CRITICAL PERSPECTIVES ON GENDER AND POLITICS

Multilevel Analysis as a Tool to Understand the Spatio-Temporality of Gender

Niels Spierings, Radboud University

> Gender:
> "culturally specific characteristics associated with masculinity and femininity . . . a wide range of . . . social roles assigned to men and women *historically and cross-culturally*"
>
> (Hawkesworth 2013, 36, emphasis added)

> Intersectionality:
> "that race, class, sexual orientation, nationality, and gender are not discrete markers of difference but rather intersection social structures of inequality experienced by individuals *in specific social locations*"
>
> (Ewig and Ferree 2013, 442, emphasis added)

As illustrated by these definitions from the *Oxford Handbook of Gender and Politics*, context dependency is a defining characteristic of gender, since gender and intersectional inequalities differ spatially and temporarily in how they are organized. It is on this issue of context dependency where gender studies scholars often collide with (the application and interpretation) of regression-based statistical analysis in debates about quantitative methods (see Spierings 2012).

   In this reflection, I will briefly address where this tension comes from, focusing particularly on how (statistical) relationships vary in occurrence and strength between contexts. Next, I discuss three rather common archetypical responses to this tension that do little to resolve the issue. Subsequently, I focus on more profound solutions, including multilevel regression models, which challenge gender scholars to specify and develop their theoretical arguments on how and why relationships are context dependent. The different, bad, and better ways in which to deal with possible context dependency are illustrated by an assessment of their

1

prevalence in the 32 studies applying inferential quantitative methods to explaining individual-level (e.g., voters, MPs) characteristics, as published in this journal between 2011 and 2015 (40% of all research articles published in volumes 7–11).[1]

THE TENSION BETWEEN REGRESSION ANALYSIS AND CONTEXT DEPENDENCY

At the core of regression models is estimating to what extent two variables (controlled for other factors) vary together and how likely it is that this correlation holds beyond the sample of studied units (e.g., voters). This can be in tension with context dependency for two reasons.

First of all, the goal of regression analysis is to draw conclusions that are generalizable beyond the cases studied. Random sampling — assuming no sampling or response bias — assures that results can be generalized to the cases the sample was drawn from, such as the U.S. or a European population.[2] Our theories, however, are often formulated more generally with, for example, all industrialized Western democracies falling within a theory's scope. Generalizing results of a study beyond the population it samples from — for example, generalizing a U.S. sampled study to European countries — brings in the assumption that the core causal mechanisms are the same across contexts.

Second, the more cases that are included in the analysis, the higher the certainty (statistical significance) that the relationships found in the sample are also present in the larger population.[3] This makes pooling data attractive, as it increases "certainty." In such studies it is common (and good) practice to control for different average levels on the dependent variable across countries. However, this practice does not account for different *effects* across countries — context dependency (Spierings 2012). Models based on pooled data might present the average effect across countries accurately, but they do not show whether the effect is actually the same in each country. Moreover, different sample sizes can lead to the results being dominated by the relationship in one oversampled or populous country.[4]

1. An overview can be obtained from the author.
2. See Note 4.
3. Technically, the denominator of the standard error of the B-coefficient includes "n." The smaller the "s.e.," the higher the "p-value"; thus, the more certain results hold for the population sampled.
4. That an effect holds for the sampled population does not mean it holds for all subsets of that population. For instance, if a negative relationship between being married and voting progressive

FALSE SOLUTIONS

The tension between regression practices and context dependency is particularly relevant for gender studies, as we start from the notion that the social meaning and consequences of one's sex are different across countries and that the structural and cultural context (e.g., welfare state regimes and societal norms) shape this meaning.

The ways in which this issue is dealt with are diverse. Before turning to some real solutions, I discuss three archetypical ways of dealing with context dependency that hardly help to develop our understanding of its (gendered) dynamics, but of which derivatives and/or combination are regularly found, as illustrated with figures for the 2011–2015 volumes of this journal:

1. The simplest way to deal with the possibility of relationships being context dependent is ignoring the issue. It is surprising how many studies frame their conclusions in general terms while not reflecting on the possibility that the results might be different in other countries (or context) or between the countries studied, and for what reasons. In the 31–40% of all reviewed studies this was the case; all were single-country studies.[5]

2. The second archetypical solution is claiming that the case(s) studied is (are) representative or, in the case of pooled samples, that results are reliable because the cases are highly similar or because they are rather diverse — both arguments are used. However, given the differences in political systems, cultural norms, welfare state regimes, and economic structures, there is generally good reason to expect that the strength or even direction of a relationship differs between countries. Albeit, this practice is not uncommon; examples were found in 15% to 20% of the single-country studies that did not ignore the issue overall (see above) and in five of the six reviewed pooled studies.

3. The third response is the opposite of the second one: explicitly mentioning that the results are context specific and cannot be generalized. While this fits the argument that is made by some qualitative-oriented or country-specialist scholars (for instance, in the role of reviewer), these reservations are often rather empty. It does not provide a thought-through reasoning of why and

exists in a proportional sample of the U.S. (319 million inhabitants) and the Netherlands (17 million), it might be absent in the Dutch subsample given its more diverse party system. The 19:1 sample ratio will obscure this in the overall model. Similarly, among the 5% LGBT citizens, the opposite relationship might hold because newlyweds support progressive parties that made marriage possible for them. The overall results will veil this too.

5. All on American politics, none of the 10 other-country studies. Providing only case-specific background information is not considered discussing context dependency.

how the results can be expected to be different in other spatial or temporal situations, which the authors as experts in their field should be able to provide. Even though claims about "absolute context specificity" might not be expected in quantitative studies, traces of this archetype were found in 4 of the 26 single-country studies reviewed.

These three responses to potential context dependency are cosmetic or even dogmatic and tell us little about *how* the context influences the gender dynamics in a particular domain.

THEORIZING AND MODELING CONTEXT DEPENDENCY

The "solutions" to the issue of potential context dependency above lack theoretical specification and empirical testing. Here, I will discuss three better ways of dealing with context dependency when applying regression analyses.

Positioned statistical case study. There can be very good reason to study data on only one country. One of those reasons is actually that one has more space to go into detailed descriptive backgrounds regarding the context and the position the country takes in the population of countries (see Seawright and Gerring 2008). For instance, if a theory about political participation is said to hold for democracies, all democracies form the population. However, results of a study on a two-party democracy can be expected to differ from those of multiparty systems. Contextualizing the results of a single-country study and thinking through the implications of the contextual situation for the causal mechanisms specified in the theory provides a fairer and more reliable test of that theory. Or, if the mechanisms in the theory do not provide enough information for a priori expectations, the author can think through how the contextual circumstances have influenced the strength of the relationship found and generate new tangible expectations for other cases, even though these cannot be tested in that particular study. Three strong examples of positioned-statistical case studies among the reviewed articles were studies on Denmark (Bækgaard and Kjear 2012), Sweden (Folke, Freidenvall, and Rickne 2015), and the UK (Bicquelet, Weale, and Bara 2012).

Context-disaggregated analyses. If multiple datasets are available, pooled analyses can always be rerun per contextual unit to test how robust the pooled relationships are. Evidently, in disaggregate analyses the number of cases is lower, leading to fewer statistically significant relationships, but

coefficients per country can be compared to the coefficients from the pooled model on their substantial significance (see Miller and van der Meulen Rodgers 2008).[6] Even if surveys on only a few contexts (space or time) are available, this provides a preliminary assessment of context dependency or might generate theoretical ideas on which contextual characteristics matter (e.g., Spierings and Zaslove 2015). This approach is less common; among the reviewed studies, Schwindt-Bayer (2011) is the best example.

Multilevel-interaction models. If surveys on at least 25 contextual units (e.g., World Value Surveys; European Social Surveys) and data on these units are present, it is possible to systematically assess whether and why gendered relationships are different by context. The disaggregation approach discussed above cannot ascribe the differences to specific characteristics of the context, such as the public childcare budget, and it cannot rule out other moderating contextual characteristics. Multilevel regression models (MLM) do allow for this. MLM allows testing how variations in strength of microlevel relationships across higher-level units correlate with differences in contextual characteristics (e.g., whether variation in the employment-political participation linkage across years, countries, or states correlates with differences in welfare expenditures), without the risk of overestimating statistical significance. Alexander (2012) and Luhiste (2015) are clear examples from the *Politics & Gender* articles reviewed.

MLM is clearly the most demanding one in terms of data and analysis as well as the most systematic one. At the same time, disaggregated models in combination with case knowledge can be particularly useful for hypothesis-generating approaches or exploring unexpected results.

GENDER AND MULTILEVEL MODELS

Because context dependency is intrinsically part of the concept of gender, the gender studies literature can most easily embrace and push our theoretical understanding of context dependency. This can be done in multiple ways if the approaches above, MLM particularly, are embraced.

We should go beyond simply claiming that patriarchy and gender regimes take different shapes in different times and places, and start specifying the dimensions of patriarchy and theorizing and testing how they shape the gendered relationship at the individual level. For

6. Comparing coefficients across logistic models is more problematic (Mood 2010).

instance, how do different patriarchal institutions influence the impact of family formation on women's social participation (e.g., Spierings 2014). In other words, we should go beyond *claiming* context dependency and start *explaining* context dependency. This implies ruffling of old paradigmatic feathers that withhold a constructive discussion about using regression models for understanding context dependency (see Spierings 2012).

Second, while politics and gender studies tend to focus on how clearly gendered contextual variables, such as the proportion of women in Parliament (Alexander 2012), shape individual-level relationships, MLM also stimulates us to formalize, systematize, and demonstrate claims about ostensible gender-neutral structures having gendered implications for individual-level relationships. For instance, it is generally acknowledged that the type of electoral system has gendered implication for representation, but how do differences between PR and majoritarian systems influence *the impact of* individual characteristics on gender differences in political participation? MLM challenges gender scholars to further specify in what way gender neutrality is a myth.

Lastly, gender studies have a strong track record in qualitative methods, leading to vast in-depth case knowledge on gender dynamics. The field is therefor particularly suited to show how mixed-methods studies and projects can combine regression analyses with case studies in order to develop systematic theories on context dependency. For instance, disaggregated regression models can show for which countries a relationship does not hold. In-depth case knowledge can then provide explanations that are translated to more general expectations, which subsequently can be tested in multilevel models.

CONCLUDING NOTE

In this contribution, I have reflected on two practices: (1) the gender criticism of not acknowledging spatiotemporal context dependency in quantitative studies and (2) the application of the gender-based theoretical notion of context dependency in regression models. It is my hope that combining gender studies' in-depth situational knowledge with multilevel regression models will stimulate gender scholars to further develop their arguments on which aspect of the context actually influences which microlevel relationships and stimulate multilevel

modelers in political and other social sciences to reflect on and test the internal and external generalizability of their results.

Niels Spierings is Assistant Professor in Political and Gender Sociology at Radboud University, Nijmegen, The Netherlands: n.spierings@maw.ru.nl

REFERENCES

Alexander, Amy C. 2012. "Change in Women's Descriptive Representation and the Belief in Women's Ability to Govern: A Virtuous Cycle." *Politics & Gender* 8 (4): 437–64.

Bækgaard, Martin, and Ulrik Kjaer. 2012. "The Gendered Division of Labor in Assignments to Political Committees: Discrimination or Self-Selection in Danish Local Politics?" *Politics & Gender* 8 (4): 465–82.

Bicquelet, Aude, Alvert Weale, and Judith Bara. 2012. "In a Different Parliamentary Voice?" *Politics & Gender* 8 (1): 83–121.

Ewig, Christina, and Myra Marx Ferree. 2013. "Feminist Organizing: What's Old, What's New? History, Trends, and Issues." In *The Oxford Handbook of Gender and Politics*, ed. Georgina Waylen, Karen Celis, Johanna Kantola, and S. Lauren Weldon. Oxford: Oxford University Press, 437–61.

Folke, Olle, Lenite Freidenvall, and Johanna Rickne. 2015. "Gender Quotas and Ethnic Minority Representation: Swedish Evidence from a Longitudinal Mixed Methods Study." *Politics & Gender* 11 (2): 345–81.

Hawkesworth, Mary. 2013. "Sex, Gender, and Sexuality: From Naturalized Presumption to Analytical Categories." In *The Oxford Handbook of Gender and Politics*, ed. Georgina Waylen, Karen Celis, Johanna. Kantola, and S. Lauren Weldon. Oxford: Oxford University Press, 31–56.

Luhiste, Maarja. 2015. "Party Gatekeepers' Support for Viable Female Candidacy in PR-List Systems. *Politics & Gender* 11 (1): 89–116.

Miller, Jane E., and Yana van der Meulen Rodgers. 2008. "Economic Importance and Statistical Significance: Guidelines for Communicating Empirical Research." *Feminist Economics* 14 (2): 117–49.

Mood, Carina. 2010. "Logistic Regression: Why We Cannot Do What We Think We Can Do, and What We Can Do about It." *European Sociological Review* 26 (1): 67–82.

Seawright, Jason, and John Gerring. 2008. "Case Selection Techniques in Case Study Research: A Menu of Qualitative and Quantitative Options." *Political Research Quarterly* 61 (2): 294–308.

Schwindt-Bayer, Leslie A. 2011. "Women Who Win: Social Backgrounds, Paths to Power, and Political Ambition in Latin American Legislatures." *Politics & Gender* 7 (1): 1–33.

Spierings, Niels. 2012. "The Inclusion of Quantitative Techniques and Diversity in the Mainstream of Feminist Research." *European Journal of Women's Studies* 19 (3): 331–47.

——. 2014. "The Influence of Patriarchal Norms, Institutions, and Household Composition on Women's Employment in Twenty-Eight Muslim-Majority Countries." *Feminist Economics* 20 (4): 87–112.

Spierings, Niels, and Andrej Zaslove. 2015. "Gendering the Vote for Populist Radical-Right Parties." *Patterns of Prejudice* 49 (1–2): 135–62.