Operationalizing Intersectionality in Quantitative Research - What Do Classification Trees Reveal about Gendered Divisions of Housework?

Jenny Chanfreau

and

Wendy Sigle

Extended Abstract

MOTIVATION AND AIMS:

In the past several decades, feminist scholars have critiqued many aspects of the ways that social science research is conceptualized and carried out. Feminist family scholars challenged the conceptualization of the family as an institutionally "separate sphere" and feminist economists problematised the idea that people allocate their time to either work or leisure. They engaged directly with and sought to modify existing conceptual frameworks with the aim of ameliorating, if not eliminating, their errors and biases. Rather than an individual-level character trait, gender was conceptualized as a dynamic and multi-level system of social stratification with "institutions such as families, states, and markets as inter-connected sites rather than separate spheres or even discrete systems" (Ferree 2010, p. 425). In publications that coined the term intersectionality, Kimberle Crenshaw (1989, 1991) used examples from law and policy to demonstrate the theoretical and practical importance of longstanding black feminist critiques of the way categories were deployed and interpreted.

Feminist critiques and innovations have repeatedly been described as a stalled revolution (Stacey & Thorne 1985; Ferree 1990; Ferree et al 2007; Sigle 2021. Mainstream scholarship did not, for the most part, adopt and incorporate these new ideas and frameworks. Although one discernible "win" can be identified at the topic-level: studies of the household division of labour, went from being discouraged to being a routine (Ferree 2010), this research, <u>often remained embedded in conceptualizations based on "roles"</u> in which the issue of women's employment was taken as a matter of overcoming "old fashioned" attitudes to adopt more "modern" views consistent with the economy's demand for women workers…" (Ferree, Kahn, and Morimoto 2007). In 2009, Alexis Walker "deplored how little feminist research entered mainstream family journals, and how difficult she found it as editor [of Journal of Marriage and the Family] to bring in a critical, socially dynamic gender analysis….her overall conclusion was one of missed opportunities for the field of family studies…." Ferree 2010 pg. 422). A recent review of the content of textbooks (Zarza 2018) confirms that after decades of engagement and intervention, the contributions of feminists remain "not quite let into the canon" (Burawoy, 2005 pg. 6).

Scholars seeking to explain the stalled feminist revolution have tended to focus on structural factors within the discipline: older scholars had not been exposed to the new frameworks, gatekeepers in publishing did not insist that authors adopt new frameworks (Ferree et al 2007; Risman 2001). In a 2021 review of the state of theory in demography, Sigle (2021) suggested that another explanation for the stalled feminist revolution might be that feminist theoretical developments outpaced methodological developments, particularly in quantitative research: "When certain theories fall out of fashion, but the methods and approaches associated with them do not, we might find that we are inadvertently carrying some unwanted (and not very explicit) conceptual baggage". (pg. S238). More specifically, the uncritical and unreflexive attachment to linear regression models – when put alongside a reliance on secondary survey data in research (Crimmins 1993) – has limited the influence of intersectionality (see also Sigle 2016).

"The term intersectionality references the critical insight that race, class, gender, sexuality, ethnicity, nation, ability, and age operate not as unitary, mutually exclusive entities, but rather as reciprocally constructing phenomena" (Collins 2015, abstract). Intersectionality, then, problematises the use and interpretation of social categorical variables as additive and separable (for example, by introducing dummy variables for sex and for different ethnic groups) and draws attention to with-in group diversity. With linear regression models, such diversity can be captured *to some extent* by introducing interaction terms or by estimating fully interacted

models. Although recent reviews of quantitative intersectionality found regression to be the most common method (Bauer et al 2021; Guan et al 2021), previous scholars have outlined some limitations of this approach. Models with complex and high-level interactions can be difficult to present and interpret. McKinzie and Richards (2019) suggest that researchers attend to the context of the study in order to determine which intersections might be most meaningful, but the small size of some sub-groups, particularly in survey data, can seriously limit *which* variables and *how many* variables can be interacted in regression models (McCall 2005; Sigle-Rushton 2014). And this strategy assumes that there is sufficient contextual information to inform decision-making – in a world where non-intersectional thinking remains paradigmatic, we might have limited information about what kinds of within category variation are likely to matter. For example, we might expect that in different sex families a male breadwinning arrangement means a very unequal allocation of housework. Predominant theories such as relative resources, time availability and "doing gender" and their associated hypotheses do not point to the need to unpack this category. Researchers interested in identifying sub-groups of male breadwinning families that are more likely to equally share housework will most probably have to look for other variables that, when introduced as separate variables, are relatively large and significant.

We agree with Bowleg (2008) who argues that "the statistical methods, even those that test interactions, were not designed with the study of intersectionality in mind" (pg. 320) and that "we need new analytical tools and strategies to assist us in understanding the complexities of intersectionality" (pg. 320). Our aim in this paper is to introduce Classification and Regression Tree (CART) as a method for exploratory quantitative analysis of intra-categorical complexity to identify (potentially) small groups which might be of theoretical interest.

DATA AND METHODS

Using the first sweep of the UK Millennium Cohort Study (MCS), a large-scale nationally representative survey of parents with babies aged 9 months, we compare results from standard linear regression analysis and from a Classification Tree grown with the same variables. We consider how the results of each might be interpreted and what aspects of gendered divisions of labour in the UK they bring into focus.

Our analytic sample includes 13,120 different-sex co-resident families where both parents are the biological parents of the baby, at least one is the biological parent of any older child(ren), and where both parents participated in the full interview (86% of all co-resident parents). Although somewhat dated, this dataset is well-suited for addressing our aims. It provides detailed information about the division of housework and childcare and paid work. The inclusion of an ethnic minority boost sample also enables analysis to explore variation across ethnic groups. Importantly, the whole sample faced similarly intense everyday childcare and housework loads associated with having an infant in the household. The main respondent was asked for each of several tasks (cooking, cleaning, laundry and looking after household finances) whether it is mostly done by the mother, the father, shared more or less evenly, or done by someone else. We derived an index of the father's share of the total housework tasks done by the two parents by assigning each task the value 2 for mostly the father, 1 for shared and 0 for mostly the mother or someone else and taking sum of his share of the four housework tasks divided by the sum of both parents' shares (reverse-coding for the mother's share). A score of 0 indicates the mother mostly does all of the tasks, 0.5 indicates relatively even sharing

between the partners and 1 that the father mostly does all of the tasks. In most families, most of the housework tasks were mainly done by the mother; mean paternal share is 0.24 and in 21% of families all four of the tasks were mostly done by the mother.

In addition to variables measuring the mother's age (<25, 25-29, 30-34, 35-39, 40+); the marital status of the parents (married or cohabiting), the number of resident children (1, 2, 3 or 4+); and the number of adults in the household (parents only, extended family incl. grandparent(s), other extended), we constructed the following dependent variables:

Parental education: We focus on whether or not the parents have degree-level qualifications (including equivalent vocational qualifications), both as a crude proxy for social class, but also as an indicator of individual investment in human capital. For the analysis, we combined the mother's and father's information into a couple-level variable identifying whether neither (48%), the father (14%), the mother (15%), or both parents have a degree-level qualification (24%).

Potential earnings trajectory: To capture a measure of forward-looking bargaining power we classify potential earnings trajectory as flat or steep based on age and occupation. We use the within-occupation proportionate change in gross hourly earnings between one age group and the next (SOC2000 2-digit sub-major occupational groupings) for men reported in the 2002 Annual Survey of Hours and Earnings and applied this to the MCS data based on the respondent's age group and (most recent) occupation. Based on the distribution we set age-specific thresholds to designate a 'steep' trajectory as 20% earnings growth for those under 30 years; 9% for those aged 30-39 and 5% or more for those aged 40 and over, combining decreases and lower increases as a 'flat' trajectory. In 60% of families both parents had 'flat' trajectories while the father alone had a steep potential trajectory in 15%, the mother in 16% and both parents in 9%of the families. We acknowledge that this is a crude measure, however it provides some indication of respondents' reasonable expectations of future earnings, given their age and occupation. Unlike current earnings, this forward-looking indicator captures a fall-back position even in hypothetical labour market (re-)entry scenarios, especially relevant for mothers with young children (although it does not reflect any downward occupational mobility some mothers may already have experienced).

Parental economic activity: We combined information on the paid working hours of each partner and categorised families as having the following 'work arrangements': male breadwinner, both parents work 21-40 hours per week, both parents work long full-time hours (>40 hours per week), female breadwinner, and at most one part-time worker.

Parental ethnic group: The MCS records each parent's self-reported ethnic group separately using standard UK categories that for racialised minorities broadly reflect the post-colonial migrant groups. Because of the high correlation between partners' ethnic group, and in order to avoid attaching undue importance to either parent's ethnicity in the interpretation by including only one parent's ethnic group, we derived a joint parental ethnic group: white (89%), Indian (1.7%), Pakistani/Bangladeshi (3.4%), Black (1.2%), mixed or other ethnicity (1.2%; including East Asian), inter-ethnic couple (3.9%).

Parental age gap: The overall average age difference between the parents is just under three years (father older), and in 52% of the couples the partners were of a similar age (which we define as ranging from the mother one year older to father three years older). In 18% of families the father was 4-6 years older, in 16% 7+ years older and in 14% of families the mother was 2+ years older.

Parenthood history: For over a third (37%) of the families, the 9-month-old is the first child (children, in the case of twins/ triplets) for both parents, while a further 46% have older resident children together and neither parent has children from previous relationships. In a minority of families, the mother (6%), father (9%) or both parents (3%) have children from a previous relationship either co-resident or living elsewhere.

We estimate an OLS model that includes all of these controls and grow a classification tree using the same set of variables. Classification and Regression Trees (CART) use the covariates to divide the sample into mutually exclusive groups, aiming to maximise homogeneity on the outcome variable. The method, which is better-established in the epidemiological literature than in the demographic literature, uses a set of covariates to stratify the analytic sample, through a series of binary partitions, into exhaustive and mutually exclusive subgroups (a set of terminal nodes). Each sub-group shares a set of characteristics related to the outcome of interest (in this case, father's share of housework). The predicted outcome for each observation in the data is the average outcome for the terminal node to which that observation belongs. A fully saturated CART will often be too large and detailed, with numerous very small terminal-nodes (groups of individuals in the data), and thus CARTs are often "pruned" using cross-validation methods to select a less complex tree that is substantively useful and interpretable (Zhang and Singer 2010).

FINDINGS:

The results suggest that for large groups, the predicted share of housework is similar in OLS and CART estimates. However, the CART results make it more readily apparent that the range of father's share of housework is really quite low on average, and, with limited variance across the sample. CART results identify only three groups with a predicted share of fathers' housework exceeding 0.33 and these four groups account for only around three percent of the sample. These small but "real" groups, some of which are a combination of variables with positive and negative OLS parameter estimates, have average shares of housework that are not well predicted using OLS parameter estimates. Moreover, they can contribute to theory and hypothesis building. For example, one more egalitarian CART profile (father's share of housework =42%) appears to differentiate "do-over" or "do-it-again" dads – well-education and economically successful men who form second families at older ages, often with much younger women. Although these men have attracted much attention in the media, where they are depicted as highly committed and involved fathers (the second time around, at least), they remain a very rare group and have not received much scholarly attention (Sigle and Kravdal 2021).

Taken together, the results suggest that CART can provide a useful alternative or complementary method to standard linear regression because the results bring directly into focus (1) small groups that are hidden or misrepresented in additive-separable models with all else equal interpretations (2) with-in category heterogeneity that might not be predicted by predominant theories (3) actual combinations of characteristics in the data and (4) information on how common or typical certain high or low probability groups are.

References

Bauer, G.R., S.M. Churchill, M. Mahendran, C. Walwyn, D. Lizotte, and A.A. Villa-Rueda, 2021. Intersectionality in quantitative research: A systematic review of its emergence and applications of theory and methods. *SSM-Population Health*, *14*: 100798.

Bowleg, L., 2008. When Black+ Lesbian+ Woman≠ Black Lesbian Woman: The Methodological Challenges of Qualitative and Quantitative Intersectionality Research. *Sex Roles*, *59*: 312-325.

Burawoy, M. 2005. For public sociology, *American Sociological Review* 70(1): 4–28.

Collins, P. H. 2015. Intersectionality's definitional dilemmas. *Annual Review of Sociology*, *41*(1): 1-20.

Crenshaw, K. 1989. Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics, *University of Chicago Legal Forum* 1989(1): Article 8, 139–167.

Crenshaw, K. 1991. Mapping the margins: Intersectionality, identity politics, and violence against women of color, *Stanford Law Review* 43(6): 1241–1299.

Crimmins, E. M. 1993. Demography: The past 30 years, the present, and the future, *Demography* 30: 579–591.

Ferree, M. M. 2010. Filling the glass: Gender perspectives on families, *Journal of Marriage and Family* 72: 420–439.

Ferree, M. M., S. Khan, and S. A. Morimoto. 2007. Assessing the feminist revolution: The presence and absence of gender in theory and practice, in C. Calhoun (ed), *Sociology in America: The American Sociological Association Centennial History*. Chicago: University of Chicago Press, pp. 354–384.

Guan, A., M. Thomas, E. Vittinghoff, L. Bowleg, C. Mangurian and P. Wesson. 2021. An investigation of quantitative methods for assessing intersectionality in health research: A systematic review. *SSM-Population Health*, *16*: 100977.

McCall, L., 2005. The complexity of intersectionality. *Signs: Journal of Women in Culture and Society*, *30*(3): 1771-1800.

McKinzie, A.E. and P.L. Richards. 2019. An Argument for Context-Driven Intersectionality. *Sociology Compass*, *13*(4), p.e12671.

Sigle, W. 2016. Why demography needs (new) theories, in Dimitri Mortelmans, Koenraad Matthijs, Elisabeth Alofs, and Barbara Segaert (eds), *Changing Family Dynamics and Demographic Evolution*. Cheltenham: Edward Elgar Publishing, pp. 217–233.

Sigle, W. 2021. Demography's theory and approach: (How) has the view from the margins changed? *Population Studies*, *75*: 235–251.

Sigle, W. and Ø. Kravdal. 2021. With age comes … ? An examination of gendered differences in the resource advantages associated with parental age in Norway, *Comparative Population Studies* 46: 363–386.

Sigle-Rushton, W. 2014. Essentially quantified? Towards a more feminist modelling strategy, *The Sage Handbook of Feminist Theory*: 431–445.

Stacey, J. and B. Thorne. 1985. The missing feminist revolution in sociology, *Social Problems* 32(4): 301–316.

Zhang, H. and B. Singer. 2010. *Recursive Partitioning in the Health Sciences* Springer New York

Zarza, J. A. 2018. Representations of feminist theory and gender issues in introductory-level sociology textbooks, Dissertations and Theses. Paper 4358.