Why Is the Accumulation of Knowledge so Hard?

Exploring Econometric Research on the Determinants of Public Social Spending in Latin America

Svenja Flechtner (University of Siegen)

Diego Sánchez-Ancochea (University of Oxford)

This version: 5 January 2021

Abstract

Many areas in applied econometric research within political economy fail to come up with conclusive findings. This is, for example, the case with studies on the determinants of public social spending in Latin America, a key area of research given the impact of social programs on poverty, inequality and welfare more generally. In this area, like in others, it is hard to identify clear answers regarding the impact of economic processes and political institutions. Two reasons explain this lack of knowledge accumulation in this case. First, each study uses different data sources and analytical models. Second, some of the empirical strategies required to solve various econometric problems may affect the results. Our analysis thus questions the role of econometric research as the only method to explore political economy questions and highlights the importance of promoting conversations between complementary methods of both quantitative and qualitative traditions.

Keywords: public social spending, Latin America, research methods, social policy

Introduction

Econometrics occupies an influential place in the study of political economy. Common arguments put forward are that econometrics is more rigorous than qualitative approaches, uses "objective" data and yields clearer, quantifiable results about relationships and effects. Yet, in practice, econometric research in many areas of the social sciences struggles to offer definite answers to important policy questions.

We illustrate this problem using the case of social policy in Latin America. This is a central area of concern for both social scientists and policymakers. From an analytical perspective, the evolution of social policy tells us much about how societies operate and how they resolve class struggles. From a policy perspective, the level and allocation of social spending has major implications on poverty, inequality and human development more generally.

Following previous studies in OECD countries, the literature on Latin America has explored the role of economic and political drivers in determining social spending in the region. Different studies have rightly emphasized the importance of democracy and the need to consider social insurance, social assistance and health care separately. Yet, overall, econometric results have been inconclusive: for example, leftist parties, trade openness or urbanization have a positive, neutral or negative effect on social spending, depending on the studies.

The fact that different papers using different data and methods produce ambivalent answers is not, as such, surprising: it is what motivates new studies on the same topic in the first place. We should obviously expect variation in research design and methodology to lead to a diversity of results. However, while some differences in results should be expected as a consequence of research design and methodology, the lack of significant agreement and the existence of important contradictions should not. In this sense, following Bird (2007), we

understand scientific progress in the following simple terms: after some time of research, there should be more common knowledge and understanding than before. This does not currently seem the case in the study of the determinants of public social spending in Latin America. Hence we need to understand better where differences come from, and why different studies do not converge to clear understandings. In order to promote knowledge accumulation, research fields should find a fair balance between promoting research creativity and avoiding unnecessary differences in research designs.

Econometric research practices in applied fields have been criticized before.[1] In contrast to these contributions, we do not focus on providing a detailed critique of specific practices or proposing better alternatives. Rather, our goal is to suggest reasons for inconclusive quantitative evidence, where econometric mistakes are only one aspect of the story. In many instances, even when econometric tools are used in a technically and theoretically satisfactory way, there remains scope for ambiguities. We hope to enable a better understanding of some limitations of quantitative contributions in applied political economy research, leading to constructive discussions about ways forward.

To do so, we review the most relevant papers published in the last two decades on the determinants of social spending in Latin America. We argue that data limitations and a variety of econometric problems related to panel studies (e.g. unobserved country characteristics, dependencies across units and time, endogeneity issues) have forced authors to make a series of technical decisions with both analytical and empirical implications. As a result, comparing results across studies and building consensus has proven particularly hard.

We focus on three kinds of decisions. First, we highlight two methodological choices with analytical implications: whether to use levels of social spending or rates of change and whether to study differences between countries or within each country over time. Second, we consider problems with data—including levels of government used and the periods

considered—and the different ways in which they have been resolved. Third, we also discuss more complex econometric problems (e.g. model specification, the characteristics of the error terms and endogeneity) and the biases that different choices can introduce in the results.

Given the limitations of econometric analysis, at the end of the paper we call for the complementary use of qualitative and quantitative research. This goal goes beyond the use of mixed methods—a technique several of the authors studying social policy in Latin America have already done (e.g. Huber and Stephens, 2012; Kaufman and Segura-Ubiergo, 2001)—and refers to the importance of having more conversations between qualitative and quantitative researchers about theory, data and the causal stories we want to develop.

What We Know about Determinants of Public Social Spending in Latin America

The study of social policy occupies a central role in political economy (Amenta, 2003). As Doyle (2018: 1) puts it, "given that distributional battles lie at the heart of politics, it is perhaps not a great surprise that one of the most researched areas of political science involves work on social policy and social spending". How much different countries spend on social policy and how they shape their welfare states tells us much about what their institutions, class balance and economic prospects (Amenta, 2003; Mkandawire, 2006). The level (and composition) of social spending is also fundamental for poverty reduction and the promotion of more equitable societies. Given its high levels of inequality, reaching agreements on the determinants of social policy and identifying the best ways to expand social spending in the future is particularly important in the Latin American context (World Bank, 2002; Sánchez-Ancochea, 2020).

The literature on the determinants of social policy first developed in the context of developed countries and emphasized the role of economic liberalization—inspired by Karl Polanyi's ideas on protective counter-movements in *The Great Transformation*—as well as

trade unions and political parties (e.g. Cameron, 1978; Garrett, 2001; Hicks and Swank, 1992; Katzenstein, 1985; Korpi, 1978). In the mid-1990s, this literature was expanded to consider the contradictory impacts of globalization.

As a result of successive waves of research, we can identify three main hypotheses: the trade and globalization hypothesis, the modernization hypothesis, and the politics hypothesis. First, after finding robust and positive partial correlations between trade openness and government expenditure, Dani Rodrik (1998) stimulated an intensive debate on the relationship between the two. In his view, government expenditure is a risk compensation mechanism for citizens in open economies (compensation hypothesis)--an argument that seems particularly valid for OECD countries (Doyle, 2018). In contrast, Garrett (2001) argued that when capital mobility is large, increases in trade could lead to pressures for a smaller rather than a bigger government (efficiency hypothesis). According to Wibbels (2006), the efficiency hypothesis may be particularly relevant to developing countries, which are capital-constraint and face more obstacles to borrowing. While these initial contributions were concerned with overall government expenditure, the debate later focused on social spending and its components.

Second, the modernization hypothesis assumes a positive link between the level of social spending and GDP per capita. As countries become more developed, both social demands and social needs increase, thus resulting in the expansion of social services. This hypothesis—developed around Wagner's Law on public spending—has received attention in the context of OECD countries (see e.g. Williamson and Fleming 1977). Drawing on this argument, some authors have also argued that a growing urban population may drive higher spending, since urbanization comes together with industrialization and labor organization (Avelino, Brown, and Hunter, 2005).[2]

Third, various studies highlight the role of political institutions and party ideology. The literature on OECD countries focused on the contribution of left-wing parties, trade unions and various institutional arrangements to higher social spending (Hicks and Swank 1992; Huber, Ragin, and Stephens, 1993; Huber and Stephens 2001). There was less attention to the role of democracy since it was considered the only "game in town" in the developed world.

Since the early 2000s, the study of the determinants of social spending extended to Latin America.[3] The region constitutes an excellent case for researchers interested in "building on extant theory and developing mid-range theories of welfare state development across regions" (Huber, Mustillo, and Stephens, 2008, 420). On the one hand, it has stronger welfare institutions and higher levels of spending than other developing countries. On the other hand, it is poorer and more institutionally diverse than the U.S. or Europe. There are thus reasons to treat social policy in Latin America as a unique body of research with its own explanatory variables (Doyle, 2018).

For this exploration, we juxtapose the results obtained by six studies. The selection criteria were as follows. We aimed for papers that study the determinants of public social spending in Latin American in light of the three main hypotheses - the trade and globalization hypothesis, the modernization hypothesis, and the politics hypothesis - and that were published after 2000. Studies before the 2000s are less sophisticated econometrically and have more data limitations, and were therefore excluded. Further, since we want to evaluate papers that "talk to each other" explicitly, we excluded research that focuses on other political economy variables and processes, such as Doyle's (2015) excellent work on the impact of remittances. We decided to focus on Latin America exclusively and avoid cross-national studies (such as Haggard and Kaufman, 2004 or Wibbels, 2006) for two of reasons. First, some of the determinants of social policy are likely to be region-specific, depending on

particular histories, institutions and economic models. Second, this decision allowed us to

focus on a smaller number of papers, making comparison easier. Finally, we did not include

two influential books (Huber and Stephens, 2012 and Segura-Ubiergo, 2007) because their

econometric chapters are extensions of two of the papers we review. To the best of our

knowledge, there exist six original papers that fulfill our criteria (see table 1). In any event, it

is not our aim to provide a full review of all this literature, but to illustrate a number of

problems arising in it, using these articles as an illustration. There is no reason to assume that

including additional articles would change our evaluation.

Table 1: Countries and years analyzed in reviewed studies

Authors	Country sample	Period
Kaufman and Segura-Ubiergo (2001)	14 countries (Argentina, Bolivia, Brazil, Chile, Costa Rica, Dominican Republic, Ecuador, El Salvador, Guatemala, Mexico, Paraguay, Peru, Uruguay, and Venezuela)	1973–1997
Avelino, Brown, and Hunter (2005)	19 countries (Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Dominican Republic, Ecuador, El Salvador, Guatemala, Honduras, Jamaica, Mexico, Nicaragua, Panama, Paraguay, Peru, Uruguay, and Venezuela)	1980–1999
Huber, Mustillo, and Stephens (2008)	18 countries (same as Avelino, Brown, and Hunter (2005) except Jamaica)	1970–2000
Zarate Tenorio (2014)	18 countries (not listed)	1970–2007
Niedzwiecki (2015)	10 countries (Argentina, Bolivia, Brazil, Chile, Colombia, Ecuador, Paraguay, Peru, Uruguay, Venezuela)	1980–2010
Martín-Mayoral and Fernández Sastre (2017)	17 countries (same as Avelino, Brown, and Hunter (2005) except Dominican Republic and Jamaica)	1990–2012

Trade and globalization

Much of the global cross-country literature argues in favor of the efficiency

hypothesis, finding that trade and capital account liberalization reduce a country's ability to

tax and spend (Wibbels, 2006; Wibbels and Arce, 2003). Several studies on Latin America confirm this emerging consensus. For example, Robert R. Kaufman and Alex Segura-Ubiergo (2001, henceforth KS) find that trade openness has a negative impact on total social spending and social security but no effect on health and education. Sara Niedzwiecki (2015, henceforth NZ) also finds negative effects.

Yet this result is by no means undisputed. In fact, George Avelino, David S. Brown, and Wendy Hunter (2005, henceforth ABH) find a positive effect on social insurance and education, but not on health. Evelyne Huber, Thomas Mustillo, and John D. Stephens (2008, henceforth HMS) find no statistically significant links between various types of social spending and trade openness. Barbara Zarate Tenorio (2014, henceforth ZT) confirms that neither levels nor changes of trade openness are related to public social spending. The evidence thus tend to confirm the negative impact of globalization—a result that also make sense theoretically given Latin America's high level of dependence—but even here there are doubts.

Modernization

The evidence on the significance of the modernization hypothesis in the Latin American context is mixed. The impact of GDP per capita on social spending is contradictory and inconsistent across studies. For example, in HMS (2008), GDP per capita has a positive but small impact on health and education, but none on overall social security and welfare—opposite to what ZT (2014) finds. For social security spending, results vary from no statistical significance to positive and even negative impacts. In the latter case, some differences may stem from different foci: for example, negative relationships over time (ABH, 2005; NZ, 2015) do not contradict the absence of a positive relationship in cross-country comparison. But contradictions still remain. Two studies estimate levels and changes separately and still find opposite results (KS, 2001; ZT, 2014).

Contrary to the literature on the OECD, Latin Americanists pay particular attention to the role of democracy, which is expected to exert a positive role in social policy (Martínez Franzoni and Sánchez-Ancochea, 2016). Elections force political parties to compete for votes among the poor and middle-income groups—most of which benefit from higher spending. Democratic institutions also open new space for social contestation and social demands through mass media.

In the reviewed studies, democracy has a positive effect on the level of social spending but not on changes from year to year. This is evident when comparing, for example, HMS (2008) with KS (2001). The first study focuses on explaining the levels of social spending as percentage of GDP and finds that the number of accumulated years of democracy is positively related to spending in social insurance and health plus education—a result confirmed by NZ (2015), as far as within-country developments are concerned. In contrast, KS (2001)—like ZT (2014) and Fernando Martín-Mayoral and Juan Fernández Sastre (2017, henceforth MMFS)— find no impact or even negative effects of democracy on social spending, depending on specifications.

All this econometric work on Latin America makes useful references to each other and provides new understandings of the determinants of social spending. Authors do read, cite, and engage with previous works. Yet the accumulation of knowledge—important to devise future research agendas and make policy recommendations—has been limited. As reflected in table 2, the six studies provided a diversity of interesting results, but few clear consensuses.

Type of public social spending	Explanatory variable / measured in levels or changes	Kaufman & Segura-Ubiergo (2001)[1,2]		Avelino et al. (2005)		Huber et al. (2008)[1]	Zarate Tenorio (2014)[1]		Niedzwiecki (2015)[3]		Martín-Mayoral & Fernández-Sastre (2017)[4]	
		levels	*changes*	*levels*	*changes*	*levels*	*levels*	*changes*	*levels*		*levels*	*changes*
									within-country	*between-country*		
Total		(/)	(/)	(+)	not tested	not tested	not tested	not tested	(+)	(/)	(+) and (-)	(/) and (/)
Education	*Democracy*	(+)	(/)	(+)	not tested	(+)	(/)	(+)	not tested	not tested	not tested	not tested
Health		(+)	(/)	(/)	not tested	(+)	(/)	(+)	(+)	(/)	not tested	not tested
Social security		(-)	(/)	(/)	not tested	(+)	(-)	(/)	(+)	(+)	not tested	not tested
Total		(+)	not tested	(/) or (-)	(/) or (-)	not tested	not tested	not tested	(/)	(+)	(/) and (/)	(+) and (-)
Education	*GDP per capita*	(+)	not tested	(/)	(/)	(+)	(-)	(/)	not tested	not tested	not tested	not tested
Health		(+)	not tested	(+)	(-)	(+)	(-)	(/)	(/)	(+)	not tested	not tested
Social security		(+)	not tested	(-)	(-)	(/)	(-)	(-)	(-)	(/)	not tested	not tested
Total		(-)	(-)	(+)	not tested	not tested	not tested	not tested	(-)	(+)	(/) and (/)	(/) and (+)
Education	*Globalization and trade openness*	(/)	(/)	(+)	not tested	(/)	(/)	(/)	not tested	not tested	not tested	not tested
Health		(/)	(/)	(/)	not tested	(/)	(/)	(/)	(-)	(/)	not tested	not tested
Social security		(-)	(-)	(+)	not tested	(/)	(/)	(/)	(-)	(+)	not tested	not tested

Results: (+) = positive and statistically significant coefficient; (-) = negative and statistically significant coefficient; (/) = no statistically significant coefficient.

[1] Health and education spending estimated jointly.

[2] Results are reported where spending per capita is the dependent variable.

[3] Reported results are those that use separate estimators for within-country and between-country variance (table 8 in original paper).

[4] Results reported are from system GMM estimator. Results refer to two different periods that were analysed comparatively: 1990-2000 and 2000-2012.

As such, this body of research constitutes a clear example of our broader claim about the problems of applied econometric research to provide definite conclusions. In the rest of the paper we try to explain why the results are contradictory, focusing on three sets of reasons. Some contradictions can be explained by differences in key analytical decisions, which are not always sufficiently acknowledged. Other differences are related to data. A final set of factors has to do with the econometric techniques used.

Diversity of Analytical Approaches

Let us begin with two kinds of decisions that are often discussed in technical terms but have significant analytical implications. First, the determinants of rates of change in social spending are likely different than those affecting levels of spending. It is thus not surprising that studies that use one or the other end up with different results. Second, the use of country fixed effects (FE), which has become standard to account for unobserved heterogeneity between countries, also affects what researchers analyze and should have more influence on the way they interpret their results.

Levels vs changes

Some of the studies reviewed in this paper focus on the level of social spending, while others concentrate on changes over time. Likewise, several independent variables are sometimes used in levels and other times in changes. The decision about levels or changes is often presented in technical terms, i.e. as a way to solve various econometric problems such as serial correlations in the error terms (MMFS 2017), while theoretical implications are often only superficially explored. However, the use of levels or changes implies different research questions, the answers to which are not easily compared.

Spending levels tell us something about the level of public welfare and allow comparisons across countries or with other types of government spending. On the other hand, researchers might be interested in understanding how, for example, a particular shock (e.g. a commodity boom) affects year-to-year changes in public social spending. Unfortunately, these differences are often not considered, and levels or changes are sometimes mixed up. For example, KS (2001) discuss the efficiency and compensation hypothesis of globalization in terms of levels, arguing that "in OECD countries this hypothesis is supported by studies that show a very strong empirical association between economic openness, large public sectors, and generous welfare systems" (KS, 2001, 557). Their assumption is thus that more open countries spend more on social spending, which is quite different than saying that opening up the economy will lead to year-to-year expansion of social programs. Yet they then estimate a model where the dependent variable is changes in social spending.

There is a similar problem in the case of the independent variables: should we assume that the variance of social spending across countries is determined by the same variables than changes across time? The answer is obviously no. As HMS (2008, 421) explain for the case of political variables, "we would not expect one year of democracy or of dominance of one political tendency or another in the legislature and/or the executive to make a major difference in the formation of social policy". There are no reasons to assume that an improvement in the quality of democracy will result in an immediate expansion of social spending. More competitive elections or better electoral tribunals will lead to more attention to voters' preference and can result in more redistributive policies, but the process will take years to materialise. Equally, GDP per capita is likely to affect social spending over the long run (since the amount spent by the government will depend on the overall resources available), but its influence on annual changes is less clear. Economic growth may have a larger effect in changes in social spending than in its level. As a result, the various studies

reviewed are less comparable than initially thought: small changes in the equation used (e.g. introducing differences) can have major implications in the results and in their interpretation.

<u>Differences between and within countries</u>

To account for unobservable country-level characteristics and avoid biased estimators, most studies introduce country FE. As the term "FE" is understood differently in different contexts (Wooldridge 2010, Kropko and Kubinec 2020), we clarify that we use it to refer to a model with case-specific intercepts. Such a model can be thought of as containing dummy variables for each country. Coefficients estimated using country FE reflect the over-time impact of each independent variable averaged across countries. The estimator can be derived via the means-centering approach (e.g. Wooldridge 2010) or the – equivalent – data sub-setting approach (Kropko and Kubinec 2020), both of which lend themselves for intuitive exposition. In the former, the FE estimator subtracts the mean across observations within each country so that the remaining variation comes only from variations within each country over time. In the latter, the regression is first performed country-wise (i.e. all variation takes place over time), before the FE coefficient is calculated as a weighted average of all country-specific coefficients. Both approaches imply that the results of regression analysis with country FE should be interpreted as the effect of changes in the explanatory variable on the dependent variable within countries over time. In contrast, they tell us nothing about differences in social policy between countries, because this dimension of the variation in the cross-sectional data has been removed.[4]

The decision to implement country FE is most often driven by the desire to address unobserved heterogeneity in panel data (see also next section), and not by the authors' wish to concentrate their analysis on within-country developments.[5] For example, in discussing the role of political regimes, MMFS (2017, 7) wonder "whether authoritarian or democratic

regimes have different levels of social spending", even though their analysis will not allow them to say anything about this. ABH (2005, 628) present their hypotheses in a similar way: "H1: Latin American governments in open economies spend more on social programs than do governments in relatively closed economies. H2: Democracies in Latin America spend more on social programs than do authoritarian regimes". These are comparative statements about different countries at a certain moment in time, which cannot be evaluated through an econometric technique that considers changes within countries over time.

Diversity of Sources

The second set of factors that we deem responsible for diverse and partly contradictory results of the literature refers to data. Here we focus our discussion on the dependent variable, but also make a few remarks about independent variables.

Data sources and measurement: public social spending

Some of the differences in the results may be due to the use of different data sources. ABH (2005) and MMFS (2017) use Economic Commission of Latin America and the Caribbean (ECLAC) data, while KS (2001) rely on information from the International Monetary Fund (IMF). ZT (2014) and NZ (2015) use the Social Policy in Latin America and the Caribbean Dataset elaborated by Evelyne Huber and John D. Stephens, which is currently available in its 2014 version.[6] This database combines data for social security and welfare spending, education and health spending from different ECLAC and IMF sources.[7]

The level of government included in each database is different. As KS (2001) indicate, the IMF data is available at central government level only. As to ECLAC data, levels of coverage have changed over time. ECLAC statistics for the period 1980–1990— based on Rossella Cominetti and Gonzalo Ruiz (1998)—covers spending by the central

government with a few exceptions. Brazil is the only country where general government spending is available, while the cases of Argentina and El Salvador include non-financial public-sector spending (Cominetti and Ruiz, 1998, 24–25).

Since 1991, ECLAC data appears in their annual flagship report *Social Panorama of Latin America* and the institutional coverage still varies by country. Bolivia (1996–2014), Brazil (2000–2015), Colombia (2009–2015), Cuba (1996–2015) and Peru (1999–2015) only report information for the central government. Non-financial public sector spending, which comprises spending for the central government, subnational governments, and non-financial public corporations, is available for El Salvador (2002–2015) and Mexico (1990–2015); public sector spending for Costa Rica (1987–2015) and Peru (1999–2015).

The use of central government spending is particularly problematic in countries with federal structures, where social spending is decentralized. In Brazil, for example, the federal government spent less than 60 percent of public social spending, while state governments and municipalities were responsible for 23 percent and 20 percent, respectively (ECLAC, 2006, 127). Additionally, since the beginning of our period of study, many Latin American countries have undergone decentralization reforms. As a result, even in non-federal countries like Bolivia or Colombia, sub-national governments account for over 70 percent of public spending in education and about 50 percent in health (Brosio and Jiménez, 2012). Not surprisingly, using data from different levels of government can lead to erroneous conclusions (Martínez and Paz Collinao, 2010, 26).

Figure 1 illustrates the differences in the values of spending and its change over time for countries for which more than one level is available from ECLAC data. For example, in Argentina there was a drop in non-financial public sector spending around 2007, but no changes in central government spending. There are also marked differences in Brazil, where the evolution of general government spending and central government spending has been

different over the whole period. Comparing spending at different levels thus poses problems

for both cross-country and within-country comparisons.

Figure 1: ECLAC public social spending data 1990–2015, by different levels of government

spending
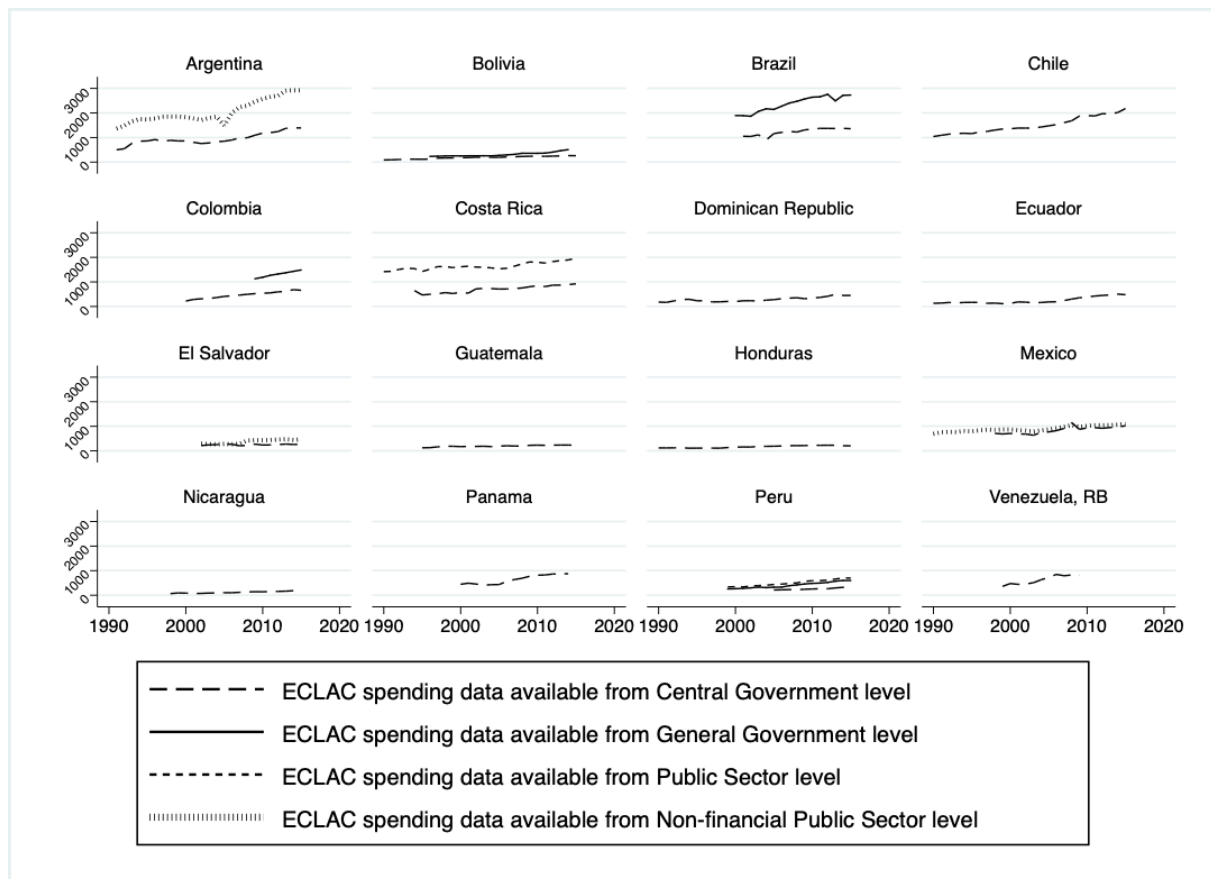
Source: own elaboration with ECLAC data

With the Social Policy in Latin America and the Caribbean Dataset, Huber and

Stephens make an outstanding effort to construct the most comprehensive database possible.

To do so, they combine four different ECLAC sources with IMF data. Unfortunately, by

merging data from these different series, their information may end up combining central

government spending at the beginning of the period with more encompassing levels later on.

How should researchers deal with this problem? If at all possible, we should only use data that fully reflects a country's effort in social policy. This would require considering information for all levels of government in federal and highly decentralized countries. We should also change sources depending on the circumstances: for example, for a country like Bolivia, it may be fine to use central government data prior to the late 1980s but better to use data on the whole public sector for more recent periods.

More broadly, this problem calls for a new approach to the construction of databases by international institutions. At the moment, there is a lot of information available for some indicators—for example, each international institution seems to have its own information on pensions—and not enough on others. Ideally, the World Bank, ECLAC and other institutions should collaborate to produce the most comprehensive historical data possible on various social spending indicators (see Martínez Franzoni and Sánchez-Ancochea, 2018 for a similar plea in the case of indicators on universalism). Working together and pooling resources, a better database could be easily built, improving quantitative (and qualitative) research significantly.

Data sources and measurement: explanatory variables

Some of the key explanatory variables are also measured in different ways and come from a variety of sources. Consider, for example, the case of trade. NZ (2015) and some others uses trade openness data from the World Bank's *World Development Indicators*. In contrast, ABH (2005) argue that measures of trade based on real exchange rates underestimate the size of some economies and propose a trade measure based on purchasing power parity instead.

Democracy has also been measured in two ways: as a yearly binary dummy variable and as the number of years of democracy that a country has accumulated over time. ABH

(2005) rely on a binary distinction between democratic and authoritarian regimes —checking for robustness with continuous Polity data. They motivate their choice by explaining that they understand democracy as "fundamentally distinguishable" from other regimes - apparently implying that this distinction can be properly expressed by a yearly dummy. HMS (2008), in contrast, rely on cumulative years of democracy from 1945 onwards and argue that this measure is able to express the "strength of the democratic record", likely implying that democracies become stronger, more stable or more impactful over time. KS (2001), MMFS (2017) and ZT (2014) use binary measures as yearly dummies, while NZ (2015) uses cumulative years of democracy based on a binary measure; these four papers only state their choice without motivating it theoretically.

The use of these distinct measures of democracy imply different theoretical understandings of the role of democracy for public social spending: even though it is not always clearly expressed, it seems that authors who use an annual score of democracy assume that the political system has an immediate impact on policy variables, while those that use accumulated years of democracy have a more complex understanding of how institutions work. Note also that besides the two common measures used by our authors, there exist other, less minimalist measures of democracy, which may convey different information still.[8]

While we do not wish to identify any one measure as superior, we emphasize that any measure chosen should be a good fit for the theoretical mechanism under study. The use of different measures of democracy, as well as the use of more complex measures, clearly comes at a cost – the lack of direct comparability. Yet it may nevertheless enhance our understanding of the role played by democratic institutions for public social spending as long as authors motivate their choices clearly and carefully discuss implications of their results with a view to the measure used.

Another difference that should be mentioned refers to periods of analysis (see table 1). Decisions on the period of analysis are likely to exert a major influence on the results but receive insufficient attention. MMFS (2017) conduct separate regressions for two different periods and obtain different results. This is not surprising: variables like commodity exports likely had a different impact on public social spending during the commodity boom of the 2000s than at other times. In fact, studies in other areas have shown that the impact of political processes varies significantly, depending on the period under study. For example, Carina Schmitt (2016) shows how periodization changes empirical results about partisan effects on policy.

Theoretically, it is clear that studies conducted upon a specific time-series should not be used to extrapolate beyond the period of the sample, but this is sometimes forgotten. Moreover, it would be useful to motivate the choice of time periods not only based on data availability, but on theoretical grounds. Here, econometric theory does not come with a user's guide: the choice of periods under study needs to be made based on case knowledge. Otherwise, the analysis could produce results that only hold for a sub-period, mask results from one period that do not hold for others, and so forth.[9]

Decisions on Econometric Techniques

All studies we discuss in this paper use time-series cross-section (TSCS) data. TSCS data come along with a number of features that contradict basic assumptions from the canonical ordinary least squares (OLS) model. This gives rise to a set of problems, among them unobserved heterogeneity, non-stationarity, endogeneity, and serial correlation, heteroscedastiscity and contemporaneous correlation of the error terms. The papers we review follow two general strategies to resolve some of these issues (see table 3): one group

uses OLS panel estimators with country FE, and a second group uses error-correction models (ECM).

Table 3: Overview of estimation techniques used in six different studies

Authors	Estimation technique
ABH (2005)	OLS with country FE and error-correction model (ECM) with OLS estimates and FE
HMS (2008)	Pooled TSCS, generalized Least Square Estimators without FE
Niedzwiecki (2015)	OLS with FE and panel model that estimates within-country and between-country coefficients separately
KS (2001)	ECM with OLS estimates and country FE
MMFS (2017)	Various techniques in comparison, including OLS with country FE and ECM with GMM estimates
Zarate Tenorio (2014)	ECM with OLS estimates and country FE

Unobserved heterogeneity

Unobserved heterogeneity is probably the primary reason to introduce country FE. Unobserved heterogeneity occurs when countries' levels of public social spending differ in ways that are not explained by the independent variables included in the estimation. For example, it could be that a specific country has a political tradition of high public social spending levels. If this tradition is not accounted for, higher levels of social spending would be—falsely—attributed to other, included factors. Unobserved heterogeneity thus results in biased estimators (Rabe-Hesketh and Skrondal, 2008; Raudenbush and Bryk, 2002).

The use of country FE offers no one-size-fits-all solution and has the problems previously discussed. HMS (2008) prefer to use pooled data precisely to explore country differences. Unfortunately, their solution does not properly deal with unobserved

heterogeneity and may result in biased estimators. These problems are illustrated by MMFS (2017), who provide estimates from a variety of techniques, among them pooled and FE OLS. Their results show marked differences between estimated coefficients and corresponding standard errors.

Non-stationarity

Many of the macro variables used in the studies we review likely exhibit non-stationarity (see e.g. Phillips and Moon, 2000), and some variables may be co-integrated. When time series exhibit non-stationarity in levels—and this is clearly the case with public social spending in Latin America or GDP per capita—, spurious regressions can be the result (Granger and Newbold 1973, Entorf 1997). In the presence of non-stationarity and/or co-integration, OLS estimators are not appropriate, and alternatives such as fully modified estimators or dynamic OLS estimators should be considered (e.g. Kao and Chiang 1999, Baltagi 2008). Where co-integration is present, co-integration approaches in the spirit of Engle and Granger (1987) may provide a viable solution.

In our sample of studies, only three papers address issues of non-stationarity and co-integration, at least partly or indirectly. KS (2001 and ZT (2014) use ECMs even though this choice is not motivated through non-stationarity concerns. MMFS (2017) are the only ones to actually test for the presence of structural long-run relationships in their time series, employing unit root and then co-integration tests. In the two other cases, it is not made clear whether the choice of an ECM was appropriate, given the time series characteristics of the data. Overall, econometric requirements of the specific data employed are hardly ever discussed in the papers we review.

<u>Serial correlation, heteroscedasticity, and contemporaneous correlation of the error terms</u>

TSCS data typically exhibit a number of cross-sectional and temporal dependencies. First, serial correlation of the error terms occurs when temporal dependencies exist in observations over time. Second, in the case of heteroscedasticity, error terms have a constant variance within, but not across countries. Third, there may be contemporaneous correlation of errors across countries, e.g. through a common shock. Different types of dependencies may easily but not necessarily occur together. Ignoring these dependencies can lead to biased inferences (e.g. Wooldridge, 2010).

Diverse methods have been proposed to adjust standard errors. It is important to note that different solutions are appropriate depending on the case of dependencies. For example, heteroscedasticity-robust standard errors can be used in the presence of heteroscedastic residuals. Clustered standard errors adjust for cross-sectional dependence of observations within clusters such as countries. The latter also accounts for temporal dependency, while the first does not (Hoechle, 2007). Moreover, different robust estimators have different requirements in terms of panel size or structures.

Nathaniel Beck and Jonathan N. Katz (1995) proposed panel-corrected standard errors (PCSE) that address heteroscedasticity, contemporaneous correlation and serial correlation of order 1. While many authors seem to have taken this procedure as a universal remedy for all sorts of situations, PCSE have their own problems. For example, they may be problematic when the panel data set consists of a rather small number of years, or in the presence of serial correlation (Reed and Webb, 2010).

Overall, our point is that the choice of appropriate standard errors is complicated, and there is hardly a one-size-fits-all solution. In practice, however, most papers in our sample employed PCSE without discussing whether PCSE was appropriate. They did not provide information on whether they tested for serial correlation, heteroscedasticity and contemporaneous correlation in their data to choose appropriate standard errors. This is not a

problem of the papers on social policy alone: Sven E. Wilson and Daniel M. Butler (2007) analyze the intellectual aftermath of Beck and Katz (1995) in the political science literature and conclude that "a nontrivial number of studies appear to be nothing more than a blind application of the method" (Wilson and Butler, 2007, 102), and others speak of a "de facto Beck-Katz standard" (Plümper et al. 2005, 327).

When reading the various papers, it is hard to know how exactly the problems were resolved and the implication that this may have for the estimated coefficients and standard errors. At the same time, the same technical solution may not always be the best for different papers, which means that the results are not necessarily comparable just because the same standard errors have been used.

Endogeneity

Endogeneity refers to situations in which the dependent variable is correlated with the error term. Endogeneity can be attributed to three types of causes: omitted variables, measurement error and simultaneity (Wooldridge, 2010). These constitute violations to the fundamental assumptions made in OLS estimations.

While many authors discuss omitted variable bias and measurement error in some form (recall the discussion about how to measure trade openness), simultaneity is less often addressed. For example, most papers in our sample measure the dependent variable as percentage of GDP, while also incorporating GDP as an independent variable on the right-hand side. As a consequence, the estimated coefficient for the effect of GDP is likely to be biased and may even change signs. Such an effect is possibly at work in KS's (2001) comparison of the results obtained using social spending as percentage of GDP and in per capita figures. When using the latter, they find a positive and statistically significant relationship between GDP per capita and public social spending per capita in accordance to

their theoretical expectations. Yet when using social spending as a percentage of GDP, most results are not statistically significant. If researchers want to incorporate GDP per capita on the right-hand side to test the modernization hypothesis, they should either use social spending figures per capita to measure the dependent variable or implement an estimation technique that takes care of this kind of endogeneity.

The assumption made about the absence of reverse or simultaneous causality in most papers is also problematic. For example, including the share of old-age population on the right-hand side of the equation can lead to biased estimators: a larger share of elderly people increases public social spending, but, at the same time, increased social spending could also increase life expectancy and thus the share of elderly people in the population. MMFS (2017) raise a potential simultaneity relation between social spending and economic growth: while a favorable economic situation could increase public budgets, social spending could also lead to higher growth, for instance because public spending in education and health increases human capital.

In our sample of papers, MMFS (2017) are the only ones dealing explicitly with endogeneity. They use system and differenced generalized methods-of-moments (GMM) estimators in order to control for potential endogeneity.[10] Yet even if the other papers considered endogeneity issues more explicitly, they could still resolve them in different ways. Furthermore, some simultaneous relationships are not immediately clear but are discovered in new research. Different considerations and expectations about potential simultaneity relationships can thus change results and render them incomparable.

In summary

The papers in our sample address technical issues in different ways, which has consequences for their results. Not all choices are equally satisfactory from a technical point

of view. In some instances, authors may have introduced biases in their estimators in their attempt to resolve some other problems. In other cases, the choice of econometric tools is inconsistent with the research question to be addressed. In these cases, it is the use of the tool rather than the tool itself that causes problems.

Still, and this is our central concern, even if all choices researchers make were technically and theoretically satisfactory, there is still some scope for incommensurability. We think that an important first step to deal with this incommensurability is that authors provide more transparent reflections of choices made. Such discussions could not only help readers to understand the extent to which results are comparable, but also ensure that choices made correspond to the authors' specific analytical goals. Further, results and findings should be presented more clearly: for example, rather than stating that trade openness benefits public social spending, authors should explain that public social spending increases in countries as these countries open up for trade, or that in cross-country comparison, countries with higher trade openness have higher spending levels.

Mixed Methods and Conversations across Methodologies as a Useful Response

Some of the problems we have identified in previous sections can be tackled by improving the methods (e.g. dealing with the standard errors appropriately) and the data researchers use. Yet, as shown at several points in this paper, many econometric challenges do not have an easy cure and contribute to the diversity of results. How can we deal with this situation? How can we advance more quickly in our understanding of the determinants of social policy? Although there is no single answer to these questions, we believe that more active conversations between qualitative and quantitative researchers would be particularly useful.

This is, of course, already done by researchers using mixed methods in a single research project. For example, in their recent review of mixed methods in the study of welfare regimes, Sara Niedzwiecki and David Nunnally (2017, 1028) argue that "incorporating multiple methodologies in a single research design has the potential to significantly advance knowledge". Let us illustrate how this interaction between methods has already taken place with a few examples that use econometrics to identify correlations and case studies to determine how the causal chain unfolds within and between countries. For example, Huber and Stephens (2012)'s econometric analysis focuses on the correlation between political regimes and left-wing parties in governments and social spending in various areas. They then select five cases within the region to evaluate how democracy and the left operate in practice, and to identify omitted variables. Segura-Ubiergo (2007) follows a similar methodology, although he focuses on three countries with similar levels of development but different degrees of openness and political institutions.

Case studies may also precede and inform subsequent econometric research. Niedzwiecki's (2018) *Uneven Social Policies* helps us explain why the relationship between politics and social policy at the national level can vary depending on the type of policy and a country's level of decentralization. Focusing on the behavior of subnational governments in Argentina and Brazil, she shows that subnational politicians will act differently depending on the characteristics of the policies. When the attribution of responsibility is clear—that is, when voters have no doubt that the central government is responsible for the new program— subnational governments ruled by the opposition will be reluctant to implement it. This is exactly what happened with *Asignacion Familiar por Hijo* and *Bolsa Familia*—Argentina's and Brazil's conditional cash transfers—in the 2000s. Niedzwiecki's insights could inform future econometric research in at least two ways. First, they highlight the need to consider the relationship between level of decentralization and subnational government ideology through

interaction effects. Second, researchers can use her work to determine when using central government statistics is appropriate and when it is not.

Candelaria Garay's (2016) *Social Policy Expansion in Latin America* is another recent book that could inform quantitative research and explain some of the confusing results we observe.[11] She uses the experience of Argentina, Brazil, Chile and Mexico to discuss the reasons behind different levels of expansion in pensions, income support and health care programs. Her work emphasizes the central role of outsiders (i.e. workers in informal jobs without social security benefits) in contemporary Latin America. In her view, it is not democracy per se, but the level of electoral competition and the presence of social mobilization what determines social policy expansion. Her work invites quantitative researchers to consider electoral results and the relative share of insiders and outsiders when considering the growth of social spending. Her research also highlights the importance of coverage as an alternative measure of social policy expansion.

Yet we are not just calling for more mixed methods research, but for more and better interactions between quantitative and qualitative researchers as well. In particular, even when not conducting a mixed-methods study themselves, quantitative researchers may benefit from actively drawing on qualitative studies. In *Multi-Method Social Science*, Jason Seawright argues that "integrating designs are a wonderful tool for evaluation and critiquing others' research, as well as for strengthening one's own causal inferences" (Seawright, 2016, 10). In his view, qualitative research can support regression analysis in at least three ways. First, it can contribute to better measurement (and theoretical understanding) of the dependent and independent variables. The above-mentioned study by Niedzwicki (2018) is a straightforward example of a study that may help reconcile contradictory results from studies that use different government levels of spending. Referring to the data and measurement problems discussed in this paper, qualitative knowledge could illuminate the extent data from different

government levels are comparable in different countries, or which political processes are at play in specific countries, indicating whether a level or change specification is more appropriate. Likewise, case knowledge may help to identify breaks in political processes over time, which could help to compare results obtained from different study periods.

Second, qualitative studies can test causal paths, thus illuminating the processes behind certain statistical correlations. This strength has been highlighted by authors like Lieberman (2005), who propose a "nested" approach where case studies illuminate some of the findings and assumptions of the econometric exercise. Given the high number of contradictory findings in the literature we review, insights from country case studies may be used to discriminate between different findings and suggest avenues for future model specifications.

Third, qualitative studies can also illuminate the role of certain omitted variables (Seawright, 2016, chapter 3). The appropriate specification of econometric models, in particular when it comes to control variables that should be included or not, requires theoretical and case knowledge. Mistaken ideas about variables to be left out or included have critical impacts on estimation results. Qualitative knowledge is crucial when setting up a specification and theorizing about relationships between relevant variables.

If the accumulation of knowledge is going to increase in this field (as in others in political economy), both qualitative and quantitative researchers should make an effort to design their work in a way that enhances communication. For quantitative researchers, this would require stating analytical goals and key decisions more clearly. Besides making their work more accessible for qualitative researchers, explicitly spelling out analytical goals could also improve the alignment of goals and technical decisions. Qualitative researchers could try to present their work (including its key dependent and independent variables) in such a way as to allow their findings to be assessed and expanded by quantitative researchers. Advancing

in this direction may be easier for researchers of Latin America's social policy than in other fields because they are part of a small but vibrant community where quantitative, qualitative and mixed method research is growing rapidly.

Conclusion

This paper has illustrated the difficulties that applied econometric research has to accumulate knowledge through the study of the literature on the determinants of public social policy in Latin America. We showed that the key studies do not offer any clear conclusion regarding the role of modernization, globalization and politics. Why is this the case?

One potential answer is that the underlying social reality is contradictory: for example, some variables may affect the level of social spending in some periods and not others, depending on how they interact with other processes. This is implicitly Doyle's (2018) view in a recent review of the literature on social spending and taxation in the region. From this perspective the goal of future research should be to elaborate new theories to account for differences in results and to add new studies.

Instead, we have argued that applied econometric research of this kind may have some inherent problems. We have focused on three factors to explain the diversity of results: different technical decisions that lead to differences in the analytical questions asked; differences in the data sources; and diverse estimation problems that can affect results. All these problems demonstrate that the craft of applied econometric work is messier and more ambiguous than often expected.

To be clear, this should not lead us to conclude that econometric research is unhelpful, but that it both can be done better and should not occupy a monopolistic position in political economy research. Instead, we need to do more to build more effective communication between qualitative and quantitative research, exploring with more detail

how one can inform the other. We should also acknowledge the problems of quantitative research (as we more often do with qualitative studies), spending more time building better databases, comparing econometric choices and explaining the implications of our research. These are all challenging tasks, but particularly important when dealing with a topic of such intellectual but also political relevance as the expansion of social policy in the most unequal region of the world.

References

Amenta, Edwin. 2003. "What We Know about the Development of Social Policy." In
Comparative Historical Analysis in the Social Science, eds. James Mahoney and Dietrich
Rueschemeyer, New York, NY: Cambridge University Press, 91-130.

Avelino, George, David S. Brown, and Wendy Hunter. 2005. "The Effects of Capital
Mobility, Trade Openness, and Democracy on Social Spending in Latin America, 1980–
1999." *American Journal of Political Science* 49 (3): 625–641.

Baltagi, Badi. 2008. *Econometric analysis of panel data*. Chichester: John Wiley & Sons.

Beck, Nathaniel, and Jonathan N. Katz. 1995. "What to Do (and not to do) with Time-Series
Cross-Section Data". *American Political Science Review* 89 (3): 634–647.

Bird, Alexander. 2007. "What is scientific progress?" *Noûs* 41(1): 64-89.

Brooks S. 2008. *Social Protection and the Market in Latin America*. Cambridge: Cambridge
University Press.

Brosio, Giorgio, and Juan P. Jiménez. 2012. 'Introduction: Approaching Recent
Transformations of Intergovernmental Relations from Multiple Profiles.' In
Decentralization and Reform in Latin America Improving Intergovernmental Relations,
ed. Giorgio Brosio and Juan P. Jiménez. Cheltenham, UK, and Northampton, MA: Edward
Elgar, 1–12.

Brunnschweiler, Christa N., and Erwin H. Bulte. 2008. "The Resource Curse Revisited and
Revised: A Tale of Paradoxes and Red Herrings." *Journal of Environmental Economics
and Management* 55 (3): 248–264.

Cameron, David R. 1978. "The Expansion of the Public Economy: A Comparative Analysis."
American Political Science Review 72 (4): 1243–1261.

Cominetti, Rossella, and Gonzalo Ruiz. 1998. *Evolución del gasto público social en América Latina, 1980–1995. [The Development of Public Social Spending in Latin America, 1980–1995.]* Cuadernos de la CEPAL 80. Santiago de Chile: ECLAC.

Doyle, David. 2015. "Remittances and Social Spending." *American Political Science Review* 109 (4): 785–802.

Doyle, David. 2018. 'Taxation and Spending in Latin America.' *Oxford Research Encyclopedia of Politics*. Available online at http://oxfordre.com/politics/abstract/10.1093/acrefore/9780190228637.001.0001/acrefore-9780190228637-e-632?rskey=Rbv4Meandresult=1.

Engle, Robert F., and Clive W.J. Granger. 1987. "Co-integration and error correction: representation, estimation, and testing." *Econometrica: Journal of the Econometric Society* 55(2): 251-276.

ECLAC. 2006. *Social Panorama of Latin America*. Santiago de Chile: ECLAC.

Entorf, Horst. 1997. "Random walks with drifts: Nonsense regression and spurious fixed-effect estimation." *Journal of Econometrics* 80(2): 287-296.

Garay, Candelaria. 2016. *Social Policy Expansion in Latin America.* New York, NY: Cambridge University Press.

Garrett, Geoffrey. 2001. "Globalization and Government Spending around the World." *Studies in Comparative International Development* 35 (4): 3–29.

Granger, Clive W. J., and Paul Newbold. 1974. "Spurious regressions in econometrics." *Journal of Econometrics* 2(2): 111-120.

Haggard, Stephan, and Robert Kaufman. 2004. Revising Social Contracts: Social Spending in Latin America, East Asia, and the Former Socialist Countries, 1980–2000. *Revista de Ciencia Política* 24 (1): 3–37.

Hicks, Alexander M., and Duane H. Swank. 1992. "Politics, Institutions, and Welfare

Spending in Industrialized Democracies, 1960–82." *American Political Science Review* 86

(3): 658–674.

Hill, Terence D., Andrew P. Davis, J. Micah Roos, and Michael T. French. 2019.

"Limitations of Fixed-Effects Models for Panel Data." *Sociological Perspectives* 63(3):

357-369.

Hoechle, Daniel. 2007. "Robust Standard Errors for Panel Regressions with Cross-Sectional

Dependence." *The Stata Journal* 7 (3): 281–312.

Holland, Alisha C. 2017. *Forbearance as Redistribution: The Politics of Informal Welfare in

Latin America*. Cambridge: Cambridge University Press.

Huber, Evelyne, Thomas Mustillo, and John D. Stephens. 2008. "Politics and Social

Spending in Latin America." *The Journal of Politics* 70 (2): 420–436.

Huber, Evelyne, Charles Ragin, and John D. Stephens. 1993. "Social Democracy, Christian

Democracy, Constitutional Structure, and the Welfare State". *American Journal of

Sociology* 99 (3): 711–749.

Huber, Evelyne, and John D. Stephens. 2001. 'Welfare State and Production Regimes in the

Era of Retrenchment.' In *The New Politics of the Welfare State*, ed. Paul Pierson. New

York: Oxford University Press, 107–145.

Huber, Evelyne, and John D. Stephens. 2012. *Democracy and the Left: Social Policy and

Inequality in Latin America*. Chicago: University of Chicago Press.

Kao, Chihwa, and Min-Hsien Chiang. "On the estimation and inference of a cointegrated

regression in panel data." *Advances in Econometrics*, 20: 179-222.

Katzenstein, Peter J. 1985. *Small States in World Markets: Industrial Policy in Europe*.

Ithaca and London: Cornell University Press.

Kaufman, Robert R., and Alex Segura-Ubiergo. 2001. "Globalization, Domestic Politics, and

    Social Spending in Latin America: A Time-Series Cross-Section Analysis, 1973–97."

    World Politics 53 (4): 553–587.

Korpi, Walter. 1978. *The Working Class in Welfare Capitalism: Work, Unions and Politics in*

    Sweden. London and Boston: Routledge and Kegan Paul.

Kropko, Jonathan, and Robert Kubinec. 2020. "Interpretation and identification of within-

    unit and cross-sectional variation in panel data models." *PLoS One* 15(4): e0231349.

Lieberman, E. 2005. "Nested Analysis as a Mixed-Method Strategy for Cross-National

    Research." *The American Political Science Research*, 99(3): 435-52.

Madrid, R. 2003. *Retiring the State: The Politics of Pension Privatization in Latin America*

    and Beyond. Stanford: Stanford University Press.

Martínez, Rodrigo, and María Paz Collinao. 2010. *Gasto social: modelo de medición y*

    análisis para América Latina y el Caribe. [Social Spending: A Model of Measurement and

    Analysis for Latin America and the Caribbean.] Serie manuales 65. Santiago de Chile:

    ECLAC.

Martínez Franzoni, Juliana, and Diego Sánchez-Ancochea. 2016. *The Quest for Universal*

    Social Policy in the South: Actors, Ideas and Architectures. Cambridge: Cambridge

    University Press.

Martínez Franzoni, Juliana, and Diego Sánchez-Ancochea. 2018. "Undoing Segmentation?

    Latin American Health Care Policy During the Economic Boom", *Social Policy &*

    Administration, 52(6): 1181-1200.

Martín-Mayoral, Fernando, and Juan Fernández Sastre. 2017. "Determinants of Social

    Spending in Latin America During and After the Washington Consensus: A Dynamic

    Panel Error-Correction Model Analysis." *Latin American Economic Review* 26 (10).

    Online publication. DOI:10.1007/s40503-017-0053-6.

Mkandawire, Thandika. 2006. *Social Policy in a Development Context*, Geneva: UNRISD

Niedzwiecki, Sara. 2015. "Social Policy Commitment in South America. The Effect of Organized Labor on Social Spending from 1980 to 2010." *Journal of Politics in Latin America* 7 (2): 3–42.

Niedzwiecki, Sara. 2018. *Uneven Social Policies. The Politics of Subnational Variation in Latin America*, New York, NY: Cambridge University Press.

Niedzwiecki, Sara, and David Nunnally. 2017. "Mixed-Methods Research in the Study of Welfare States." *PS: Political Science and Politics* 50 (4): 1028–1031.

Phillips, Peter C.B., and Hyungsik R. Moon. 2000. "Nonstationary panel data analysis: An overview of some recent developments." *Econometric Reviews* 19 (3): 263-286.

Plümper, Thomas, Vera E. Troeger VE, and Philip Manow. 2005. "Panel Data Analysis in Comparative Politics: Linking Method to Theory." *European Journal of Political Research* 44 (2): 327–354.

Rabe-Hesketh, Sophia, and Anders Skrondal. 2008. *Multilevel and Longitudinal Modeling Using Stata*. Texas: Stata Press.

Raudenbush, Stephen W., and Anthony S. Bryk. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Thousand Oaks, London, and New Delhi: Sage.

Reed, W. Robert, and Rachel Webb. 2010. "The PCSE Estimator is Good – Just Not as Good as You Think." *Journal of Time Series Econometrics* 2 (1): 1–26.

Rodrik, Dani. 1998. "Why do More Open Economies Have Bigger Governments?' *Journal of Political Economy* 106 (5): 997–1032.

Sánchez-Ancochea, Diego. 2020. *The Costs of Inequality in Latin America: Lessons and Warnings for the Rest of the World*. London: Bloomsbury.

Seawright, Jason. 2016. *Multi-Method Social Science. Combining Qualitative and Quantitative Tools*. Cambridge: Cambridge University Press.

Segura-Ubiergo, Alex. 2007. *The Political Economy of the Welfare State in Latin America. Globalization, Democracy, and Development.* Cambridge: Cambridge University Press.

Schmitt, Carina. 2016. "Panel Data Analysis and Partisan Variables: How Periodization does Influence Partisan Effects." *Journal of European Public Policy* 23 (10): 1442–1459.

Wilson, Sven E., and Daniel M. Butler. 2007. "A Lot More to Do: The Sensitivity of Time-Series Cross-Section Analyses to Simple Alternative Specifications." *Political Analysis* 15 (2): 101–123.

Wibbels, Erik. 2006. "Dependency Revisited: International Markets, Business Cycles, and Social Spending in the Developing World." *International Organization* 60 (2): 433–468.

Wibbels, Erik, and Moisés Arce. 2003. Globalization, Taxation, and Burden-Shifting in Latin America. *International Organization* 57 (1): 111–136.

Williamson, John B., and Jeanne J. Fleming. 1977. "Convergence Theory and the Social Welfare Sector." *International Journal of Comparative Sociology* 18 (3–4): 242–253.

Wooldridge, Jeffrey M. 2010. *Econometric analysis of cross section and panel data.* Cambridge, Massachusetts: MIT Press.

Zarate Tenorio, Barbara. 2014. "Social Spending Responses to Organized Labor and Mass Protests in Latin America, 1970–2007." *Comparative Political Studies* 47 (14): 1945–1972.

---

[1] For example, Thomas Plümper, Vera E. Troeger, and Philip Manow (2005) provide a critique of four frequent sources of problems in panel data analysis in comparative politics, and criticise that researchers too often rely on supposed one-size-fits-all solutions, even when these may not be adequate for their specific purposes. Another example is the empirical literature on the resource curse (e.g. Brunnschweiler and Bulte, 2008).

[2] Other demographic variables can also influence different types of social spending. A growing elderly population will exert an upward pressure on social insurance, while the number of children will affect spending on education. Although the evidence is mixed, in general the former effect seems stronger than the latter.

[3] Although there were some studies before the 2000s, they were less sophisticated econometrically and also had more data limitations. In this paper, we thus focus on papers published after 2000.

[4] For a more detailed exposition, the interested reader is referred to Kropko and Kubinec (2020), who decompose case and time FE estimators and discuss their interpretations very clearly.

[5] Kropko and Kubinec (2020) argue that this is actually a misunderstanding and misuse of FE models, as these do not offer a control for omitted variables or confoundedness in the same way as including a control variable would. While they do remove time-variant and/or case-invariant omitted variables, this is a side effect of the transformation that the use of FE introduces, and this transformation impacts on the research question that

is being studies, and hence on the interpretation of the estimated coefficients. See also Hill et al. (2019) for a discussion of how FE models as used in the social sciences do not always correspond to the research question they were built for.

[6] Zarate Tenorio (2014) used a version containing data from 1970-2000 and updated until 2008 using ECLAC data. Niedzwiecki (2015) used a version where data are available until 2006 and updated using ECLAC data.

[7] The dataset is available online here:
http://huberandstephens.web.unc.edu/files/2016/06/LA-Welfare-Jan2016.xlsx.
The codebook is available here:
http://huberandstephens.web.unc.edu/files/2016/06/Codebook_LAWelfare-Jan2016.doc

[8] For example, it is likely that the different indicators of democracy presented in the Varieties of Democracy database result in different estimators. The extent to which this is driven by theory or by measurement errors demands detailed attention in any econometric study.

[9] Dividing samples into different periods will often be problematic, as sample sizes might become very small.

[10] KS (2001) report that they also used alternative estimation techniques, among them GMM, without obtaining substantially different results.

[11] There are many other examples that we wished we could also review with some detail. For example, Pribble (2013) shows how the impact of the left on welfare policies will be contingent on the type of political parties (more or less programmatic) and their links to civil society. Her work, however, focuses on explaining universalism and not social spending.   researchers can avoid mechanistic understanding of policies. Alisha C. Holland (2017)'s study of forbearance and informal welfare calls for more attention to the informal mechanisms of social intervention and how they can have a crowding-out effect on public social spending. Of course, how to the measure these types of mechanisms quantitatively is a major challenge.