universalism they stand (Nelson 2007). Similar conceptions of universalism as a single point along a spectrum remain a widely accepted standard (Kenworthy 2011; Marx, Salanauskaite, and Verbist 2013). The main consequence corresponding with the spectrum approach is that reference to another concept is always required for defining universalism.

The second strategy for conceptualizing universalism I refer to as the 'distinct concept approach'. In contrast to the scholars who conform to the spectrum framework are those academics that advance universalism as a concept that can – and ought to – be defined independently of any reference or contrast to related but external concepts. Put differently, imagine a remote uninhabited island with a self-sustaining ecosystem, untouched by, or reliant, on any foreign species. Our imaginary island represents universalism itself, while the absence of foreign influence on the nature and functioning of the island's ecosystem symbolizes how the absence of targeting (in this scenario, playing the role of foreign influence) is both inconsequential for fully defining and understanding the concept of universalism. This hypothetical scenario provides not only a straightforward explanation of the conceptualization approach from the 'distinct concept' vantage point, but also works to further reinforce the fundamental differences guiding these two conceptual approaches.

Following this sentiment, scholars of this opinion reject the notion of tethering universalism to a secondary concept in the form of a spectrum serving as a prerequisite for pin-pointing the definition and successful assignment of meaning conferred upon the term universalism. Opponents of the spectrum approach rebuff this conceptual approach first on the belief that it is unnecessary – with some scholars advancing independently conceptualized alternatives of universalism (Brady and Bostic 2015). Relatedly, rather than framing targeting as universalism's opposite, the inverse of universalism is instead

understood as residualism (Jacques and Nöel 2021). For example, universalism can be conceptualized as homogeneity with respect to social benefits (Brady and Bostic 2015). According to Brady and Bostic's proposed conceptualization, the more uniform benefits and their corresponding eligibility and coverage characteristics are for the entire population, the more universal a welfare state becomes (2015). Conversely, when benefits differ across different groups, and eligibility requirements and coverage rates vary, a welfare state no longer meets the conceptual requirements to be considered universal (Brady and Bostic 2015). Conceptualizing universalism in this manner invalidates arguments insistent on the necessity of defining universalism through reference to targeting (Brady and Bostic 2015). To summarize, the distinct concept approach assumes the straightforward view that welfare programs can be identified as simply more or less universal.

A second important, and perhaps more consequential case to be made against the use of the spectrum approach, is the argument that such an approach to conceptualizing universalism may induce a chain-reaction with the potential to contaminate the operationalization of universalism (Jacques and Nöel 2018). Such contamination – which will be discussed in greater detail below – may run the risk of constructing and centering analyses around indicators that fail to measure what they are intended to measure, exposing researchers to the threat of tainted findings (Stiglitz, Sen, and Fitoussi 2009; Marchal and Van Lancker 2019; Lindh and McCall 2022; Scruggs and Tafoya 2022). Overall, it is important to remember that the concept of universalism remains complex and multidimensional, making it a challenging idea to capture in a single, comprehensive conceptualization. As a final note on conceptualization, it is worth acknowledging the growing emphasis on individualism within social policies; this entails a shift from universal "one-size-fits-all" solutions towards an increasing demand for personalised services (Anttonen and Sipilä 2014).

2.3.2   Operationalizing Universalism

Importantly, these two competing conceptual perspectives also come with distinct operational strategies, thus perpetuating a body of literature characterized by both conceptual and methodological division. I assess the methodological aspect of this divide

by advancing that current measures of universalism can be grouped into two clusters: (1) those focused on outcomes (Korpi and Palme 1998; Kenworthy 2011; Marx, Salanauskaite, and Verbist 2013), and (2) those concerned with institutional characteristics (Jacques and Nöel 2018; Rasmussen and Knutsen 2021).

First, measures that prioritize outcomes can be defined as those that operationalize universalism by capturing the endgame of universalist policy – that is, through the resulting distribution of benefits to varying segments of the population. Essentially, the concentration of social benefits serves as a proxy for universalism rather than accounting for other potentially relevant characteristics. This approach assumes that the greater the diffusion of social transfers across the entirety of a population is, the more universal a welfare state becomes. Operationalizing universalism, however, is a challenge given that the range of choices available to individuals could be limited by welfare state policies (Lindbeck, 1988). This makes the task of defining and measuring universalism quite complex. This complexity is further highlighted by Lindbeck, Nyberg, and Weibull's (2003) suggestion that social norms and the number of beneficiaries could influence the effectiveness of universal policies versus targeted approaches.

Other scholars have approached operationalizing universalism through an emphasis on the institutional design elements that precede spending and transfer outcomes. Examples of indicators following this approach include measures that examine the extent of means-testing or private services (Jacques and Nöel 2018), or indicators that evaluate the frequency and magnitude of social programs available to different sub-groups of the population (Rasmussen and Knutsen 2021). In a related vein, Huber and Stephens' (2001) analysis of the Social Democratic Welfare State regime illustrates the practical operation of universal policies in reducing wage dispersion and promoting a high-wage, high–labor productivity competitive niche. Additionally, Esping-Andersen (2015) provides a vivid illustration of operationalizing universalism, showing how Scandinavian welfare regimes have achieved significant equalization effects in educational attainment.

2.3.3   A New Targeting-Universalism Paradox? The Ubiquitous yet Confined Debate

A crucial insight regarding these methodological patterns is that scholars who conform to a spectrum-based view of universalism are often restricted – whether knowingly or unknowingly – in terms of the methodological camp they work within. I suggest that this occurs because once accepting that targeting and universalism are linked in such a way that the increasing presence of one corresponds with a decrease in the other, measuring universalism according to 'outcomes' achieved in the form of an index of concentration appears as most logical and operationally-sound.

Indeed, scholars following the 'distinct concept approach' are naturally oriented towards the measurement styles of the institutional camp. Emphasizing diverse institutional features of welfare states, these researchers often generate innovative measures of universalism (Brady and Bostic 2015; Jacques and Nöel 2018). This flexibility can be attributed to their perception of universalism and targeting as not necessarily existing on a singular spectrum, thus allowing a wide range of possibilities for measuring universalism backed by a solid conceptual backbone.

Yet, the division of the literature according to the two presented strategies for measuring universalism leads to the issue of reduced comparability between studies. When one study uses a concentration index to measure universalism while another uses an institutional measure, it becomes difficult for one to come to any conclusions when the causal mechanisms may vary or be disrupted depending on the type of universalism measure adopted. Additionally, the conventional practice of only using a single measure of universalism in one's study further obscures conclusions on the targeting-universalism debate. This methodological choice dominates the literature, with the exception of a few prominent studies that have incorporated multiple indicators of universalism (Brady and Burroway 2012; Rasmussen and Knutsen 2021). However, even within such exceptions, no more than two universalism measures are included. Given this pattern of scarce operational diversity, I contend that the line of literature sparked by the targeting-universalism debate has now found itself in a phase of stagnation; one that I believe can

be overcome through the incorporation of the untapped potential of conducting analyses that prioritize greater attention to measurement issues.

This stalemate within the literature has led to what could be considered a new 'paradox' in the targeting-universalism debate. In their critique of KP's index, Gugushvili and Laenen (2021) highlight the need for more empirical rigor, arguing that the concentration index as a measure of redistribution is influenced by demographic factors and pre-transfer incomes. They suggest that these factors complicate the operationalization and interpretation of measures. This critique further extends the call for operational diversity and highlights the need for a more rigorous and involved understanding of universalism.

2.4   Back to the Drawing Board

Clearly, the longstanding debate over conceiving the most effective and efficient welfare state continues to hold considerable weight, both within academic circles as well as public discourse. This profound and complex question prompts us to reevaluate the core concepts of welfare state studies. Accordingly, we must continuously refine and reassess our understanding of fundamental constructs like universalism and targeting. Although targeting and universalism can be compared for the purpose of answering a variety of questions about the functioning of the welfare state, the heart of the debate concerns unveiling the association between a welfare state's strategic logic and the extent of its generosity. Upon investigating whether universal welfare states fare better than selective systems in terms of output levels, I measure the concept of generosity in terms of social spending. As a straightforward and widely accessible indicator, studying the potential association between universalism and social spending allows one to cut straight to the foundation of the targeting-universalism question. Therefore, with the continuing salience of this debate, my first research question asks: *"Are welfare states that are more universalistic associated with higher level of social spending?"*

After working to untangle the broader relationship between universalism and generosity, universalism's function within the welfare state can be further broken down through a comprehensive analysis of alternative outcomes. This is reinforced by Marx, Salanauskaite, and Verbist (2016), who conclude that for a system to have a strong

redistributive impact, it needs to cater to a broad segment of the electorate but also incorporate effective redistribution towards the poor. Recognizing that many welfare state outcomes beyond social spending may be implicated within the role of universalism, and that variety in terms of welfare outcomes present a further point of contention over universalism's merit, my second research question asks: *"How important is universalism for other welfare state outcomes, namely 'social transfers', 'poverty', and 'inequality'?"*

Through the process of analyzing the development of the welfare state scholarship, an interesting pattern emerges with respect to how the attempts at resolving the debate are approached. I propose that much like KP's paradox of social policy – where the more we target social benefits to the poor the less effective we are at ameliorating poverty and inequality – the discipline may currently be held back by a 'targeting-universalism debate' paradox of its own creation. I argue that the more prevalent the concept of universalism and the debates over it have become, the more scholars seem to stray from concern over truly defining universalism and its distinct characteristics. Such a phenomenon then influences the extent to which alternative forms of measuring universalism are considered in analyses, leaving substantial unexplored terrain for advancing these debates. Given the variation between universalism measures and the tendency within welfare state literature to refrain from extending beyond commonly used indicators, my third research question asks: *"How sensitive are the potential relationships between universalism and these welfare state outcomes to the ways in which universalism is operationalized?"*

In closing, the following chapter presents an examination of key aspects related to the concept of universalism in welfare state literature, focusing on its conceptual and operational challenges. The discourse shed light on different perspectives and methods, revealing the need for further rigorous exploration of universalism and its impact on the welfare state. Moving forward, the next chapter, Methodology, will delve into the chosen analytical strategies and research design to address the three posed research questions. Here, the focus will shift from the broad conceptual and operational issues in universalism to the more specific, technical aspects of how this study plans to measure universalism and evaluate its relationship with various welfare state outcomes.

Chapter 3: Methodology

3   Data

To test the relationship between various operationalizations of universalism and welfare state outcomes, I combine six sources of data on 22 OECD countries from 1960 to 2020. OECD social spending data and welfare state data from the Comparative Welfare States (CWS) Dataset are used for both my four dependent variables and welfare state controls. My five independent variables – each coming from a separate data source – contain different measures of universalism. Table 1 summarizes the data sources and country-year availability for each welfare state and universalism measure. Although each of the datasets merged for my analysis vary with respect to countries and years available, my final sample size (N=1,342) is restricted to the 22 countries covered by CWS from 1960–2020.

Table 1: List of 'welfare state outcome' measures (dependent) and 'universalism' measures (independent), including country-year availability and data sources.

Measure	Countries available	Years available	Data sources
Social spending [†]	22	1960-2018	OECD, Social Expenditures Database, 1985. OECD, Social Expenditures Database, 2016.
Inequality [‡]	22	1960-2017	Comparative Welfare State Dataset, 2020. Standardized World Income Inequality Database, Version 7.1, 2018.
Social transfers [§]	22	1960-2017	Comparative Welfare State Dataset, 2020. OECD, Historical Statistics, various years. OECD, National Accounts Statistics (database), 2020.
Poverty	19	1967-2016	Comparative Welfare State Dataset, 2020. Luxembourg Income Study, LIS Database, 2010a. Luxembourg Income Study, LIS Key Figures, 2010b.
VDEM uni	22	1960-2020	Varieties of Democracy Project Dataset, Version 11.1, 2021a.
SPIN uni	21	1960-2015	Social Insurance Entitlements Dataset, 2020.

SPaW uni [#]	10	1960-1999	Social Policy around the World Database, 2016.
J&N uni	20	2000-2011	OECD, Social Expenditures Update, 2014. OECD, Social Expenditures Database, 2016.
KP uni [ǀ]	10	1974-2004	Luxembourg Income Study, LIS Database, 2010a.

[†]Social spending only available starting in 1980 for Luxembourg, Portugal, and Spain; [‡]Inequality only available starting in 1980s for Switzerland, New Zealand, and Luxembourg (1980, 1982, 1985); [§]Social transfers only available until 2011 for Switzerland; [#]Measure only available until 1988 for Switzerland; [ǀ]Data available in 5-year intervals.

Due to variation in country-year availability across universalism measures, sample size varies from my main sample size (N=1,342) depending on the independent variable deployed. As a result, sample sizes across universalism measures range from 117 to 1,168 Country-Years. Given the nature of my analysis – comparing across measures of universalism from different data sources – this variation is an inherent limitation to acknowledge. Nonetheless, my dataset innovates in (1) utilizing the best longitudinal cross-national data available for wealthy countries (CWS 2020; Brady, Huber, and Stephens 2020), and (2) combining four distinct measures of universalism that, based on my in-depth review of the literature, have not yet been applied simultaneously in addressing the targeting-universalism debate and questions over the merit of universalism.

3.1   Dependent Variables

I use four main dependent variables across my three research questions. My first research question uses *'social spending'* as its dependent variable, while research questions two and three use '*social transfers*', '*inequality*', and '*poverty*'. Distinguishing the size and structure of government programs, including *'social spending'* and *'social transfers'*, is crucial when estimating their impact on inequality and poverty (Esping-Andersen 1990). This underscores the importance of delineating 'social spending' and 'social transfers' as separate dependent variables in this research. Additionally, 'inequality' and 'poverty', though often conflated in public discussions, are distinct issues, highlighting the importance of measuring them independently (Esping-Andersen 1990; Bradshaw and Finch 2002).

'*Social spending*' is measured as gross public expenditures as a percentage of GDP. The measure includes data on 22 countries from 1960–2018, sourced from OECD (1985) for the years 1960–1979, and from the OECD Social Expenditures Database for 1980–onward (OECD 2016). Esteban Ortiz-Ospina and Max Roser are responsible for compiling these two data sources into this single measure of social spending. The measure is inclusive of all government spending where, (1) the intentions of benefit provision are socially orientated, and (2) where benefit allocation entails either redistribution or mandatory participation (Lindert 2004; Obinger 2021). It includes programs ranging from basic assistance to poor families, unemployment compensation, public noncontributory pensions, public health expenditures, to housing subsidies (Lindert 2004). The nine policy realms covered by the measure include: health, family, housing, old age, incapacity benefits, survivors' payments, unemployment insurance, active labour market policies, and several other policy programs. Last, acceptable forms of social benefits incorporated in the measure are inclusive of cash, benefits in-kind, and tax expenditures.

'*Social transfers*' is also measured as public expenditures on social benefits as a percentage of GDP and includes data on 22 countries from 1960–2017. However, this measure differs from 'social spending' in that it excludes social transfers in-kind. The variable is inclusive of all government provided cash transfers, such as social assistance and social insurance benefits (Lindert 2004). Given their redistributive nature, these transfers are frequently controversial as they are often targeted at aiding the poor (Lindert 2004). Cash transfers have been identified as more redistributive than the taxes and contributions that finance them (OECD 2015) and are instrumental in impacting income equality and poverty (OECD 2015). This measure, in line with Lindert's (2004) definition, excludes government expenditures on education, which distinguishes it from 'social spending'. These transfers are more controversial due to their redistributive nature, often targeted at aiding the poor (Lindert 2004). Notwithstanding potential controversies surrounding these types of programs, Obinger (2021) underscores the impact of social transfers, particularly cash transfers, as potent redistributive instruments significantly influencing poverty and inequality. This measure, alongside two of my dependent

variables outlined below, is drawn from the CWS data set assembled by David Brady, Evelyne Huber, and John D. Stephens in 2020.

'Inequality' is measured as the post-tax and transfer Gini coefficient of household income. This CWS variable utilizes data from The Standardized World Income Inequality Database (SWIID), compiled by Solt (2016). The decision to use post-tax and transfer data for measuring 'inequality' aligns with the objective of comprehensively capturing the total impact of welfare states, an advantage over pre-tax and transfer studies such as those by Piketty and Saez (2003). The OECD Income Distribution Database provides Gini coefficients for incomes before and after taxes and transfers for a comprehensive comparison across countries (OECD 2015). Labour market factors including earnings inequality, employment and unemployment rates, and labour market institutions also influence income inequality (OECD 2015). Utilizing post-tax and transfer data provides a more accurate reflection of individuals' actual resources by analyzing income distribution *after* the implementation of redistributive policies (Smeeding 2005).

'Poverty' is measured as the post-tax and transfer relative poverty rate for a nation's total population, based on a threshold median income of 50%. The measure includes data on 19 of the 22 countries included in the analysis – excluding Japan, New Zealand, and Portugal – ranging from 1967–2016. The choice of a 50% threshold for the poverty line is consistent with most studies in rich (OECD) countries, despite controversies surrounding poverty line setting (Bradshaw and Finch 2002). This variable was sourced by CWS from the Luxembourg Income Study (2010a; 2010b), which offers access to databases of harmonized income data. As with my 'inequality' measure, the post-tax and transfer approach to 'poverty' helps to better encapsulate the ultimate effects of welfare states' interventions. This methodology reveals the actual poverty rates *after* governmental redistribution, thus better demonstrating how successful different welfare states are in achieving poverty reduction (Kenworthy 2011).

3.2 Independent Variables

Each of my independent variables is a unique indicator used in past literature to measure the concept of universalism. While I include five universalism measures for univariate

and bivariate analyses, only four measures are used in my multivariate analyses. First, *'VDEM universalism'* draws on the 'universal programs' measure from the Varieties of Democracy (V-Dem) Project and is available for the full 1960–2020-time frame for all 22 countries (Coppedge et al. 2021a; 2021b). This extensive timeframe of data makes it an ideal choice in line with Korpi (2010), who emphasizes the importance of longitudinal empirical data when analyzing welfare states. The measure is constructed based on the response of expert coders to the question: "How many welfare programs are means-tested and how many benefit all (or virtually all) members of the polity?" (Coppedge et al. 2021b). 'Universal programs' is based on an initial ordinal scale ranging from 0 to 5; 0 indicating none or very few welfare programs exist, and 5 indicating the majority of welfare programs are universal in nature and minimal means testing is found (Coppedge et al. 2021b). VDEM then converts this measure to an interval scale using their measurement model (Pemstein et al. 2021). For my main sample (N=1,342), the minimum score reported is -1.795 and the maximum score reported is 3.042, with higher values corresponding to higher universalism scores.

Second, *'SPIN universalism'* draws on data from the Social Insurance Entitlements Dataset (SIED), a component of the Social Policy Indicator (SPIN) Database (Nelson et al. 2020). *'SPIN universalism'* measures universalism as the coverage ratio of labour force sickness benefits, calculated by dividing the total number of individuals entitled to state sickness benefits by the total labour force (in hundreds of thousands). Although restricting the operationalization of universalism solely to sickness benefits may illicit skepticism from some, the measure has previously been used as a universalism indicator within recent literature where it was found to be associated with a range of other universalism measures (Rasmussen and Knutsen 2021). Additionally, this measure aligns with the perspective of Öktem (2020), who studies welfare states through the lens of universalism, specifically operationalizing this as old age and sickness through health and pension coverage. This universalism measure covers 21 countries from 1960–2015 and offers a unique perspective in that it prioritizes insurance against the risk of illness for all workers as a core principle of universalism. Since SPIN reports on variables in five-year intervals, I use linear interpolation to smooth these gaps across time and maximize sample size.

In addition to the above *'SPIN universalism'* measure, I also incorporate a secondary version of SPIN universalism made possible by the Social Insurance Entitlements Dataset's additional coverage ratio measures for unemployment insurance and workplace accident coverage. The sickness benefits and unemployment benefits variables are highly correlated at r=0.634, while I found the workplace accident measure to be weakly correlated with its two SPIN counterparts. Thus, my secondary SPIN universalism variable is calculated as the average of an observation's score for sickness benefits and unemployment benefits. As with my primary SPIN universalism measure, this secondary SPIN measure also covers 21 countries from 1960–2015.

Third, *'SPaW universalism'* uses the universalism index from the Social Policy Around the World (SPaW) Database, consisting of data for just 10 countries from 1960–1999 (Rasmussen 2016). This index aggregates universalism scores for the following six major policy programs: old age, maternity benefits, sickness insurance, work accident, unemployment insurance, and family allowance. Disaggregated universalism scores range from 0 to 9; 0 indicating no major program exists, and 9 indicating the program is inclusive of all citizens. Once combined, the universalism index ranges from 0 to 54, with higher values corresponding to higher universalism scores. In my main sample, the minimum score reported is 5.5, while the maximum score reported is 43.

Fourth, *'J&N universalism'* offers the shortest time frame, capturing only the years from 2000 to 2011. The measure was constructed by Jacques and Nöel (2018), using data from both the Social Expenditures Update (OECD 2014) and the Social Expenditures Database (OECD 2016). Using these OECD data, Jacques and Nöel (2018) combined the following two indicators: the percentage of means or income-tested social benefits, and the proportion of total social expenditures accrued to private spending. The universalism index ranges from -2 to 2, with higher scores corresponding to higher levels of universalism.

Last, *'KP universalism'* is the targeting-universalism measure originally used in KP's 1998 analysis. Using income data from the Luxembourg Income Study (2010a), their measure is an index of concentration ranging from -1 (targeted) to 0 (universal). A score

of -1 indicates that all government transfers are funneled towards a single household, while a score of 0 implies that all households receive the same amount of government transfers. This measure was later extended by Kenworthy (2011) up to the mid-2000s, consisting of data for just 10 countries from 1974–2004. Due to an insufficient sample size (N=66), I exclude this measure from the multivariate stage of my analysis.

3.3   Controls

My multivariate models incorporate the following five control variables standard to comparative welfare state analyses: *'population'* measured in thousands; *'share elderly'* measured by dividing population aged 65+ over total population; *'immigration'* measured by the migrant stock as a percentage of population; *'unemployment rate'* measured by the number of unemployed as a percentage of total labour force; and *'GDP'* measured as expenditure-side real GDP at chained PPPs, in millions of 2011 USD.

Including controls for the elderly and unemployed populations are important within the welfare literature, as doing so accounts for needs-based shifts in social spending (Green-Pedersen 2004; Shelton 2007). For example, controlling for *'share elderly'* accounts for changes in old-age dependency over time. The demographic composition of countries at a given time influences social spending levels, as public spending on pensions accounts for a sizeable portion of total social transfers (Jacques and Nöel 2021). Similarly, controlling for *'unemployment rate'* ensures social spending and other welfare state outcomes are not impacted by labour market growth cycles (Esping-Andersen 1990). Last, controlling for *'GDP'* ensures that national differences in economic prosperity are avoided so that some countries will not falsely appear to have higher levels of social spending than others (Roine, Vlachos, and Waldenström 2009). Importantly, these covariates account only for time-varying country characteristics. Capturing time-invariant country characteristics is achieved by applying fixed-effects techniques.

3.4   Analytical Approach

First, I report the unstandardized descriptive statistics for my four 'welfare state'
measures (dependent variables), my five 'universalism' measures (independent
variables), alongside my five control variables used in my multivariate analyses (Table
2).

**Table 2: Descriptive statistics of all dependent, independent, and control variables
(unstandardized).**

	N	Mean	SD
Dependent variables			
Social spending as % of GDP	1212	18.30	6.09
Inequality [†]	1001	28.97	4.16
Social transfers as % of GDP [‡]	1266	12.57	4.02
Poverty [§]	505	9.80	3.71
Independent variables			
VDEM universalism	1342	1.40	.84
SPIN universalism	1122	.69	.33
SPaW universalism	383	28.52	9.45
J&N universalism	240	$8.3 \ (\times 10^{-9})$	1.00
KP universalism	66	-.20	.11
Control variables			
Population (in thousands)	1201	35,265	55,453
Population share of elderly	1200	.13	.03
Immigration [#]	1202	9.83	7.54
Unemployment rate	1276	5.81	4.29
Real GDP [l]	1210	1,025,196	2,165,864

[†]Household Income Gini Coefficient (Post-Tax-and-Transfer); [‡]Excludes social transfers in-kind;
[§]Total population poverty rate based on 50% median-income threshold; [#]Migrant stock as % of
population; [l]Expenditure-side real GDP at chained PPPs, in millions of 2011 USD.

Second, to understand initial trends, I assess time trends by country from 1960–onward
for my four universalism measures (Figure 1), alongside my four welfare state measures
(Figure 2).

Figure 1: Time trends for four universalism measures.
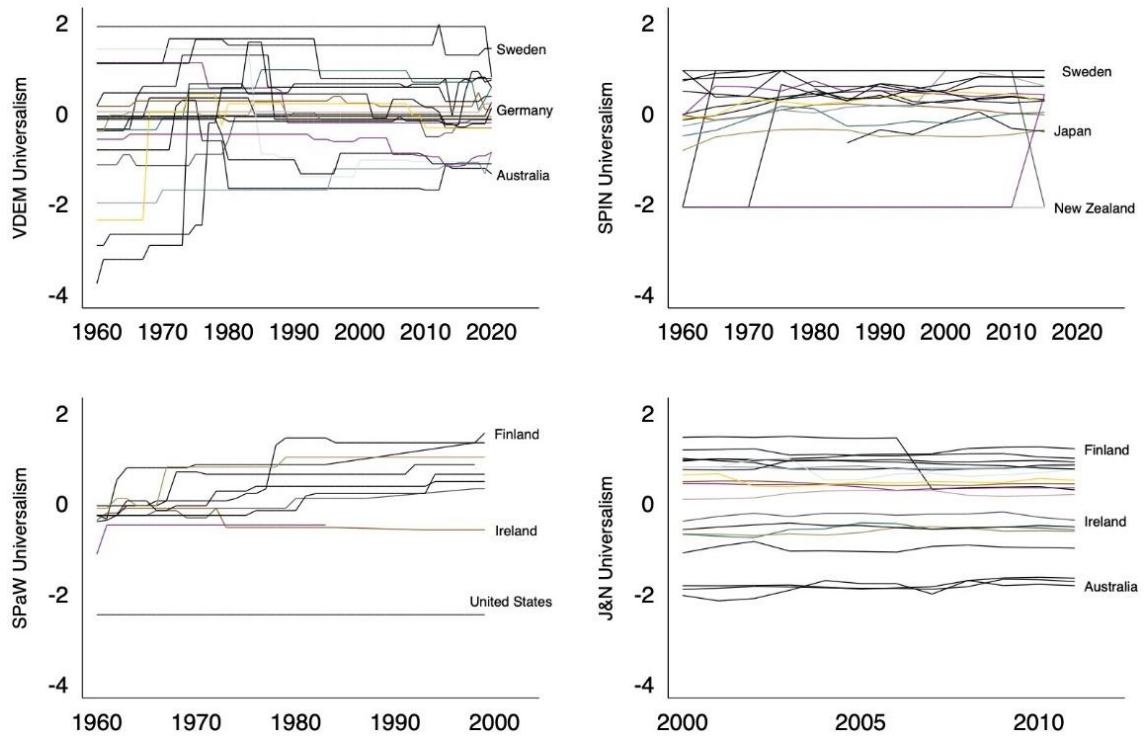
All universalism measures displayed are standardized. Data sources: Michael Coppedge and colleagues (V-Dem Project 2021) used for *VDEM* (1960-2018); Kenneth Nelson and colleagues (SIED 2020) used for *SPIN* (1960-2015); Magnus Rasmussen (SPaW 2016) used for *SPaW* (1960-1999); Olivier Jacques and Alain Noël (OECD 2014; OECD 2016) used for *J&N* (2000-2011). Sweden, Denmark, Finland, Greece, Ireland, and Norway all tie for universalism score in 2015.

Figure 2: Time trends for four welfare state outcome measures.
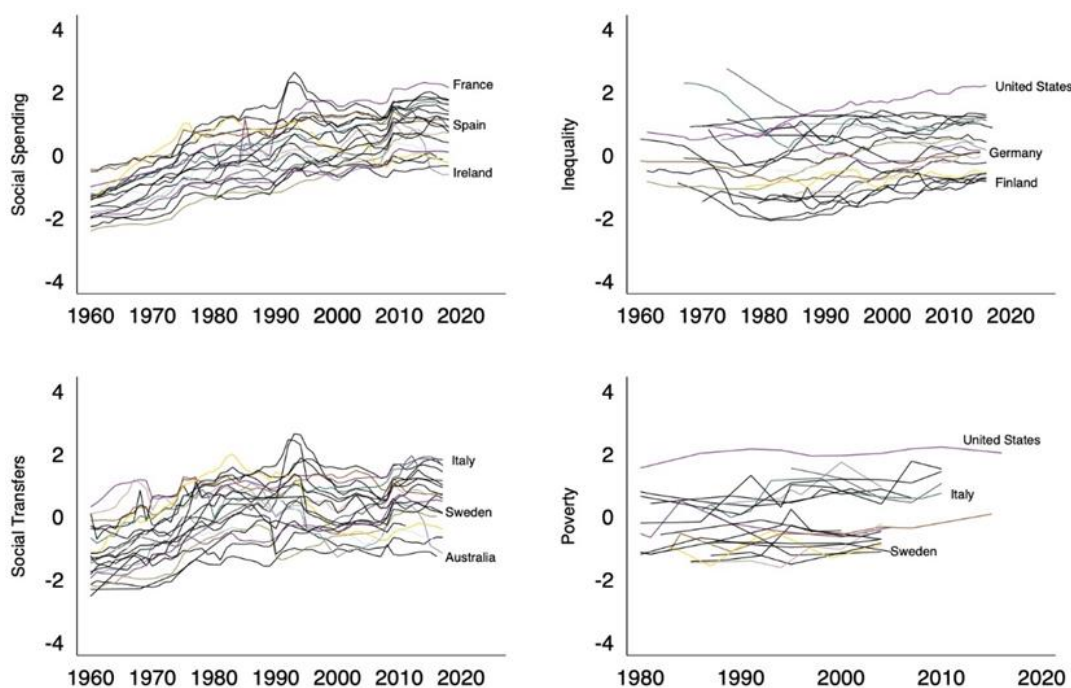
All welfare state outcome measures displayed are standardized. Data sources: Esteban Ortiz-Ospina and Max Roser (OCED 1985; OECD 2016) for *Social Spending* (1960-2018); David Brady, Evelyne Huber, and John D. Stephens using data from Frederick Solt (CWS 2020; SWIID 2018) for *Inequality* (1960-2017), David Brady, Evelyne Huber, and John D. Stephens using data from the OECD (CWS 2020; OECD 2020) for *Social Transfers* (1960-2017); and David Brady, Evelyne Huber, and John D. Stephens using data from the Luxembourg Income Study (CWS 2020; LIS 2010a; 2010b) for *Poverty* (1980-2016).

Prior to plotting trend lines, I standardized all variables for ease of comparison – labelling only the highest, lowest, and median scoring countries for each measure. As highlighted by Smeeding (2005), using comparable data is vital to drawing firm conclusions about policy impacts on inequality and poverty. Therefore, standardizing all variables prior to plotting trend lines is crucial for the validity of the analyses. The use of Luxembourg Income Study (LIS) data to measure 'poverty' aligns with this need for standardization (Gottschalk and Smeeding 1997; 2000; Jäntti and Danziger 2000). The importance of comparable data is further underscored by the analytical approach's reliance on cross-national comparisons to increase variable differences and policy variations (Kenworthy and Pontusson 2005). This approach aligns with calls within welfare state scholarship for a return to clear visual presentations of descriptive data prior to regression analyses

(Kenworthy 2009). I then conducted bivariate analyses to determine the association between social spending and each universalism measure (Figures 3–4), alongside my remaining three measures of welfare state outcomes by each of the five universalism measures (Figure 5). *'VDEM universalism'* offers the longest look at the relationships of interest over time, ranging from 1960–2018. On the other hand*, 'J&N universalism'* captures the shortest timeframe, from 2001–2011. To enhance the clarity of each scatterplot, the plotted values represent *decade* clustered means for each country used in the analysis. Again, I label only a subset of country-decades for ease of interpretation of broader trends.

Third, I use an ordinary least squares (OLS) fixed-effects approach to estimate the potential association of four universalism measures on my four dependent variables. This stage of analysis is central in addressing each of my three research questions. The advantage of incorporating fixed-effects is its ability to eliminate bias stemming from difficult to capture and stable characteristics of cases (Allison 2005; 2009), making country fixed-effects an ideal methodological strategy for cross-national analyses. The fixed-effects models I estimate can be expressed using the following formula:

$$\gamma_{it} = \beta_1 Univ_{it} + \beta_2 Pop_{it} + \beta_3 ShareElderly_{it} + \beta_4 Imm_{it} + \beta_5 Unemp_{it} + \beta_6 Gdp_{it} + \alpha_i + \mu_{it}$$

The presented model can be interpreted as the base model from which all my subsequent models conform. Here, γ_{it} represents one of my four welfare state outcomes (i.e., social spending, social transfers, inequality, or poverty), for country i at time t. *Univ* represents the average effect size of universalism net of the five time-varying controls. The variable α_i represents country-specific fitted constants which allow for the application of fixed-effects. Thus, the effects of time-invariant country characteristics are eliminated, thereby allowing my model to account for all unobservable characteristics unique to a given country. Lastly, μ_{it} represents the error term. This variable varies across both country and time, indicating that any relevant time-varying variables that are not explicitly controlled for in the model will lead to omitted variable bias. For example, without including a

control for GDP, the variable would be in the error term – thereby introducing bias to the findings according to differences in country wealth over time.

Chapter 4: Results

4 Findings

In line with my first research question, bivariate analyses depict a positive correlation between social spending and universalism *across* all five universalism measures (Figures 3–4). However, the strength of this correlation varies largely by universalism indicator, where r=0.24 for *'VDEM universalism'*, r=0.38 for *'SPIN universalism'*, r=0.48 for *'KP universalism'*, r=0.53 for *'SPaW universalism'*, and r=0.74 for *'J&N universalism'*.

Figure 3: Scatterplots illustrating 'social spending' by four measures of universalism.
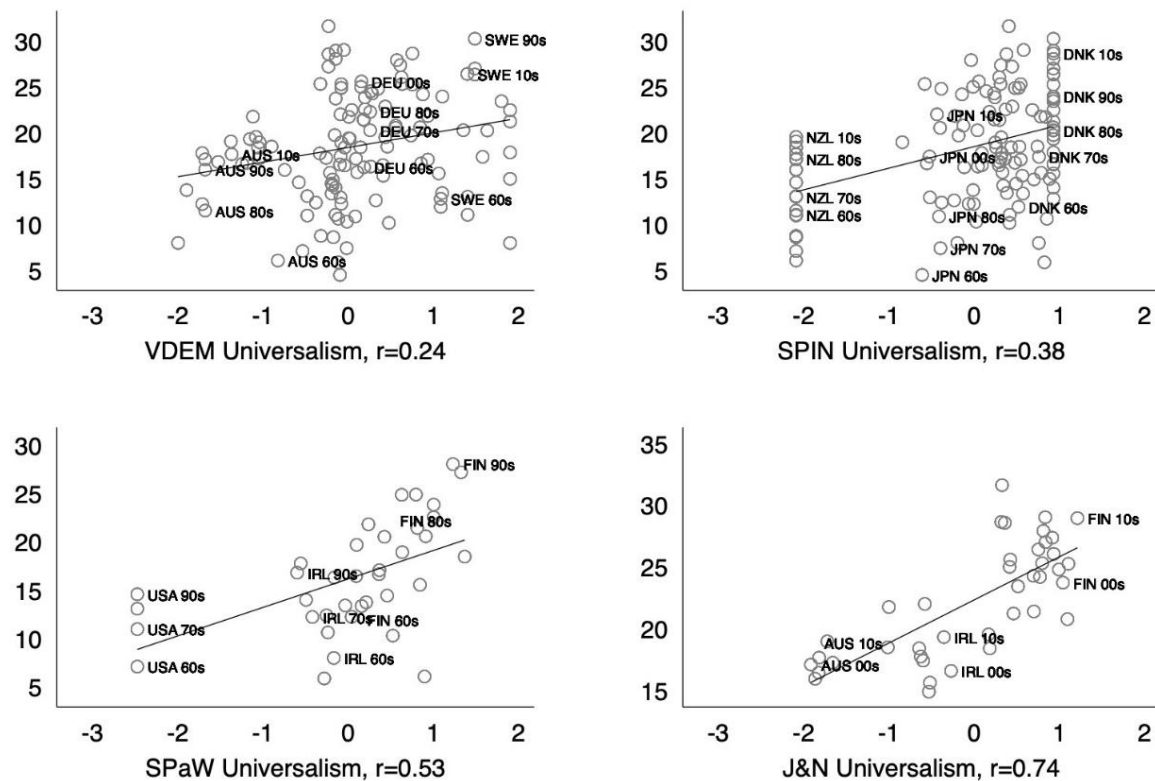
Plotted values represent decade clustered means. Data sources: Esteban Ortiz-Ospina and Max Roser (OCED 1985; OECD 2016) used for *Social Spending*; Michael Coppedge and colleagues (V-Dem Project 2021) used for *VDEM* (1960-2018); Kenneth Nelson and colleagues (SIED 2020) used for *SPIN* (1960-2015); Magnus Rasmussen (SPaW 2016) used for *SPaW* (1960-1999); Olivier Jacques and Alain Noël (OECD 2014; OECD 2016) used for *J&N* (2000-2011).

Figure 4: Scatterplot illustrating 'social spending' by KP universalism.
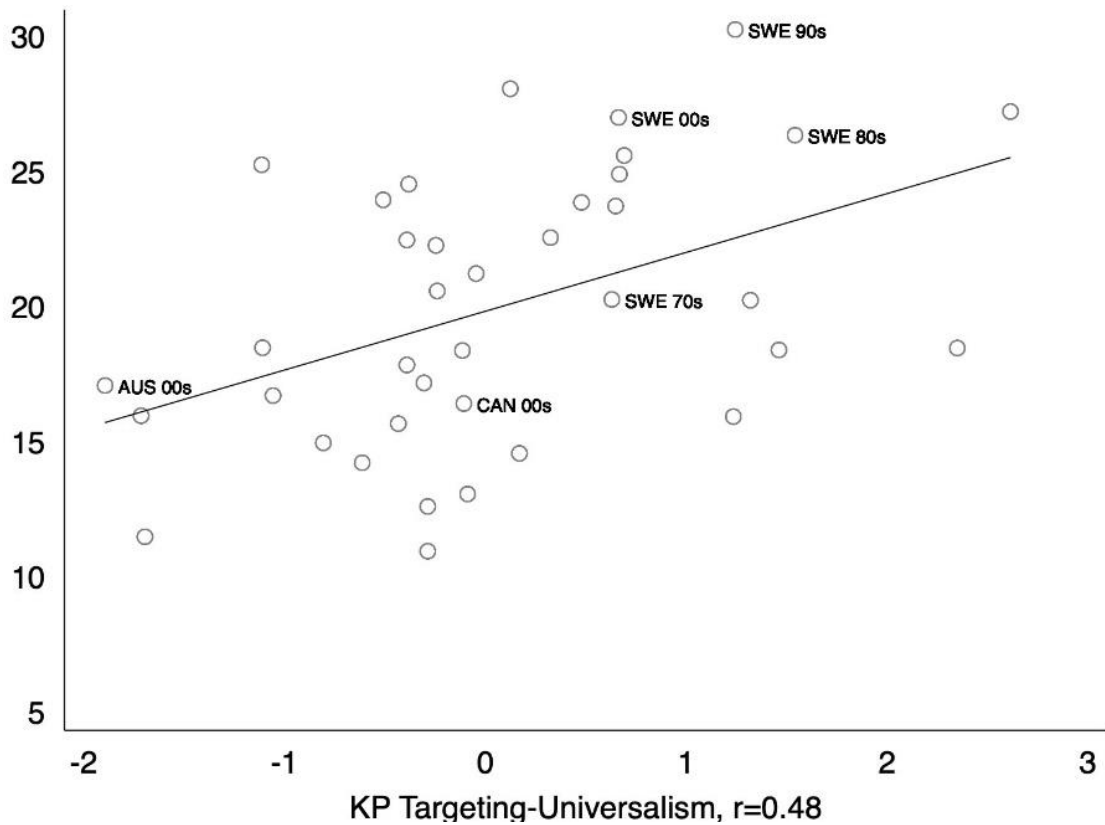
Plotted values represent decade clustered means. Data accessed from: Lane Kenworthy's *Progress for the Poor* (Figure 6.1), using data from LIS 2010a.

While a positive association was observed among all five universalism measures, the strength of the association varies significantly (from r=0.24 to r=0.74), thus hinting at the relevancy of measurement choices to the targeting-universalism debate. The positive association found between social spending and universalism will be expanded upon with the interpretation of results from my fixed-effects models.

Turning to my second research question, Figures 5 through 9 extend Pearson's correlation coefficients to all four welfare state outcomes for each universalism measure.

Figure 5: Welfare state outcome measures by VDEM universalism.

Plotted values represent decade clustered means. Data sources: Esteban Ortiz-Ospina and Max Roser (OCED 1985; OECD 2016) for *Social Spending* (1960-2018); David Brady, Evelyne Huber, and John D. Stephens using data from Frederick Solt (CWS 2020; SWIID 2018) for *Inequality* (1960-2017), David Brady, Evelyne Huber, and John D. Stephens using data from the OECD (CWS 2020; OECD 2020) for *Social Transfers* (1960-2017); and David Brady, Evelyne Huber, and John D. Stephens using data from the Luxembourg Income Study (CWS 2020; LIS 2010a; 2010b) for *Poverty* (1970-2007).

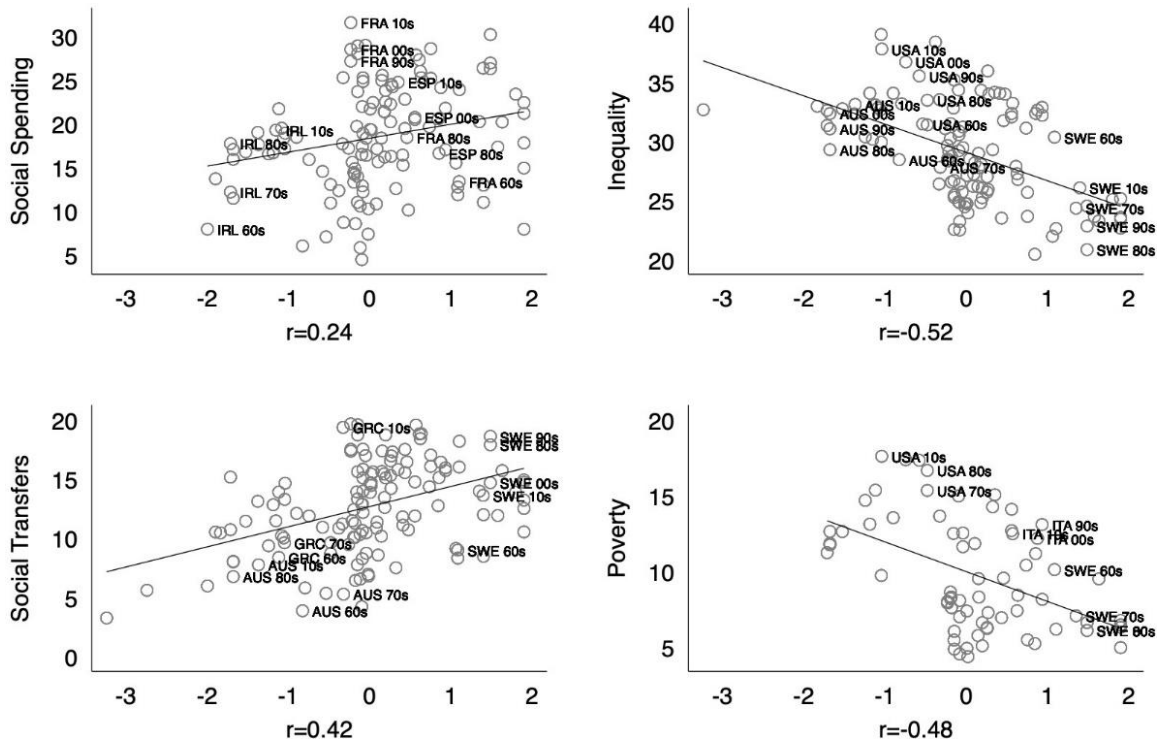**Figure 6: Welfare state outcome measures by KP universalism.**
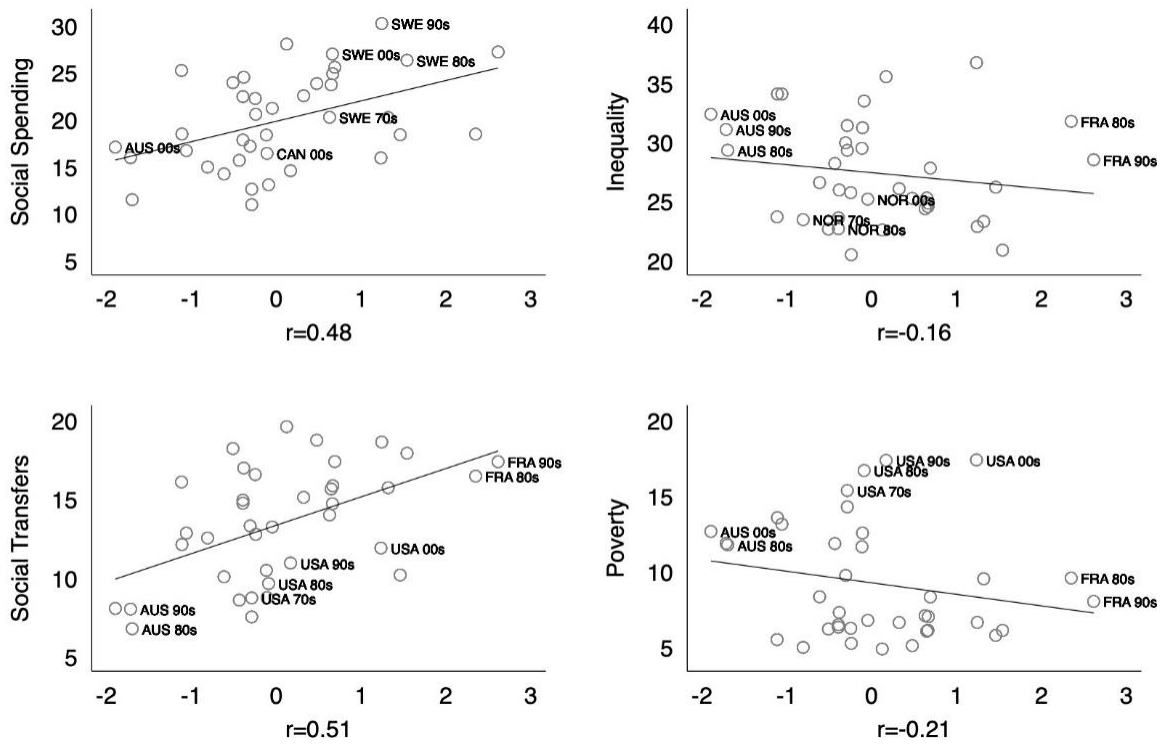
Plotted values represent decade clustered means. Data accessed from: Lane Kenworthy's *Progress for the Poor* (Figure 6.1), using data from LIS 2010a.

Figure 7: Welfare state outcome measures by SPIN universalism.
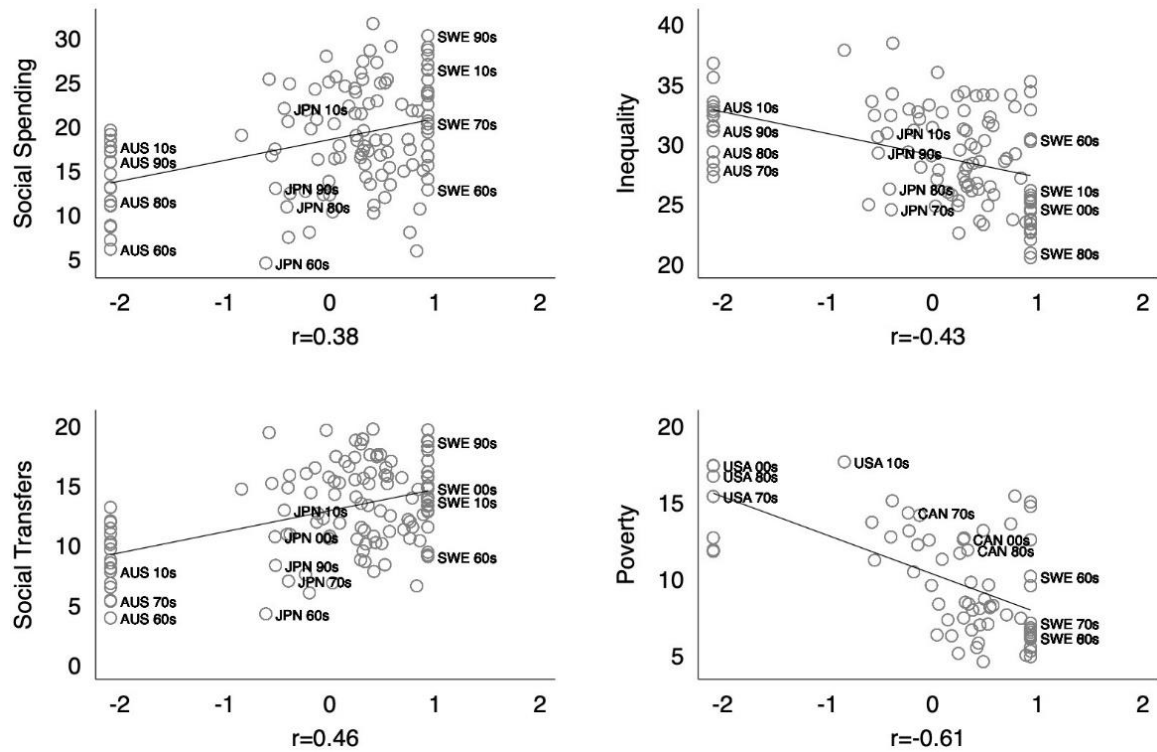
Plotted values represent decade clustered means. Data source: The Social Insurance Entitlements Data Set (SIED 2020), a component of the Social Policy Indicator (SPIN) database, developed by Kenneth Nelson and colleagues in 2020.

Figure 8: Welfare state outcome measures by SPaW universalism.
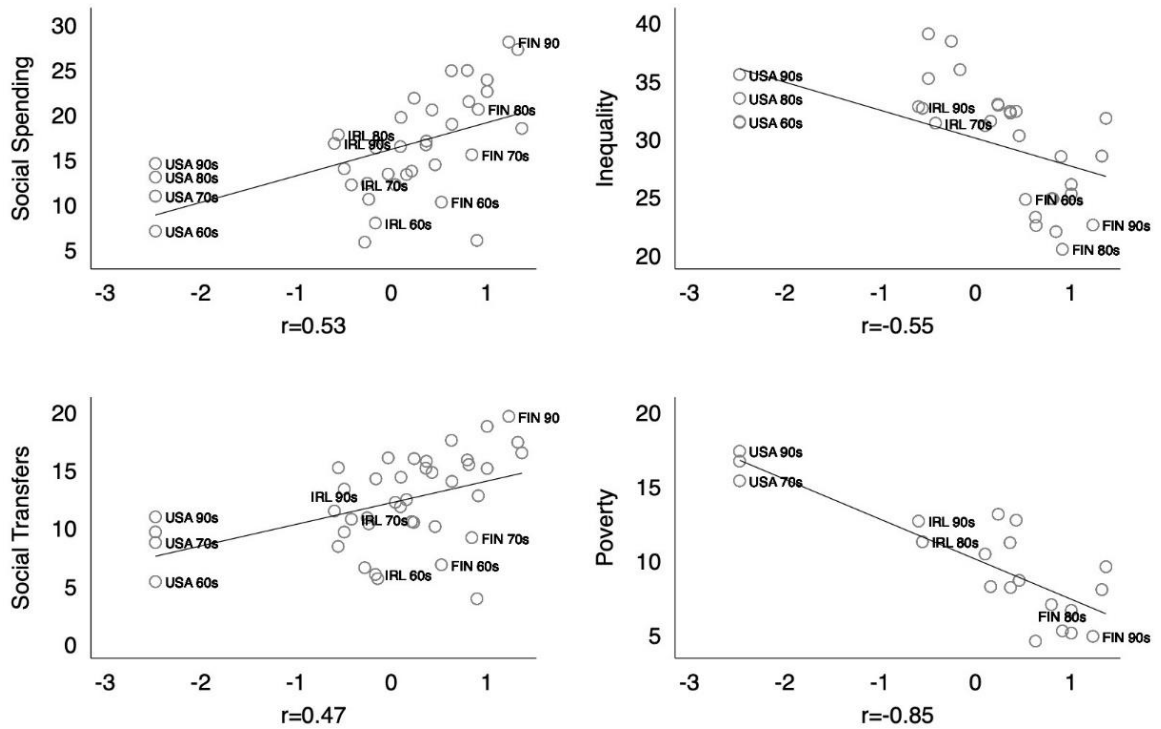
Plotted values represent decade clustered means. Data sources: Social Policy around the World Database (SPaW 2016), developed by Magnus Rasmussen in 2016. Data accessed from: Lane Kenworthy's *Progress for the Poor* (Figure 6.1), using data from LIS 2010a.

Figure 9: Welfare state outcome measures by J&N universalism.
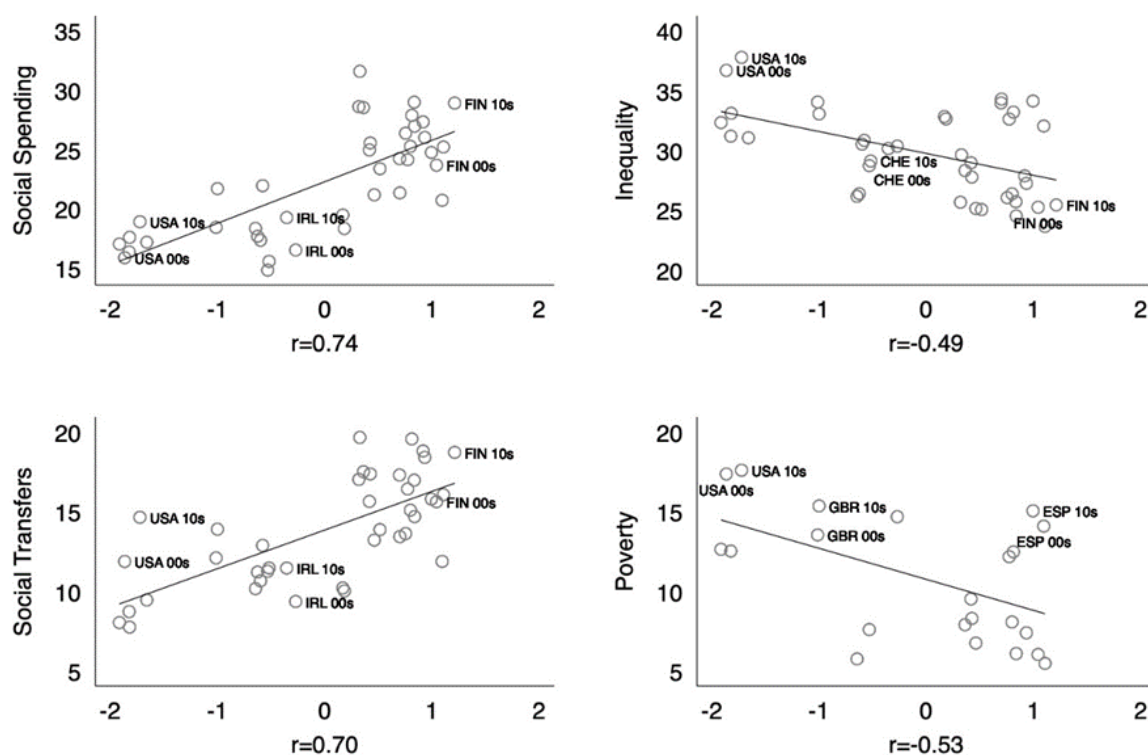
Plotted values represent decade clustered means. Data sources: OECD 2014 (Social Expenditures Update) and OECD 2016 (Social Expenditures Database), developed by Olivier Jacques and Alain Noël in 2018.

Here, I highlight two general trends across universalism measures (Figure 5). First, poverty and inequality have a moderately negative to strongly negative correlation with universalism. The association for poverty across universalism measures is weakest for *'VDEM universalism'* at r=-0.48, and strongest for *'SPIN universalism'* at r=-0.85. Similarly, for inequality, the weakest association found was for *'SPIN universalism'* at r=-0.43, while the strongest was for *'SPaW universalism'* at r=-0.55. Second, as noted above with social spending, social transfers have a moderately positive association with universalism – the weakest association being *'VDEM universalism'* at r=0.42, and the strongest association being *'J&N universalism'* at r=0.70. Notably, the first of the above trends was far less pronounced with the *'KP universalism'* measure (Figure 6). While associations with social spending (r=0.48) and social transfers (r=0.51) did not veer far from the other measures, associations with inequality (r=-0.16) and poverty (r=-0.21) were much weaker than with the other four universalism indicators. These findings

suggest some support for the distinction between 'outcome' and 'institutional' measures of universalism. It may be the case that universalism indicators capable of capturing outcomes have weaker associations with certain welfare state variables than those based on institutional characteristics. However, with the restricted amount of data available, this is merely suggestive. At the bivariate level universalism is associated with increases in social spending and social transfers, alongside decreases in poverty and inequality.

My third research question addresses the measurement sensitivity of the above relationships between universalism and each welfare state outcome directly (excluding *'KP universalism'*), making it the central and most involved component of my analysis. Evidence supporting the case for greater attention to measurement choice emerges in univariate, bivariate, and multivariate findings respectively. Substantial differences in trend lines are observed among universalism measures (Figures 7–9). For example, the VDEM measure shows a steeper increase in universalism in most countries between the years 1970–1980, in contrast to the SPIN measure which remains relatively flat during the same period. Moreover, the country reported as the 'most' and 'least' universal varies across universalism measures. According to V-DEM, in 2020, the most universal country was Sweden while the least universal was Australia. In contrast, the SPIN measure looking at 2015 illustrates New Zealand as the least universal country, while Sweden remains the most universal. Although one could point to the similarities between Australia and New Zealand as a justification for dismissing the importance of these measurement differences, further analysis supports and strengthens this theme.

The bivariate findings reported in relation to my second research question are also relevant when assessing measurement sensitivity. As mentioned above, while the association between social spending and universalism is consistently positive, the strength of these positive associations fluctuates by universalism measure. To put this fluctuation into perspective, using the *'J&N universalism'* measure (strongest correlation, r=0.74), over the *'VDEM universalism'* measure (weakest correlation, r=0.24), significantly strengthens the positive association observed between social spending and universalism – an increase of 0.50.

Multivariate models are estimated by accounting for bias derived from unobservable country-specific characteristics and time-varying factors to further interpret the role and significance of measurement. 16 linear fixed-effects models were constructed, with each model predicting one of the four welfare state outcomes using one of the four universalism measures (Tables 3–4).

Thus, four models are estimated for each universalism measure. I apply the same set of demographic controls to all models but exclude the presentation of coefficients and standard errors for control variables to emphasize findings on measurement differences in universalism. I report both standardized and unstandardized coefficients for each model, as standardized coefficients are necessary for comparing across measures whose original units of analysis vary. The predicted effects of universalism on each of the four welfare state measures are presented using models from VDEM and SPIN data (Table 3), alongside models from SPaW and J&N data (Table 4). Findings from these models highlight two key themes discussed below.

4.1   Universalism and Welfare State Outcomes

Statistically significant variation is observed in the predicted effect of universalism on welfare state outcomes across the four measures of universalism, thus supporting that these relationships *are* sensitive to measurement choice (Figure 10).

Figure 10: Coefficient plot of regression of universalism indicators on welfare state outcome measures.
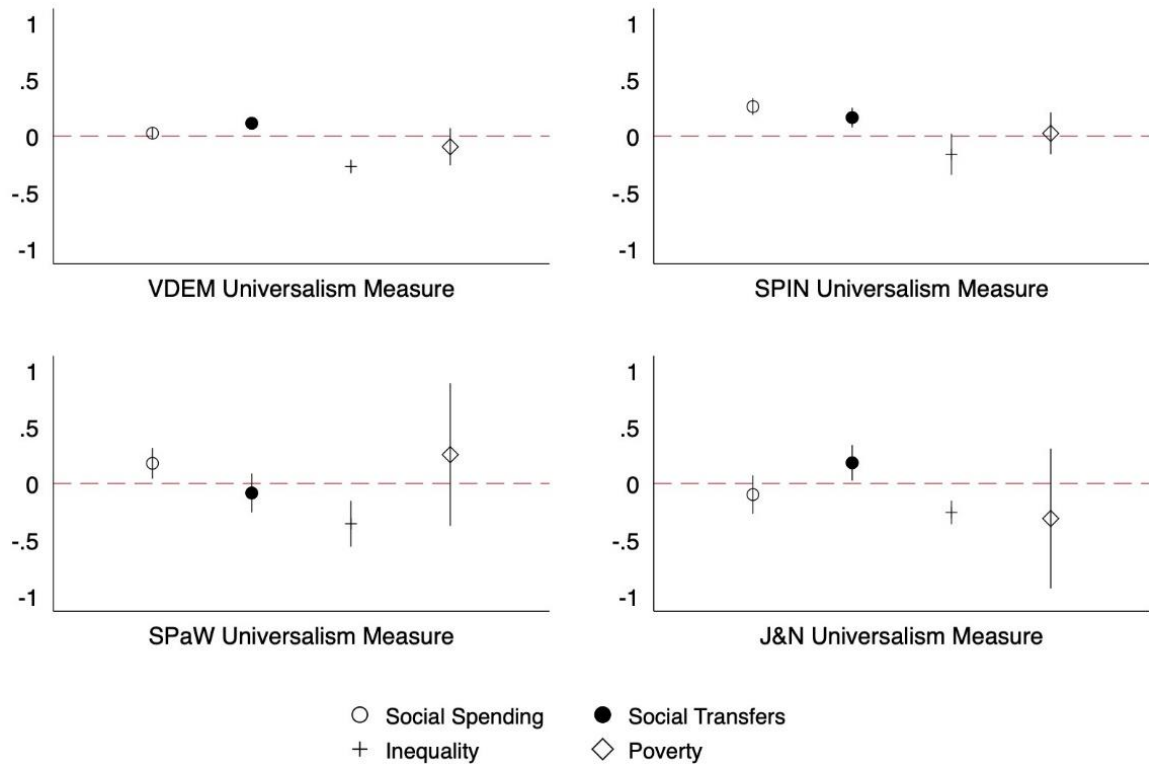
All models control for population, share of population 65+, immigration, unemployment, and GDP. Years available: *VDEM* (1960-2018); *SPIN* (1960-2018); *SPaW* (1960-1999); *J&N* (2000-2011).

I proceed by first highlighting findings depicting the greatest variability, and subsequently the least variability in measurement choice.

Table 3: Country fixed effects – V-DEM and SPIN universalism models.

Data source	V-DEM								SPIN							
Dep. Var.	Social Spending		Social transfers		Inequality		Poverty		Social spending		Social transfers		Inequality		Poverty	
Model	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)	
Coeff.	b	β	b	β	b	β	b	β	b	β	b	β	b	β	b	β
Universalism	.190 (.184)	.026 (.025)	.545*** (.098)	.114*** (.020)	-1.330*** (.154)	-.269*** (.031)	-.414 (.369)	-.094 (.084)	4.816*** (.700)	.262*** (.038)	2.000*** (.533)	.165*** (.044)	-2.028 (1.160)	-.161 (.093)	.275 (1.052)	.025 (.094)
R^2	.763		.633		.220		.115		.773		.599		.174		.099	
Observations	1,090		1,168		925		485		1,060		1,058		876		460	
Countries	22		22		22		19		22		22		22		18	
Time period	1960-2018		1960-2017		1960-2017		1970-2007		1960-2018		1960-2017		1960-2017		1970-2007	
Country FE?	Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes	

Standard errors in parentheses below parameter estimates; b = unstandardized coefficient; β = standardized coefficient. All models control for population, share of population 65+, immigration, unemployment, and GDP (coefficients not shown).

*p < 0.05, **p < 0.01, ***p < 0.001.

Table 4: Country fixed effects – SPaW and J&N universalism models.

Data source	SPaW								J&N							
Dep. Var.	Social Spending		Social transfers		Inequality		Poverty		Social spending		Social transfers		Inequality		Poverty	
Model	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)	
Coeff.	b	β	b	β	b	β	b	β	b	β	b	β	b	β	b	β
Universalism	.116* (.046)	.179* (.070)	-.037 (.038)	-.087 (.089)	-.159*** (.046)	-.357*** (.103)	.101 (.127)	.256 (.319)	-.599 (.526)	-.098 (.086)	.738* (.323)	.183* (.080)	-1.069*** (.221)	-.257*** (.053)	-1.150 (1.156)	-.310 (.312)
R^2	.614		.160		.087		.663		.559		.624		.303		.157	
Observations	343		383		262		156		240		240		240		117	
Countries	10		10		10		9		20		20		20		17	
Time period	1960-1999		1960-1999		1960-1999		1970-1999		2000-2011		2000-2011		2000-2011		2000-2011	
Country FE?	Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes	

Standard errors in parentheses below parameter estimates; b = unstandardized coefficient; β = standardized coefficient. All models control for population, share of population 65+, immigration, unemployment, and GDP (coefficients not shown).

*p < 0.05, **p < 0.01, ***p < 0.001.

Inequality is significantly and negatively associated with *'VDEM universalism'* and *'SPaW universalism'* at p<0.001, alongside *'J&N universalism'* at p<0.01. In fact, the largest association between universalism and any welfare state outcome is observed between *'SPaW universalism'* and inequality (Table 4, Model 3). Here, the standardized coefficient for *'SPaW universalism'* shows that a one-standard deviation increase in universalism is associated with an inequality decrease of 0.357 standard deviations on average—or 35.7% of an inequality standard deviation at p<0.001—holding all else constant in the model. In contrast, a statistically insignificant association is reported between *'SPIN universalism'* and inequality (Table 3, Model 7). The welfare state outcome varying the least by universalism measurement choice is social transfers. Social transfers were positively and significantly associated with universalism when using all but the *'SPaW universalism'* measure, producing standardized coefficients of 0.114 at p<0.001 (*'VDEM universalism'*), 0.165 at p<0.001 (*'SPIN universalism'*), and 0.183 at p<0.05 (*'J&N universalism'*), holding all else constant. Here, the standardized coefficient for *'J&N universalism'* depicts that a one-standard deviation increase in universalism is associated with an increase of 0.183 standard deviations in social transfers—or 18.3% of a social transfers standard deviation at p<0.05—holding all else constant (Table 4, Model 6).

4.2   Universalism and Social Spending

The strongest association was observed for *'SPIN universalism'* where a one-standard-deviation increase in universalism is associated with a social spending increase of 0.262 standard deviations on average—or 26.2% of a social spending standard deviation at p<0.001—holding all else constant (Table 3, Model 5). Further, a statistically significant association was also observed for *'SPaW universalism'*, where a one-standard-deviation increase in universalism is associated with a social spending increase of 0.179 standard deviations on average—or 17.9% of a social spending standard deviation at p<0.05—hold all else constant (Table 4, Model 1).

A statistically insignificant association is reported between universalism and social spending among *'VDEM universalism'* (Table 3, Model 1) and *'J&N universalism'*

(Table 4, Model 5). The effect size of each universalism measure for each welfare state outcome was plotted to intuitively illustrate patterns in measurement differences (Figure 10). Each plot represents the standardized coefficient for universalism on welfare state outcomes at a 95% confidence level. To interpret the role and significance of measurement, multivariate models demonstrate modest support for a significant positive association between universalism and social spending.

Figure 11: Coefficient plot of regression of welfare state outcomes across universalism measures.
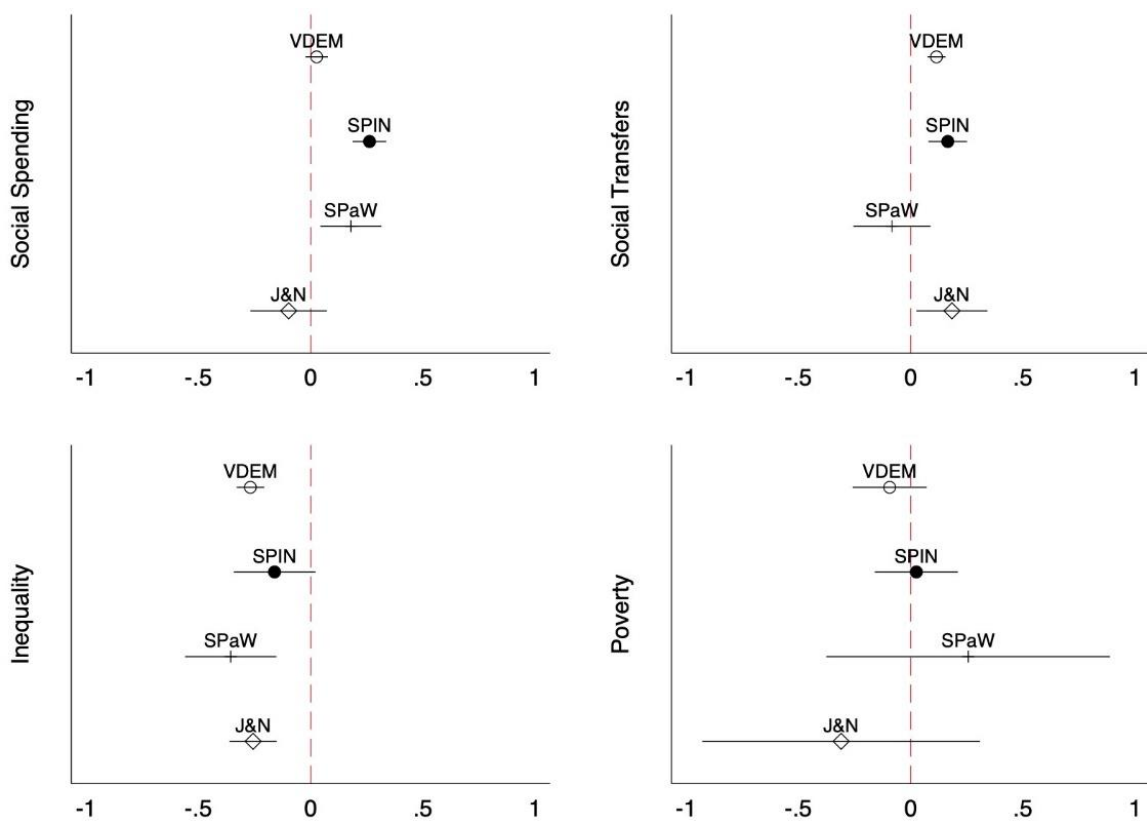
All models control for population, share of population 65+, immigration, unemployment, and GDP. Years available: *VDEM* (1960-2018); *SPIN* (1960-2018); *SPaW* (1960-1999); *J&N* (2000-2011).

Similarly, Figure 11 depicts the same information organized according to welfare state outcomes, thereby allowing for a straightforward comparison of how each welfare outcome varies by the universalism measure used. Not a single instance is observed where a grouping of four models predicting a given welfare state outcome remains both

statistically significant and occurs in the same direction across all measures of universalism.

4.3 Robustness Check

Figure 12: Coefficient plots of welfare state outcomes on universalism organized by model type.
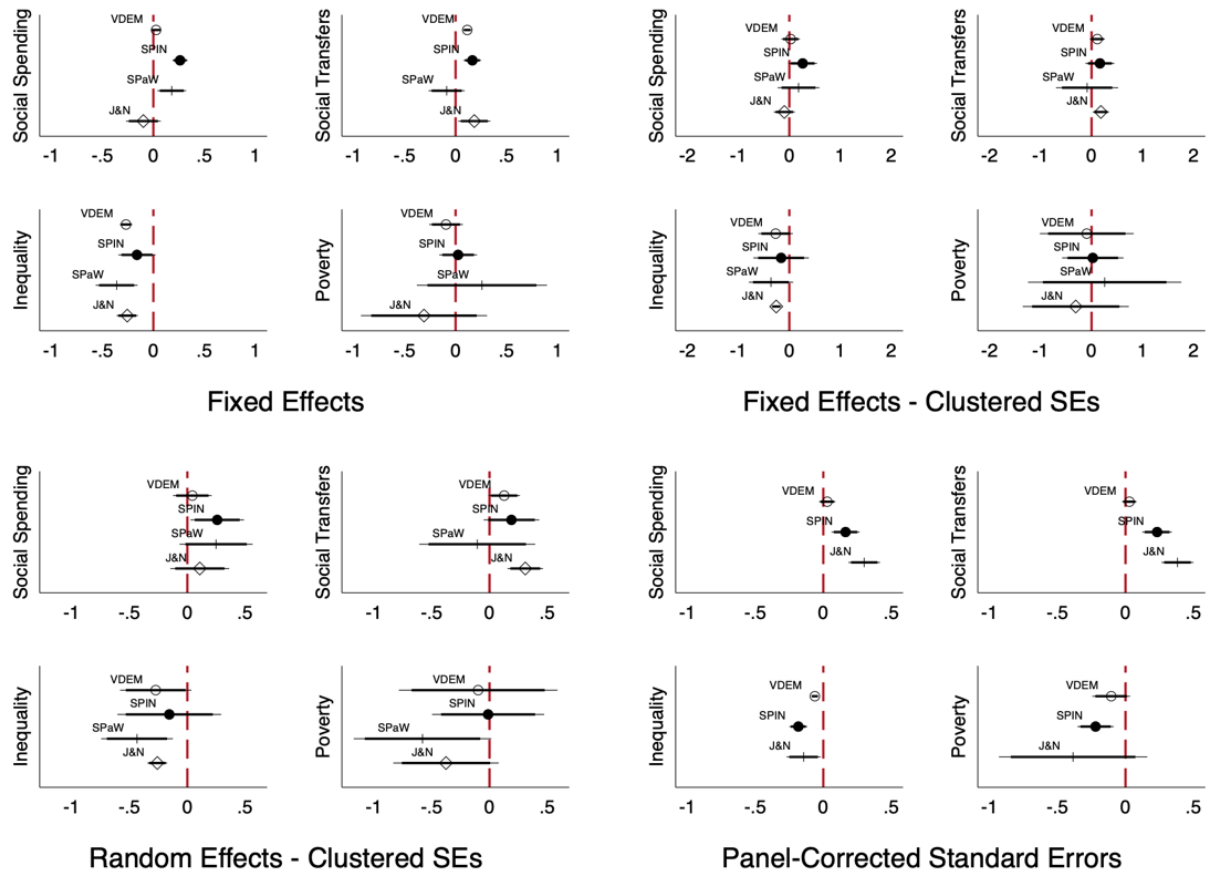
All models control for population, share of population 65+, immigration, unemployment, and GDP. Years available: *VDEM* (1960-2018); *SPIN* (1960-2018); *SPaW* (1960-1999); *J&N* (2000-2011).

Following my initial analysis, I proceed to scrutinize the robustness of the fixed effect (FE) models. My approach involved deploying both fixed and random effect estimations, each incorporating clustered standard errors. To this end, I employ panel-corrected standard error models for all parameters, barring the SPaW universalism measure which was characterized by an unbalanced panel (Figure 12). I exclude models that employed

random effects without clustered standard errors, given that their outcomes were akin to those of the fixed effect models.

In alternative models, the significance of social spending for the SPaW measure weakened. Despite this, social spending was relatively consistent for the SPIN measure. Further, the predicted influence of universalism on inequality demonstrated substantial fluctuations across different model types. Interestingly, this effect often lost its significance within the fixed and random effects clustered models, while its importance was upheld in the panel-corrected models. In addition, the J&N measure displayed an impressive level of consistency across all model types. The only departure from this consistency was for social spending; although it did not show significance for the initial three models, it gained significance in the panel-corrected models.

At the bivariate level, my findings suggested that welfare state outcomes, while variable in their magnitudes, displayed consistency in direction across universalism measures. However, this consistency appeared to wane in multivariate models, which demonstrated heightened variation across different measures. This was notably pronounced in the case of social spending. Reflecting on the relationship between universalism and social spending, my results reinforce the ongoing debate between targeting and universalism. As such, the choice between fixed and random effects models should be dictated by the nature of the universal measures in use.

Fixed effects models seem more suitable for universal measures that display substantial intra-country variations over time. Conversely, random effects models might perform better with universalism measures that exhibit less variability within, as opposed to between, countries. Regardless of the chosen model specification, findings from Figure 12 clearly underlines the substantial impact that the choice of measurement can have on welfare state outcomes. Therefore, a large portion of the understanding regarding universalism's role is dependent on the ways in which we define and measure core concepts.

Table 5: Country fixed effects – Comparing original and expanded version of SPIN universalism models.

Data source	SPIN 1								SPIN 2							
Dep. Var.	Social Spending		Social transfers		Inequality		Poverty		Social spending		Social transfers		Inequality		Poverty	
Model	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)	
Coeff.	b	β	b	β	b	β	b	β	b	β	b	β	b	β	b	β
Universalism	4.816*** (.700)	.262*** (.038)	2.000*** (.533)	.165*** (.044)	-2.028 (1.160)	-.161 (.093)	.275 (1.052)	.025 (.094)	10.837*** (.000)	.488*** (.036)	6.426*** (.619)	.438*** (.042)	-7.225*** (1.158)	-.476*** (.076)	-.966 (1.410)	-.071 (.104)
R²	.773		.599		.174		.099		.799		.632		.207		.100	
Observations	1,060		1,058		876		460		1,060		1,058		876		460	
Countries	22		22		22		18		22		22		22		18	
Time period	1960-2018		1960-2017		1960-2017		1970-2007		1960-2018		1960-2017		1960-2017		1970-2007	
Country FE?	Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes	

Standard errors in parentheses below parameter estimates; b = unstandardized coefficient; β = standardized coefficient. All models control for population, share of population 65+, immigration, unemployment, and GDP (coefficients not shown).

*p < 0.05, **p < 0.01, ***p < 0.001.

Figure 13: Coefficient plot of regression of welfare state outcomes comparing original and expanded SPIN universalism measures.
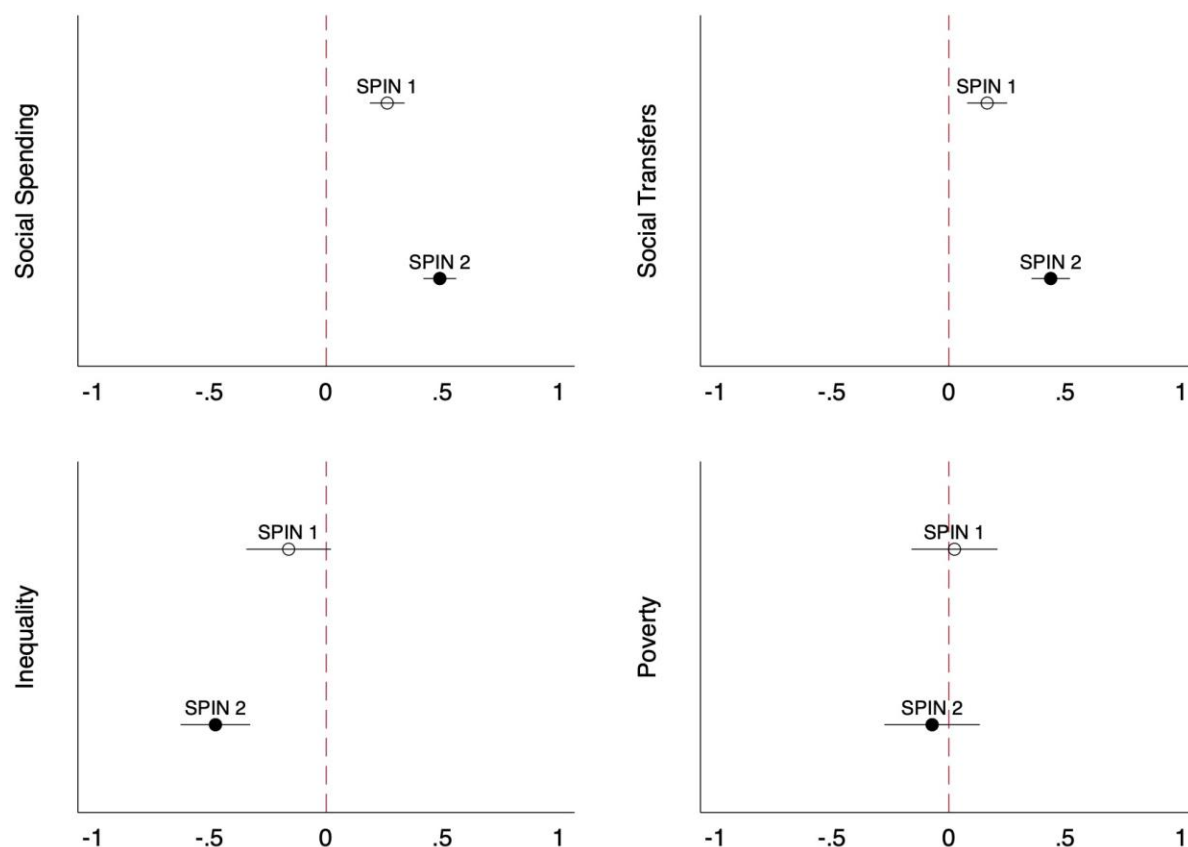
All models control for population, share of population 65+, immigration, unemployment, and GDP. *SPIN 1* measures universalism as the coverage ratio of labour force sickness benefits; *SPIN 2* measures universalism as the average score of both the coverage ratio of labour force sickness benefits *and* the coverage ratio of unemployment insurance benefits. Years available: *SPIN 1* (1960-2018); *SPIN 2* (1960-2018). Data source: The Social Insurance Entitlements Data Set (SIED 2020), a component of the Social Policy Indicator (SPIN) database, developed by Kenneth Nelson and colleagues in 2020.

In examining the sensitivity of relationships between universalism and welfare state outcomes, it is apparent that operationalization plays a substantial role in shaping the robustness of findings. This observation was discerned from the comparison between the original 'SPIN universalism' model (SPIN 1) and the expanded 'SPIN 2 universalism' model.

In the SPIN 1 model, a one-standard-deviation increase in universalism is associated with a 0.262 standard deviation increase in social spending and 0.165 in social transfers, both significant at $p<0.001$ (Table 5). The correlation with inequality remains insignificant, and poverty, although positively correlated, is also statistically insignificant. When transitioning to the expanded SPIN 2 model, the correlation with social spending and social transfers increases significantly to 0.488 and 0.438 standard deviations, respectively. Notably, the correlation with inequality becomes significant, whilst the correlation with poverty, though it changes from positive to negative, remains statistically insignificant (Table 5).

Figure 13 further emphasizes these differences in sensitivity, visually representing how the expanded SPIN 2 model yields stronger associations with each welfare state outcome. In understanding Figure 13, note that statistical significance is interpreted through the placement of the confidence interval relative to the dotted vertical line at 0. When a confidence interval crosses this line, the association is reported as statistically insignificant. Positive associations are identified to the right of the line, and negative to the left. SPIN 2's stronger association with welfare state outcomes is clearly indicated by its larger distance from the 0 line. Particularly notable is the significant association of 'inequality' under SPIN 2, contrasting with its insignificance under SPIN 1. 'Poverty', although remaining statistically insignificant under both models, exhibits a stronger correlation in SPIN 2 (Figure 13; Table 5).

Additionally, Appendices C and D, alongside Figures 3 and 7, further illustrate the increased sensitivity of the SPIN 2 model. When comparing scatterplots of 'social spending' by measures of universalism, SPIN 2 exhibits a higher correlation coefficient ($r=0.41$) than SPIN 1 ($r=0.38$). Welfare state outcomes, measured by the SPIN 2 model in Appendix D, also display stronger correlations than those measured by the original SPIN model in Figure 7.

To summarize, these findings emphasize the crucial role of measurement in studying the relationships between universalism and welfare state outcomes. Even minor variations in operationalization can significantly alter the robustness and directionality of the results,

emphasizing the need for careful consideration of the operationalization methods employed in universalism studies to ensure reliable findings.

Chapter 5: Discussion

5 Measurement Matters

While the targeting-universalism debate is likely to persist, my analyses offer clear evidence of the importance of measurement choices in analyzing the relationships between universalism and welfare state outcomes. My research delves into the intricacies of this debate by applying a more holistic perspective, scrutinizing the diverse impacts of universalism in terms of variation over both the reach and manifestation of its consequences.

One element of analysis that warrants discussion is why different measures of universalism are differentially related to the social welfare outcomes. The nature of these relationships may be explained by how the universalism measures are constituted. For example, the SPIN and SPaW measures – which are more institutionally-oriented – demonstrate stronger associations with social spending, showing different levels of statistical significance. In contrast, measures that perhaps capture other aspects of universalism might relate differently to the outcomes. Understanding the underlying dimensions that each measure reflects – be they institutional, outcome-oriented, or something else – can provide insights into their differing levels of association with welfare state outcomes.

Speaking to the first, my findings offer mixed support for the importance of universalism for generous social spending. Two of four tested universalism measures were found to be positively and significantly associated with social spending. This can be interpreted in two ways. On the one hand, in terms of past literature, my models provide more support for the stance taken by scholars who cast doubt on the existence of an association between universalism and social spending, and the relevance of universalism for the overall success and efficacy of welfare states (Kenworthy 2011, Marx, Salanauskaite, and Verbist 2013, Brady and Bostic 2015).

On the other hand, universalism is significantly and positively associated with social spending for two universalism measures, although at varying levels of statistical

significance (SPIN at p<0.001; SPaW at p<0.05). Accordingly, the debate remains for the targeting-universalism question with respect to social spending generosity. It is important to consider, however, that these mixed findings could be reflective of the complexity inherent in measuring universalism, and not necessarily a reflection of the theoretical construct itself.

In contrast, my findings provide clear support for the relationships between universalism and my four welfare state outcomes being sensitive to measurement choices. The extent to which choice of universalism indicator alters universalism's association with each welfare state outcome is conveyed most clearly by Figures 10 and 11. These findings offer a point of entry for further consideration of the various ways in which methodological choices may influence findings within welfare state literature. This aligns with concerns over the 'garden of forking paths' argument advanced by Gelman and Loken (2013), who demonstrate how model type, control choices, and many other steps within the analysis process can accumulate and combine in many ways to influence quantitative results. Resolving such scholarly debates and further uncovering the role of universalism seem to involve a more contorted than linear path. Amenta and Hicks (2021) echo these sentiments, pointing out that different methodological approaches have their unique strengths and weaknesses, thus stressing the need for scholars to employ more than one to maximize advantages while minimizing the limitations of any single approach. These perspectives foreground the necessity of being mindful of measurement and methodological choices when interpreting outcomes related to universalism and welfare state dynamics.

5.1   Revisiting the Drawing Board: A Way Forward for Comparative Welfare State Studies

The ongoing debate within comparative welfare state studies is actively reevaluating the merit and desirability of universalism (Korpi and Palme 1998). While scholars grapple with finding the superior political strategy for addressing poverty, inequality, and enhancing overall social spending outcomes, this study underscores the significance of the methodological choices in shaping welfare state outcomes.

In building upon the criticism from Gugushvili and Laenen (2021) of the KP index, this analysis supports the notion that often-overlooked factors such as demographic characteristics and pre-transfer incomes, do influence the operationalization and interpretation of universalism measures. The substantial variance observed in the correlation between universalism and social spending across different universalism measures (r=0.24 to r=0.74, Figures 3–4) supports this, as does the similar variance found in the correlation between universalism and other welfare state outcomes like poverty, inequality, and social transfers (Figures 5–9).

As the discourse over universalism intensifies, an intriguing pattern unfolds: the defining characteristics of universalism often become sidelined. This study, however, places these characteristics front and centre, demonstrating how the sensitivity of relationships between universalism and welfare state outcomes hinge on the operationalization of universalism. Country-specific trends over time and between countries varied significantly depending on the universalism measures employed (Figures 7–9).

Moreover, this study reaffirms findings from the literature that the effects of welfare states on poverty and inequality frequently hinge on the characteristics and organization of their social expenditure programs (Saunders 2021). For example, 'universal' programs, known to garner more backing due to their broad-based benefits (Goodin and Le Grand 1987), as opposed to 'targeted' schemes, which can heighten market inequality (Korpi and Palme 1998), can drastically shift the distributive outcomes of government policies (Saunders 2021).

This detailed examination of the operationalization of universalism and its consequent effects on the relationship with various welfare state outcomes addresses one of the core concerns raised in the literature review, and further corroborates the proposed 'targeting-universalism debate' paradox. The paradox suggests that as the concept of universalism and debates over it have become more prevalent, scholars may be straying from concern over truly defining universalism and its distinct characteristics. This study makes a substantial contribution to overcoming this paradox by highlighting the relevance and potential of alternative measures of universalism (Figures 3-9), thereby advancing the

debate. Additionally, conducting robustness checks further exemplified the pivotal role of operational choices. Findings differed notably across models and measures, pointing to a strong dependency of the relationship between universalism and welfare state outcomes on the specific definitions and measures of universalism used.

In reflection, this study echoes the key questions raised in the literature review regarding the potential association of universalism with social spending and its importance for other welfare state outcomes such as social transfers, poverty, and inequality. The findings reveal that the answer to these questions is not straightforward and are highly contingent upon how universalism is defined and measured. These insights align closely with concerns raised in recent literature regarding the reduced comparability between studies due to varying universalism measurement strategies (Rasmussen and Knutsen 2021). Rather than merely contesting prevailing practices, these findings propose a solution: fostering operational diversity and enhancing methodological rigor in future research, which Van Kersbergen and Vis (2015) pinpointed as critical to advancing the targeting-universalism debate.

Given these findings and the broadening definitions and metrics of poverty and inequality in current research (Saunders 2021), it is key to emphasize the need for clarity on the scale and framework of government schemes when assessing their influence on these variables (Esping-Andersen 1990). Also, we need to understand that the public's perception of inequality can be mistaken and could sway policy decisions (Osberg and Smeeding 2006). Therefore, a public conversation that reflects evidence-based findings is a necessity for successful policy making.

The future of comparative welfare state studies hinges upon a careful reconsideration of the methodological choices and assumptions. Such a consideration acknowledges the strength of different methodological approaches, their potential drawbacks, and the potential to combine them, as Amenta and Hicks (2021) suggest. While universalism remains central in understanding welfare state outcomes, how it's conceptualized and measured shapes the insights drawn. The debate on the most effective and efficient welfare state, still as profound as ever, can gain much from this study's findings, as it

urges future research to address measurement issues, incorporate diverse universalism measures, and consider influential factors in operationalization.

5.2   Directions for Future Works

Of course, this work is not without limitations. First, I acknowledge the problem of comparability across different universalism measures stemming from country-year inconsistencies. As Amenta and Hicks (2021) note, the literature on welfare states and social policy has seen various methodological approaches, with their own unique achievements and drawbacks. The focus on social spending as the primary measure has been a consistent trend, but this does not negate the issues that arise due to country-year inconsistencies that I've noted in my study. Differences in the years and countries included in each measure are not ideal and may bias findings in certain ways. Specifically, the issues of measurement and data availability in universalism metrics could have influenced the results, as hinted at by the varying correlation strengths across different universalism measures.

For example, while the *'J&N universalism'* measure produced the strongest correlation with social spending of the four measures, its correlation with social spending was found to be statistically insignificant in the fixed-effects model (Table 4). Findings may have more closely approximated those of the bivariate analysis if this measure was not limited to the years 2001–2011. Moreover, these methodological issues may also impact how we understand the degree of redistribution, which is related to the structure and size of individual programs (Korpi and Palme 1998). However, my thesis utilizes the best data currently available for conducting a comparison of this nature and does the most that can be done with such data. Strategies such as standardizing coefficients are used as one means of alleviating comparability issues. In this regard, future studies should strive to devise robust universalism measures that are not subject to country-year inconsistencies, which would improve the accuracy of the findings. This aligns with Saunders' (2021) call for the use of comparable data to draw robust conclusions about the impact of policy measures on inequality and poverty. A key concern here is that the lack of standardized measures and comparable data can lead to different outcomes. This reinforces the need

for a more coherent and systematic approach in universalism measurement, incorporating a broader set of indicators that capture the multi-dimensional aspects of universalism.

Second, it is important to note limitations associated with the application of fixed-effects (FE) techniques. One limitation of FE estimation is that it erases between-country variation (Alderson and Nielsen 1999). As a result, if a predictor variable – in this case, universalism – varies substantially between countries but varies little over time within each country, the precision of FE estimates will be compromised (Allison 2009). For this reason, some scholars view random-effect (RE) estimations to be preferable in such circumstances (Brady 2005). While I have conducted Hausman tests to determine that fixed-effects models are most preferable to random effects in my case (Bell and Jones 2015), the relatively flat trend lines depicted for the *'J&N universalism'* measure in Figure 1 suggests that the addition of RE estimation would be fruitful (Jacques and Nöel 2018). This corresponds to the suggestion by Amenta and Hicks (2021) that the use of pooled cross-sectional and time-series analyses of social spending can provide valuable insights. They also suggest that adjusting for nonindependence of errors over time and variability of error variance over time can be beneficial in addressing the challenges to unbiased and precise estimation. Thus, future works should apply approaches used by Huber and Stephens (2014) – where both FE and RE models are estimated – to the study of a wide range of universalism measures. This would add another dimension to the forked paths argument (Gelman and Loken 2013), alongside the sensitivity angle of this research agenda.

Future works should build on this focus of measurement by exploring how sensitivity findings differ by other modeling choices, such as with the incorporation of year fixed-effects, or with lagged dependent and independent variables. Another useful direction for future research would be the application of newly emerging methodologies like machine learning to this research question. This approach could potentially reveal intricate patterns in the data that traditional statistical methods might not detect. Despite these limitations and highlighted areas for future development, my M.A. thesis aims to carve out the initial groundwork for future research agendas focused on measurement issues in answering questions within the comparative welfare state literature. Although

universalism is significantly associated with social spending and other welfare state measures in some cases, my study provides evidence that much in terms of answers over universalism's role within the welfare state comes down to navigating the dense and weedy garden of forking paths.

References

Alderson, Arthur, and François Nielsen. 1999. "Income Inequality, Development, and Dependence: A Reconsideration." *American Sociological Review* 64(4):606–31. doi: 10.2307/2657259.

Allison, Paul. 2005. *Fixed Effects Regression Methods for Longitudinal Data Using SAS.* Cary, NC: SAS Institute.

Allison, Paul. 2009. *Fixed Effects Regression Models*. Thousand Oaks, CA: Sage Publications.

Amenta, Edwin, and Alexander Hicks. 2021. "Research Methods." Pp. 133–52. in *The Oxford Handbook of the Welfare State. 2nd edition.* Oxford University Press. doi: 10.1093/oxfordhb/9780198828389.013.8.

Anttonen, Anneli and Jorma Sipilä. 2014. "New Directions in Social Policy: Alternatives from and for the Global South." Pp. 1–18 in *Varieties of Universalism*. United Nations Research Institute for Social Development.

Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A. Olken, Ririn Purnamasari, and Matthew Wai-Poi. 2013. "Ordeal Mechanisms in Targeting: Theory and Evidence from a Field Experiment in Indonesia." NBER Working Paper No. 19127. http://www.nber.org/papers/w19127.

Baldini, Massimo, Giovanni Gallo, Manuel Reverberi, and Andrea Trapani. 2016. "Social Transfers and Poverty in Europe: Comparing Social Exclusion and Targeting Across Welfare Regimes." CAPP Paper No. 0145. http://155.185.68.2/CappPaper/Capp_p145.pdf.

Banfield, Edward. 1974. "The Imperatives of Class." Pp. 52–76 in *The Unheavenly City Revisited*. Boston, MA: Little, Brown, and Company.

Barrientos, Armando. 2019. "Does a General Theory of Welfare Institutions Explain the Expansion of Social Assistance in Low- and Middle-Income Countries?"

SOCIUM SFB 1342, Discussion Paper No. 1. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3652255.

Bell, Andrew, and Kelvyn Jones. 2015. "Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data." *Political Science Research and Methods* 3(1):133–53. doi: 10.1017/psrm.2014.7.

Besley, Timothy. 1990. "Testing versus Universal Provision in Poverty Alleviation Programmes." *Economica* 57(225):119–29. doi: 10.2307/2554085.

Besley, Timothy, and Ravi Kanbur. 1991. "The Principles of Targeting." Pp. 69–90 in *Current Issues in Development Economics*, edited by V. N. Balasubramanyam and S. Lall. London, UK: Macmillan Education.

Besley, Timothy and Stephen Coate. 1992. "Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty-Alleviation Programs." *American Economic Review* 82(1):249–61.

Blomqvist, Paula, and Joakim Palme. 2020. "Universalism in Welfare Policy: The Swedish Case Beyond 1990." *Social Inclusion* 8(1):114–23. doi: 10.17645/si.v8i1.2511.

Bradshaw, Jonathan, and Naomi Finch. 2002. "A Comparison of Child Benefit Packages in 22 Countries." Department for Work and Pensions Research Report No. 174. Leeds: Corporate Document Services.

Bradshaw, Richard. 2012. "The Case for Family Benefits." *Children and Youth Services Review* 34(3):590–96. doi: 10.1016/j.childyouth.2011.10.009.

Brady, David. 2005. "The Welfare State and Relative Poverty in Rich Western Democracies, 1967–1997." *Social Forces* 83(4):1329–64. doi: 10.1353/sof.2005.0056.

Brady, David, Evelyne Huber, and John Stephens. 2020. *Comparative Welfare State Data Set.* University of North Carolina and WZB Berlin Social Science Center.

Brady, David, and Rebekah Burroway. 2012. "Targeting, Universalism, and Single-Mother Poverty: A Multilevel Analysis Across 18 Affluent Democracies." *Demography* 49(2):719–46. doi: 10.1007/s13524-012-0094-z.

Brady, David, and Amie Bostic. 2015. "Paradoxes of Social Policy: Welfare Transfers, Relative Poverty, and Redistribution Preferences." *American Sociological Review* 80(2):268–98. doi: 10.1177/0003122415573049.

Brooks, Clem, and Jeff Manza. 2007. *Why Welfare States Persist: The Importance of Public Opinion in Democracies.* Chicago, IL: University of Chicago Press.

Busemeyer, Marius, and Torben Iversen. 2020. "The Welfare State with Private Alternatives: The Transformation of Popular Support for Social Insurance." *The Journal of Politics* 82(2):671–86. doi: 10.1086/706980.

Byun, Young-Hwan. 2022. "Welfare Expansion without Inequality Reduction: Institutional Explanation of Old-Age Poverty in Korea." *Journal of Social Policy* 1–21. doi: 10.1017/S0047279422000460.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, Nazifa Alizada, David Altman, Michael Bernhard, Agnes Cornell, M. Steven Fish, Lisa Gastaldi, Haakon Gjerløw, Adam Glynn, Allen Hicken, Garry Hindle, Nina Ilchenko, Joshua Krusell, Anna Lührmann, Seraphine F. Maerz, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Juraj Medzihorsky, Pamela Paxton, Daniel Pemstein, Josefine Pernes, Johannes von Römer, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Aksel Sundström, Eitan Tzelgov, Yi-ting Wang, Tore Wig, Steven Wilson and Daniel Ziblatt. 2021a. "V-Dem [Country–Year/Country–Date] Dataset v11.1" Varieties of Democracy (V-Dem) Project. doi: 10.23696/vdemds21.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Michael Bernhard, Agnes Cornell, M. Steven Fish, Lisa Gastaldi, Haakon Gjerløw, Adam Glynn, Allen Hicken, Anna Lührmann, Seraphine F. Maerz, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Pamela Paxton,

Daniel Pemstein, Johannes von Römer, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Aksel Sundtröm, Eitan Tzelgov, Luca Uberti, Yi-ting Wang, Tore Wig, and Daniel Ziblatt. 2021b. "V-Dem Codebook v11.1" Varieties of Democracy (V-Dem) Project.

Devereux, Stephen. 2016. "Is Targeting Ethical?" *Global Social Policy* 16(2):166–81. doi: 10.1177/1468018116643849.

Esping-Andersen, Gøsta. 1990. *The Three Worlds of Welfare Capitalism.* Cambridge, UK: Polity Press.

Esping-Andersen, Gøsta. 2015. "Welfare Regimes and Social Stratification." *Journal of European Social Policy* 25(1):124–34. doi: 10.1177/0958928714556976.

Garcia-Fuente, Xabier. 2021. "The Paradox of Redistribution in Time: Social Spending in 53 Countries, 1967–2018." LIS Working Paper Series No. 815. http://www.ecineq.org/wp-content/uploads/papers_EcineqLSE/EcineqLSE-331.pdf.

Garland, David. 2016. *The Welfare State: A Very Short Introduction*. Oxford: Oxford University Press.

Gelman, Andrew, and Eric Loken. 2013. "The Garden of Forking Paths: Why Multiple Comparisons can be a Problem, Even When There is No 'Fishing Expedition' or 'p-Hacking', and the Research Hypothesis was Posited Ahead of Time." Retrieved http://www.stat.columbia.edu/~gelman/research/%0Aunpublished/p_hacking.pdf.

Gilbert, Neil. 2002. *Transformation of the Welfare State: The Silent Surrender of Public Responsibility*. New York, NY: Oxford University Press.

Goodin, Robert, and Julian Le Grand. 1987. *Not Only the Poor: The Middle Classes and the Welfare State.* London: Allen & Unwin.

Gornick, Janet, and Timothy Smeeding. 2018. "Redistributional Policy in Rich Countries: Institutions and Impacts in Nonelderly Households." *Annual Review of Sociology* 44:441–68. doi: 10.1146/annurev-soc-073117-041114.

Gottschalk, Peter, and Timothy Smeeding. 1997. "Cross-National Comparisons of Earnings and Income Inequality." *Journal of Economic Literature* 35(2):633–87.

Gottschalk, Peter, and Timothy Smeeding. 2000. "Empirical Evidence on Income Inequality in Industrialized Countries." Pp. 261–308 in *Handbook of Income Distribution.* Amsterdam: North Holland.

Green-Pedersen, Christoffer. 2004. "The Dependent Variable Problem within the Study of Welfare State Retrenchment: Defining the Problem and Looking for Solutions." *Journal of Comparative Policy Analysis: Research and Practice* 6(1):3–14. doi: 10.1080/1387698042000222763.

Greenstein, Robert. 1991. "Universal and Targeted Approaches to Relieving Poverty: An Alternative View." Pp. 437–59 in *The Urban Underclass*, edited by C. Jencks and P. E. Peterson. Washington, DC: The Brookings.

Greve, Bent. 2020. "There is Still a Need for a Universal Welfare State." *IPPR Progressive Review* 27(1):69–74. doi: 10.1111/newe.12186.

Greve, Bent. 2022. "Individualism Versus Collectivism." Pp. 59–65 in *Rethinking Welfare and the Welfare State*. Cheltenham, UK: Edward Elgar Publishing.

Gugushvili, Dimitri, and Tijs Laenen. 2021. "Two Decades After Korpi and Palme's "Paradox of Redistribution": What Have We Learned so Far and Where do We Take it From Here?" *Journal of International and Comparative Social Policy* 37(2):112–27. doi: 10.1017/ics.2020.24.

Guillaud, Elvire, Matthew Olckers, and Michaël Zemmour. 2020. "Four Levers of Redistribution: The Impact of Tax and Transfer Systems on Inequality Reduction." *Review of Income and Wealth* 66(2):444–66. doi: 10.1111/roiw.12408.

Hannah, Adam, Jeremiah Brown, and Andrew Gibbons. 2020. "Welfare Capabilities: Evaluating Distributional Inequalities and Welfare Policy in Advanced Democracies." *Journal of European Social Policy* 30(3):293–305. doi: 10.1177/0958928719868447.

Huber, Evelyne, and John D. Stephens. 2001. *Development and Crisis of the Welfare State: Parties and Policies in Global Markets.* Chicago, IL: University of Chicago Press.

Huber, Evelyne, and John Stephens. 2014. "Income Inequality and Redistribution in Postindustrial Democracies: Demographic, Economic, and Political Determinants." *Socio-Economic Review* 12(2):245–67. doi: 10.1093/ser/mwu001.

Jacques, Olivier, and Alain Noël. 2018. "The Case for Welfare State Universalism, or the Lasting Relevance of the Paradox of Redistribution." *Journal of European Social Policy* 28(1): 70–85. doi: 10.1177/0958928717700564.

Jacques, Olivier, and Alain Noël. 2021. "Targeting within Universalism." *Journal of European Social Policy* 31(1):15–29. doi: 10.1177/0958928720918973.

Jäntti, Markus, and Sheldon Danziger. 2000. "Income Poverty in Advanced Countries." Pp. 309–78 in *Handbook of Income Distribution.* Amsterdam: North Holland.

Katz, Michael. 2013. *The Undeserving Poor: America's Enduring Confrontation with Poverty.* Oxford, UK: Oxford University Press.

Kenworthy, Lane, and Jonas Pontusson. 2005. "Rising Inequality and the Politics of Redistribution in Affluent Countries." *Perspectives on Politics* 3(3):449–71. doi: 10.1017/S1537592705050292.

Kenworthy, Lane. 2009. "The Effect of Public Opinion on Social Policy Generosity." *Socio-Economic Review* 7(4):727–40. doi: 10.1093/ser/mwp014.

Kenworthy, Lane. 2011. *Progress for the Poor.* Oxford, UK: Oxford University Press.

Korpi, Walter, and Joakim Palme. 1998. "The Paradox of Redistribution and Strategies of Equality: Welfare State Institutions, Inequality, and Poverty in the Western Countries." *American Sociological Review* 63(5):661–87. doi: 10.2307/2657333.

Korpi, Walter. 2010. "Class and Gender Inequalities in Different Types of Welfare States: The Social Citizenship Indicator Program (SCIP)." *International Journal of Social Welfare* 19(1):14–24. doi: 10.1111/j.1468-2397.2010.00730.x.

Lee, Dwight. 1987. "The Trade-off between Equality and Efficiency: Short-Run Politics and Long-Run Realities." *Public Choice* 53:149–65. doi: 10.1007/BF00125846.

Lindbeck, Assar. 1988. "Individual Freedom and Welfare State Policy." *European Economic Review* 32:295–318. doi: 10.1016/0014-2921(88)90174-2.

Lindbeck, Assar. 1994. "The Welfare State and the Employment Problem." *American Economic Review* 84(2):71–5.

Lindbeck, Assar, Sten Nyberg, and Jörgen Weibull. 2003. "Social Norms and Welfare State Dynamics." *Journal of the European Economic Association* 1(2–3):533–42. doi: 10.1162/154247603322391170.

Lindert, Peter. 2004. "Patterns and Puzzles." Pp. 3–19 in *Growing Public: Social Spending and Economic Growth since the Eighteenth Century*. Vol. 1. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511510717.

Lindh, Arvid, and Leslie McCall. 2022. "Bringing the Market in: An Expanded Framework for Understanding Popular Responses to Economic Inequality." *Socio-Economic Review* 0(0):1–21. doi: 10.1093/ser/mwac018.

LIS (Luxembourg Income Study). 2010a. "LIS Database." http://lisproject.org/techdoc.htm.

LIS (Luxembourg Income Study). 2010b. "LIS Key Figures." http://lisproject.org/key-figures/key-figures.htm.

Marchal, Sarah, and Wim van Lancker. 2019. "The Measurement of Targeting Design in Complex Welfare States: A Proposal and Empirical Applications." *Social Indicators Research* 143(2):693–726. doi: 10.1007/s11205-018-1995-z.

Marx, Ive, Lina Salanauskaite, and Gerlinde Verbist. 2013. "The Paradox of Redistribution Revisited: And That it May Rest in Peace?" IZA Discussion Paper No. 7414.

Marx, Ive, Lina Salanauskaite, and Gerlinde Verbist. 2016. "For the Poor, but Only the Poor: On Optimal Pro-Poorness in Redistributive Policies." *Social Forces* 95(1):1–24. doi: 10.1093/sf/sow058.

Mead, Lawrence. 2003. "A Biblical Response to Poverty." Pp. 53–106 in *Lifting Up the Poor: A Dialogue on Religion, Poverty, and Welfare Reform*, edited by M. J. Bane and L. Mead. Washington, DC: Brookings Institution Press.

Murray, Charles. 1984. *Losing Ground: American Social Policy 1950-1980*. New York, NY: Basic Books.

Nelson, Kenneth. 2004. "Mechanisms of Poverty Alleviation: Anti-Poverty Effects of Non-Means-Tested and Means-Tested Benefits in Five Welfare States." *Journal of European Social Policy* 14(4):371–90. doi: 10.1177/0958928704046879.

Nelson, Kenneth. 2007. "Universalism Versus Targeting: The Vulnerability of Social Insurance and Means-Tested Minimum Income Protection in 18 Countries, 1990–2002." *International Social Security Review* 60(1):33–58. doi: 10.1111/j.1468-246X.2007.00259.x.

Nelson, Kenneth, Daniel Fredriksson, Tomas Korpi, Walter Korpi, Joakim Palme, and Ola Sjöberg. 2020. "The Social Policy Indicators (SPIN) Database." *International Journal of Social Welfare* 29(3):285–89. doi: 10.1111/ijsw.12418.

Obinger, Herbert. 2021. "Social Expenditure and Welfare State Financing." Pp. 453–72 in *The Oxford Handbook of the Welfare State. 2nd edition.* Oxford University Press. doi: 10.1093/oxfordhb/9780198828389.013.26.

OECD. *Income Distribution Database (IDD).* N.d. www.oecd.org/social/income-distribution-database.htm.

OECD. 1985. "Social Expenditure 1960–1990. Problems of Growth and Control." *Social Policy Studies.* Paris, France: OECD Publishing.

OECD. 2014. *Social Expenditure Update.* Paris, France: OECD Publishing.

OECD. 2015. *In It Together—Why Less Inequality Benefits All.* Paris, France: OECD Publishing. doi: 10.1787/9789264235120-en.

OECD. 2016. *Social Expenditure Database*. Paris, France: OECD Publishing.

OECD. 2020. "National Accounts at a Glance (Edition 2019)." *OECD National Accounts Statistics* (database). doi: 10.1787/b6691544-en.

Öktem, Kerem. 2020. "The Welfare State as Universal Social Security: A Global Analysis." *Social Inclusion* 8(1):103–13. doi: 10.17645/si.v8i1.2509.

Osberg, Lars, and Timothy Smeeding. 2006. ""Fair" Inequality? Attitudes Toward Pay Differentials: The United States in Comparative Perspective." *American Sociological Review* 71 (3):450–73. doi: 10.1177/000312240607100305.

Pemstein, Daniel, Kyle Marquardt, Eitan Tzelgov, Yi-ting Wang, Juraj Medzihorsky, Joshua Krusell, Farhad Miri, and Johannes von Römer. 2021. *The V-Dem Measurement Model: Latent Variable Analysis for Cross-National and Cross-Temporal Expert-Coded Data.* V-Dem Working Paper No. 21, 6th edition. University of Gothenburg: Varieties of Democracy Institute. doi: 10.2139/ssrn.3799967.

Pierson, Paul. 1993. "When Effect Becomes Cause: Policy Feedback and Political Change." *World Politics* 45(4):595–628. doi: 10.2307/2950710.

Pierson, Paul. 1994. *Dismantling the Welfare State? Reagan, Thatcher, and the Politics of Retrenchment*. Cambridge, UK: Cambridge University Press.

Pierson, Paul. 1996. "The New Politics of the Welfare State." *World Politics* 48(2):143–79. doi: 10.1353/wp.1996.0004.

Piketty, Thomas, and Emmanuel Saez. 2003. "Income Inequality in the United States, 1913–1918." *Quarterly Journal of Economics* 118(1):1–39.

Rasmussen, Magnus. 2016. *The Social Policy Around the World (SPaW) Database, 1870–2010.* Oslo, Norway: University of Oslo and Bjørknes University College.

Rasmussen, Magnus, and Carl Knutsen. 2021. "Party Institutionalization and Welfare State Development." *British Journal of Political Science* 51(3):1203–29. doi: 10.1017/S0007123419000498.

Roine, Jesper, Jonas Vlachos, and Daniel Waldenström. 2009. "The Long-Run Determinants of Inequality: What Can We Learn from Top Income Data?" *Journal of Public Economics* 93(7–8):974–88. doi: 10.1016/j.jpubeco.2009.04.003.

Rothstein, Bo. 1998. *Just Institutions Matter: The Moral and Political Logic of the Universal Welfare State.* Cambridge, UK: Cambridge University Press.

Rothstein, Bo. 2001. "The Universal Welfare State as a Social Dilemma." *Rationality and Society* 13(2):213–33. doi: 10.1177/104346301013002004.

Saez, Emmanuel. 2006. "Redistribution Toward Low Incomes in Richer Countries." Pp. 187–201 in *Understanding Poverty*, edited by A. Bannerjee, R. Benabou, and D. Mookherjee. New York, NY: Oxford University Press.

Sainsbury, Diane. 1996. *Gender, Equality, and Welfare States*. Cambridge: Cambridge University Press.

Sánchez, Angélica, and Thomas Goda. 2018. "Corruption and the 'Paradox of Redistribution.'" *Social Indictors Research* 140(1):675–93. doi: 10.1007/s11205-017-1781-3.

Saunders, Peter. 2021. "Inequality and Poverty." Pp. 717–32 in *The Oxford Handbook of the Welfare State. 2nd edition.* Oxford University Press. doi: 10.1093/oxfordhb/9780198828389.013.41.

Scruggs, Lyle, and Gabriela Tafoya. 2022. "Fifty Years of Welfare State Generosity." *Social Policy & Administration* 56(5):791–807. doi: 10.1111/spol.12804.

Sefton, Tom. 2006. "Distributive and Redistributive Policy." Pp. 607–23 in *The Oxford Handbook of Public Policy*, edited by M. Moran, M. Rein, and R. E. Goodin. New York, NY: Oxford University Press.

Shelton, Cameron. 2007. "The Sizes and Composition of Government Expenditure." *Journal of Public Economics* 91(11–12):2230–60. doi: 10.1016/j.jpubeco.2007.01.003.

Skocpol, Theda. 1992. *Protecting Soldiers and Mothers: The Political Origins of Social Policy in the United States*. Cambridge, MA: Harvard University Press.

Smeeding, Timothy. 2005. "Public Policy, Economic Inequality, and Poverty: The United States in Comparative Perspective." *Social Science Quarterly* 86(1):955–83. doi: 10.1111/j.0038-4941.2005.00331.x.

Solt, Frederick. 2016. "The Standardized World Income Inequality Database." *Social Science Quarterly* 97(5):1267–81. SWIID Version 7.1, August 2018. doi: 10.1111/ssqu.12295.

Squire, Lyn. 1993. "Fighting Poverty." *American Economic Review* 83(2):377–82.

Stiglitz, Joseph, Amartya Sen, and Jean-Paul Fitoussi. 2009. "Report by The Commission on The Measurement of Economic Performance and Social Progress." The Commission: Paris. https://web.archive.org/web/20161001093447/http://ec.europa.eu/eurostat/documents/118025/118123/Fitoussi+Commission+report.

Sumino, Takanori. 2018. "Identity Priming and Public Opinion on Income Inequality: Robustness Testing of the Micro-Level Mechanism of the Paradox of Redistribution." *Socio-Economic Review* 16(3):545–65. doi: 10.1093/ser/mwx038.

Thompson, Simon, and Paul Hoggett. 1996. "Universalism, Selectivism and Particularism: Towards a Postmodern Social Policy." *Critical Social Policy* 16(46):21–42. doi: 10.1177/026101839601604602.

Van Kersbergen, Kees, and Barbara Vis. 2015. "Three Worlds' Typology: Moving Beyond Normal Science?" *Journal of European Social Policy* 25(1):111–23. doi: 10.1177/0958928714556974.

Appendices

Appendix A: Table 2 replication with addition of SPIN 2 independent variable descriptive statistics.

	N	Mean	SD
Dependent variables			
Social spending as % of GDP	1212	18.30	6.09
Inequality [†]	1001	28.97	4.16
Social transfers as % of GDP [‡]	1266	12.57	4.02
Poverty [§]	505	9.80	3.71
Independent variables			
VDEM universalism	1342	1.40	.84
SPIN 1 universalism	1122	.69	.33
SPIN 2 universalism	1122	.66	.27
SPaW universalism	383	28.52	9.45
J&N universalism	240	$8.3 \ (\times 10^{-9})$	1.00
KP universalism	66	-.20	.11
Control variables			
Population (in thousands)	1201	35,265	55,453
Population share of elderly	1200	.13	.03
Immigration [#]	1202	9.83	7.54
Unemployment rate	1276	5.81	4.29
Real GDP [l]	1210	1,025,196	2,165,864

[†]Household Income Gini Coefficient (Post-Tax-and-Transfer); [‡]Excludes social transfers in-kind; [§]Total population poverty rate based on 50% median-income threshold; [#]Migrant stock as % of population; [l]Expenditure-side real GDP at chained PPPs, in millions of 2011 USD.

Appendix B: Figure 1 replication using SPIN 2 measure.
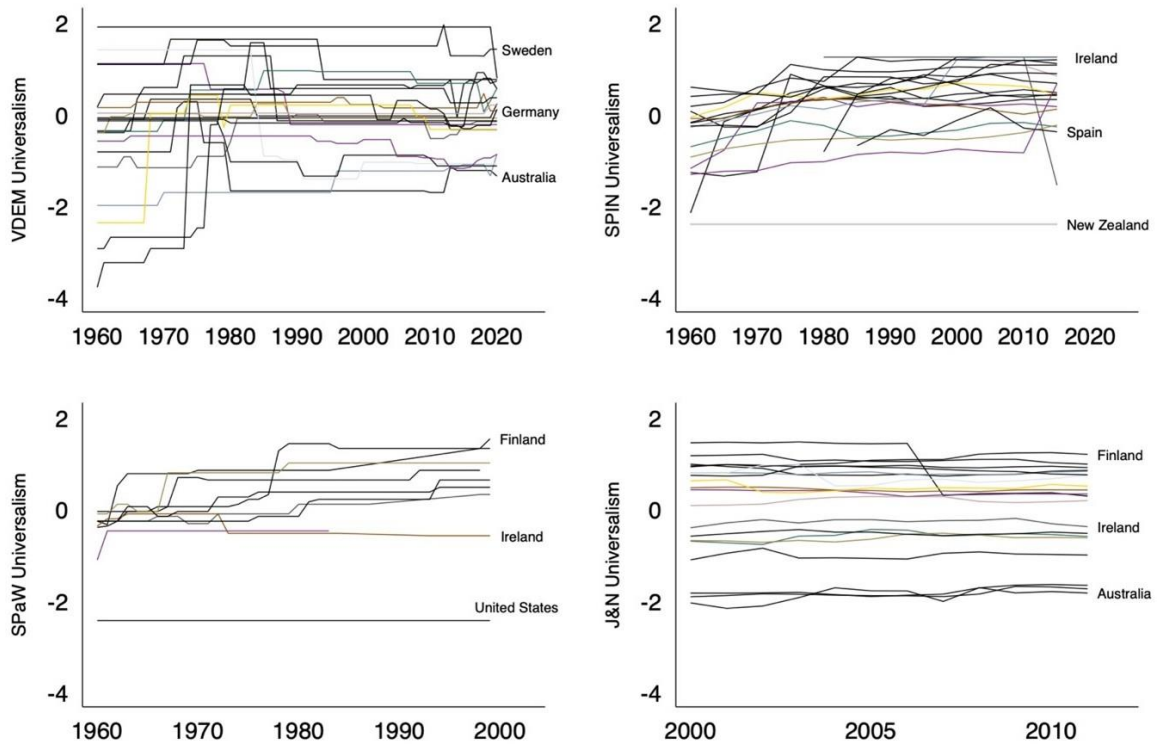
All universalism measures displayed are standardized. Data sources: Michael Coppedge and colleagues (V-Dem Project 2021) used for *VDEM* (1960-2018); Kenneth Nelson and colleagues (SIED 2020) used for *SPIN* (1960-2015); Magnus Rasmussen (SPaW 2016) used for *SPaW* (1960-1999); Olivier Jacques and Alain Noël (OECD 2014; OECD 2016) used for *J&N* (2000-2011). Sweden, Denmark, Finland, Greece, Ireland, and Norway all tie for universalism score in 2015.

Appendix C: Figure 3 replication using SPIN 2 measure.
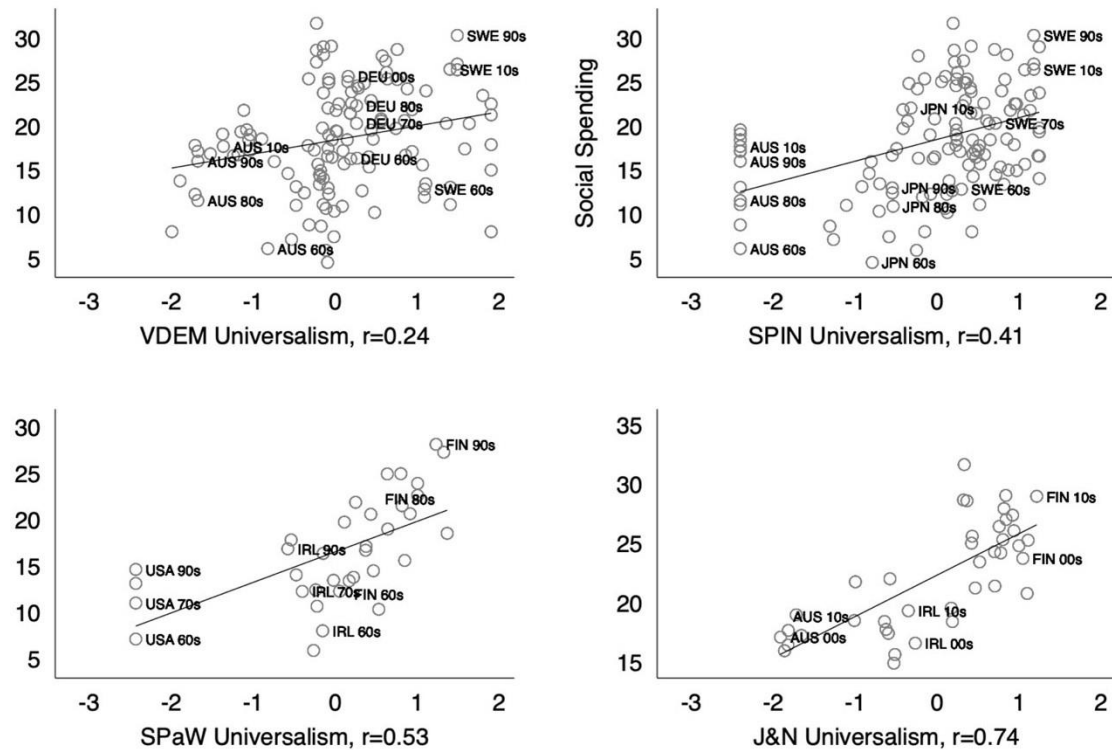
Plotted values represent decade clustered means. Data sources: Esteban Ortiz-Ospina and Max Roser (OCED 1985; OECD 2016) used for *Social Spending*; Michael Coppedge and colleagues (V-Dem Project 2021) used for *VDEM* (1960-2018); Kenneth Nelson and colleagues (SIED 2020) used for *SPIN* (1960-2015); Magnus Rasmussen (SPaW 2016) used for *SPaW* (1960-1999); Olivier Jacques and Alain Noël (OECD 2014; OECD 2016) used for *J&N* (2000-2011).

Appendix D: Figure 7 replication using SPIN 2 measure.
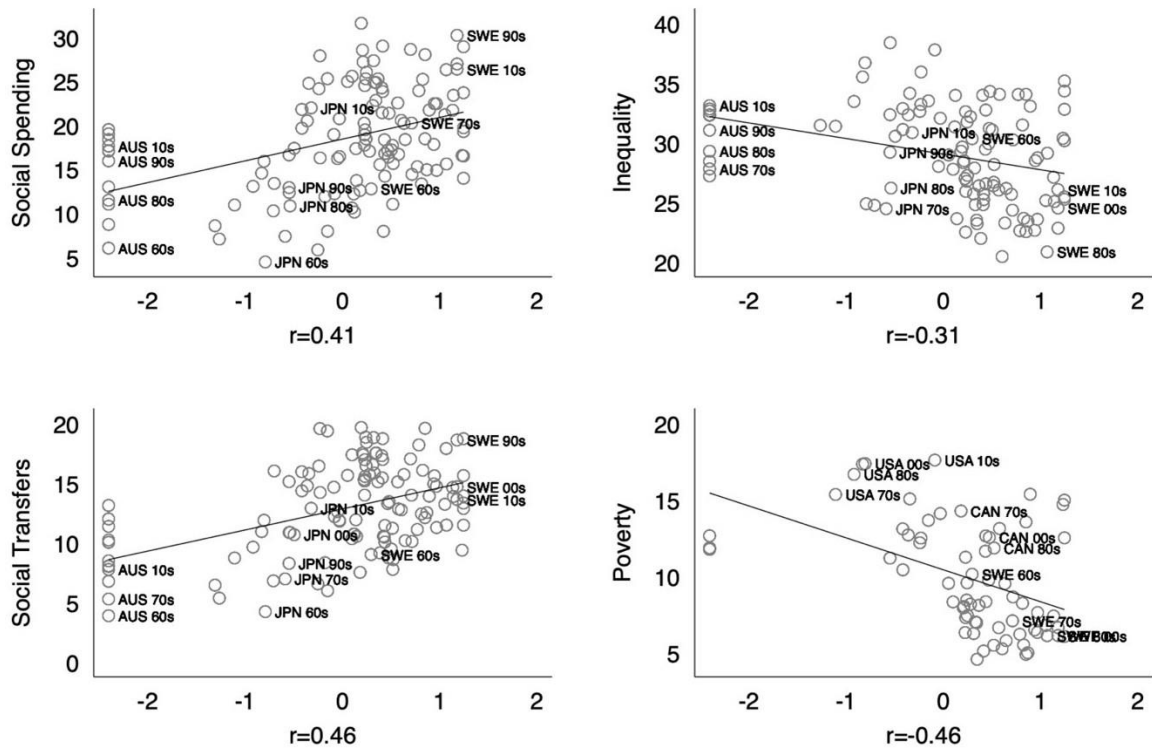
Plotted values represent decade clustered means. Data source: The Social Insurance Entitlements Data Set (SIED 2020), a component of the Social Policy Indicator (SPIN) database, developed by Kenneth Nelson and colleagues in 2020.

Curriculum Vitae

Name: Kaitlin Pauline Wannamaker

Post-secondary   The University of Western Ontario
Education and   London, Ontario, Canada
Degrees: 2021-2023 M.A.

                                          The University of Western Ontario
                                          London, Ontario, Canada
                                          2016-2021 B.A.

Honours and   SSHRC Doctoral Research Award
Awards: Social Sciences and Humanities Research Council
                                          2023-2024 *[waitlisted]*

                                          Ontario Graduate Scholarship – Master's
                                          The University of Western Ontario
                                          2022-2023

                                          Canada Graduate Scholarship – Master's
                                          Social Sciences and Humanities Research Council
                                          2021-2022

                                          Ontario Graduate Scholarship – Master's
                                          The University of Western Ontario
                                          2021-2022 *[declined]*

                                          Western Scholarship of Excellence
                                          The University of Western Ontario
                                          2016

Related Work   Graduate Student Assistant
Experience: The University of Western Ontario
                                          2022-2023

                                          Graduate Teaching Assistant
                                          The University of Western Ontario
                                          2021-2022

                                          Undergraduate Research Assistant
                                          The University of Western Ontario
                                          2020-2021

Publications:

Wannamaker, Kaitlin. 2023. Review of "*Poorly Understood: What America Gets Wrong About Poverty,* by Mark Rank, Lawrence Eppard, and Heather Bullock." *Canadian Journal of Sociology* 47(1):95–98. doi: 10.29173/cjs29912.

Adams, Tracey, and Kaitlin Wannamaker. 2022. "Professional Regulation, Profession-State Relations and the Pandemic Response: Australia, Canada, and the UK Compared." *Social Science & Medicine* 296:114808. doi: 10.1016/j.socscimed.2022.114808.

Wannamaker, Kaitlin. 2020. Review of "*Poverty and Power: The Problem of Structural Inequality,* by Edward Royce." *Canadian Journal of Sociology* 45(4):377–80.

Conference Presentations:

"Evaluating Welfare Policy: A 'Universal' Ticket to the Good Life?" 2023. *Accepted* for oral presentation at the Society for the Advancement of Socio-Economics (SASE) annual conference, Federal University of Rio de Janeiro, Brazil. *[Unable to attend]*

"Universal in What Sense? Mapping out the Conceptual and Operational Terrain of Welfare State Universalism." 2023. *Accepted* for oral presentation at the International Sociological Association (ISA) World Congress of Sociology, Melbourne Convention and Exhibition Centre, Melbourne, Australia. *[Unable to attend]*

"Universal in What Sense? Mapping out the Conceptual and Operational Terrain of Welfare State Universalism." 2022. Society for the Advancement of Socio-Economics (SASE) annual conference, University of Amsterdam, The Netherlands.

"Professional Healthcare Workers and Regulation under COVID-19", with Tracey Adams. 2021. Canadian Sociological Association conference, held virtually.