



# How to measure parenting styles?

Christopher Rauh <sup>1</sup> · Laëtitia Renée<sup>2</sup>

Received: 16 September 2021 / Accepted: 27 July 2022 / Published online: 22 August 2022  
© The Author(s) 2022

## Abstract

In this paper, we measure parenting styles through unsupervised machine learning in a panel following children from age 5 to 29 months. The topic model, which is a statistical model originally developed to discover the latent semantic structures in text, classifies parents into two parenting styles: “warm” and “cold”. Parents of the warm type tend to respond to children’s expressions in a supportive manner, while parents of the cold type are less likely to engage with their children in an encouraging manner. Warm parenting is more likely amongst educated and older mothers. Although styles reveal some persistence, the share of parents with a warm style decreases with the age of the child, in particular for boys. Children of warm parents achieve higher cognitive and non-cognitive scores at later ages. We find that the topic model estimated on different sample splits, such as by education or child age, reveal additional information while maintaining robust overall patterns.

**Keywords** Parenting styles · Human capital · Latent Dirichlet allocation · Inequality · Machine learning

## 1 Introduction

Early childhood investments have been shown to be crucial for children’s human capital development (Attanasio et al., 2020; Cunha et al., 2010; Del Boca et al., 2014), but the measurement of interactions between different dimensions of investment is challenging due to their complexity (Attanasio, 2015). Parental time investments generally are captured through different activities parents engage in with their children, such as visits to museums or the frequency of a parent reading to their children. The number of

---

✉ Christopher Rauh  
cr542@cam.ac.uk

✉ Laëtitia Renée  
laetitia.renee@umontreal.ca

<sup>1</sup> University of Cambridge, Cambridge, UK

<sup>2</sup> Université de Montréal, Montréal, QC, Canada

activities considered is either vast or restricted arbitrarily. When many investments are considered, they generally are combined (log-)linearly in latent factor models or using principal component analysis. More recently, the debate about parenting styles has emerged in economics emphasizing that not only investments, but also the style of investing matters (Doepke & Zilibotti, 2017, Fiorini & Keane, 2014).<sup>1</sup> In developmental psychology, Baumrind (1967) already classified parenting styles and related them to behavioral traits among pre-school children. This approach was extended by McCoby and Martin (1983) to four styles along two dimensions: warmth and control.

Parenting is characterized by a complex set of interactions and decisions. Draca and Schwarz (2021) discuss why linear combinations of features with the highest degrees of variance in the data, as is the case for principal component analysis or factor analysis, may not provide optimal summaries of complex data-generating processes. This paper proposes a new methodology to summarize parenting styles based on parental interactions with their children using unsupervised machine learning.<sup>2</sup> One advantage of this approach is that it allows aggregating any number of granular parental activities in a non-linear fashion. The algorithm learns from the co-occurrence of parental actions, and given that it is a mixed-member model, the same action can be assigned to different types.<sup>3</sup> If, on the one hand, a parent regularly checks on a child, and combines this with hugs and kisses, this could be considered warm control in the words of McCoby and Martin (1983). If, on the other hand, regularly checking on a child co-occurs with yelling at the child, the parenting style could be considered controlling combined with a lack of warmth.

When we restrict the algorithm to identify two styles, we find that parents' styles can be classified into "warm" and "cold" types.<sup>4</sup> The reason we limit the estimation to two styles is twofold. First, we only observe ten different parental actions in our data, which complicates the identification of more types. Second, two types simplify the exposition.<sup>5</sup> Warm parents are more likely to be supportive of their children's progress and speak directly to their children, while cold parents are characterized by hardly interacting with their children in the presence of the interviewer, and if they do, they tend to do so in a negative manner. In line with ad-hoc classifications in developmental psychology, the two parenting styles we discover can be interpreted as high warmth and high control versus low warmth and low control. Although parenting styles exhibit some persistence over time, we find that parents are more likely to adopt a warm parenting style when the children are younger. While there are

---

<sup>1</sup> In the earlier economics literature, parenting styles were defined as an index of punitive/aversive parenting using four questions about how the parent responds when the child misbehaves (Burton et al., 2002).

<sup>2</sup> With the availability of the necessary data, the approach could easily be applied to summarize parental investments rather than styles as well.

<sup>3</sup> This stands in contrast to k-means clustering where attachment to one cluster comes at the expense of the attachment to the other.

<sup>4</sup> The labels warm and cold are not necessarily meant to be interpreted as good versus bad. While we do believe (and find correlational evidence), that, on average, the warm parenting style is better, too much of a warm parenting style might have drawbacks (Haidt & Lukianoff, 2018) and the institutional environment might prime parents to adopt different styles (e.g. (Elliott & Reid, 2019)).

<sup>5</sup> In Section 4 we expand to three parenting styles and find that the algorithm can separate a 'high control' from the cold type. This emerging style is more active than the cold type but not as affectionate as the warm type.

no gender differences when the children are 5 months old, boys are considerably less likely to be exposed to the warm parenting style when they are 29 months old.

We contribute to two strands of literature. First, we contribute to the literature concerned about parenting styles.<sup>6</sup> They generally draw the distinction between parenting styles in terms of permissive, authoritarian, or authoritative. The empirical approaches tend to classify parenting styles based on a single binary response to a survey question, such as how important obedience is for a respondent (e.g., (Agostinelli et al., 2020)) or latent factor models (e.g., (Falk et al., 2021)). Our approach allows capturing parenting styles based on many questions with complex interactions. Moreover, an advantage of our data on parental activities is that they are not self-reported, but are observed and recorded by the enumerator, which should help to reduce systematic measurement error, and are the same set of actions observed across multiple survey waves. Our new interpretable measure summarizing the large dimensionality and complexity of parental activities is predictive of human capital above and beyond the predictive power of parental socio-economic characteristics, and we find that styles at the age of 5 months exhibit a higher correlation with cognitive and non-cognitive outcomes at age 6 than the styles at older ages. Despite the intuitive results we cannot claim causal effects due to the lack of an exogenous shock to parenting styles. This is a common feature of studies measuring the impact of parenting styles or human capital accumulation more generally.

Second, we add to the rapidly growing use of machine learning in Economics to classify behavioral types. The latent Dirichlet allocation (LDA) was originally developed by computer scientists Blei et al. (2003). The underlying idea is to classify text documents into a mixture of a small number of topics. One key is that the topics are not predefined but are backed out through co-occurrence. We apply the same idea of topics to behavioral types. Other approaches to classifying behavioral types using LDA are Bandiera et al. (2020) who classify CEOs using detailed time-use surveys and find that CEOs distinct behavior affects firm performance. Draca and Schwarz (2021) use LDA to measure political ideology. We contribute to this literature by using LDA to classify parenting styles and look at its relation to human capital accumulation in very early childhood.

## 2 Data

We use the Québec Longitudinal Study of Child Development (QLSCD), a detailed panel of a representative sample of families from Québec, a province in Canada, with a baby born between October 1997 and July 1998. More specifically, we focus our work on the 1,985 families who participated in the first three waves of the panel, conducted when the designated baby was 5, 17 and 29 months old.

---

<sup>6</sup> One can roughly separate this literature into two strands: First, the literature relating parenting styles to child development (e.g., (Agostinelli et al., 2020; Cobb-Clark et al., 2019; Cunha, 2015; Doepke & Zilibotti, 2019; Doepke et al., 2019)). Second, the literature studying the role of parenting styles in the intergenerational transmission of traits (e.g., (Brenøe & Epper, 2022; Falk et al., 2021; Zumbuehl et al., 2021)). Further, Del Boca et al. (2019) propose a model in which parental types are not merely the outcome of utility maximization by the parents but the result of a bargaining process with the children. Kiessling (2021) studies how parents perceive the returns to parenting styles in terms of warmth and control using hypothetical scenarios.

We rely on the Observations of Family Life (OFL) instrument filled by the enumerator at the end of the annual interview. It includes observations made during the interview about the behavior of the key respondent –the mother in 99% of the cases– and her interactions with her baby. This has the advantage of not relying on self-reported behavior which is common in the human capital literature and a potential source of bias.

We exclude mother-children pairs for whom the OFL instrument was not completed at child ages 5, 17 or 29 months because the child was sleeping. We end up with a sample of 1,443 mother-children pairs. Table 1 describes the socio-economic characteristics of the families.

We focus our analysis on the ten variables from the OFL instrument that assess the behavior of the interviewed mother toward her child. Table 2 displays descriptive statistics for these variables. We see that some parental actions are highly dependent on the age of the child. For instance, the share of parents regularly checking on their child decreases from 72% when the child is 5 months old to 32% when the child is 29 months old.

### 3 Discovering latent parenting styles

In the next step, the different features of parental behavior are summarized into interpretable behavioral styles using a machine learning algorithm based on the latent Dirichlet allocation. The topic model developed by Blei et al. (2003) is a clustering algorithm for discrete data, which traditionally was meant to reduce the high dimensionality of text into an arbitrary number of topics specified by the user. Each parental action can be featured with differing importance in each style, and each parent can be a mixture of styles. We note that for each action, we include both the appearance of an action as a count of one and the lack of appearance of an action as a count of one.<sup>7</sup>

The algorithm learns from the co-occurrence of counts through Bayesian updating. The idea is that if certain actions tend to appear together, they are likely to be linked to each other. The user only has to make three choices: The number of topics, i.e. parenting styles in our application, and the hyperparameters  $\alpha$  and  $\beta$  for the priors of the Dirichlet distributions which govern whether few actions should be dominant within each style, and whether few styles should be dominant amongst each parent. We set these hyperparameters to standard values of  $\alpha = 0.9$  and  $\beta = 0.9$ . In Appendix A we explain the technical details. For the sake of simplicity and interpretation, our main specification is based on two parenting styles.<sup>8</sup> The final output of the algorithm is the distribution of actions for each style and the style distributions for each parent.

---

<sup>7</sup> We need to do this as otherwise mothers who engage in no action would be eliminated from the estimation and not be assigned any style. The idea is that we interpret the absence of an action as a manifestation as well. Draca and Schwarz (2021) follow a similar approach and subsequently scale up the probabilities which we refrain from.

<sup>8</sup> There is no consensus about the choice of hyperparameters in topic models. In Section 4 we show that the styles are robust concerning our choice of Dirichlet priors, and we also present the results for three styles. Allowing for four styles does not give any additional insights as the algorithm cannot levy enough information in order to distinguish a fourth style, wherefore the types look very much like with three types, just that one of the types is duplicated. Therefore, the results are not shown.

**Table 1** Descriptive statistics

	<i>N</i>	Prop.
<b>Number of siblings</b>		
No sibling	656	45.5
One	560	38.8
Two or more	227	15.7
<b>Household type</b>		
Two-parent	1,179	81.7
Blended	159	11.0
Single-parent	102	7.1
Missing	3	0.2
<b>Mother's age</b>		
Less than 25	332	23.0
25-29	446	30.9
30-34	469	32.5
35 and more	195	13.5
Missing	1	0.1
<b>Mother born outside Canada</b>		
No	1,301	90.2
Yes	140	9.7
Missing	2	0.1
<b>Language spoken at home</b>		
French	1,176	81.5
Other	265	18.4
Missing	2	0.1
<b>Mother education</b>		
High school degree or less	380	26.3
Some college education	681	47.2
College degree	380	26.3
Missing	2	0.1
<b>Parental working status</b>		
Two-parents: both work	984	68.2
Two-parents: one works	304	21.1
Two-parents: none work	45	3.1
Single-parent: works	38	2.6
Single-parent: does not work	58	4.0
Missing	14	1.0
<b>Below poverty threshold</b>		
No	1,106	76.6
Yes	317	22.0
Missing	20	1.4
<i>N</i>	1,443	100

*Note:* The table shows descriptive statistics for families in our sample at the time of the first interview in 1998, when the designated child is 5 months old

**Table 2** Parental actions across waves

	Proportion of mothers who ...		
	Wave 1	Wave 2	Wave 3
	5 months	17 months	29 months
Regularly checks on her child	71.7	47.8	31.9
Speaks spontaneously to her child	40.6	43.2	46.3
Answers to her child	45.0	46.8	57.4
Kisses and hugs her child	42.6	17.3	13.9
Screams toward her child	<0.5	4.9	6.8
Is annoyed by her child	1.6	7.1	10.5
Reprimands her child	<0.5	4.3	5.5
Supports her child's progress	38.0	26.1	25.5
Organises play time	58.5	53.6	43.6
Gives pedagogical toys	68.2	59.0	43.7
Observations	1,443	1,443	1,443

*Note:* The table describes the behavior of the respondents and their interactions with their children during the annual QLSCD interview in terms of percentages. Behaviors are evaluated by the enumerator during the interview. Statistics are presented for the three first waves, when the designated child is 5, 17 and 29 months old

With this information at hand, we can then relate parenting styles to parental characteristics and human capital accumulation.

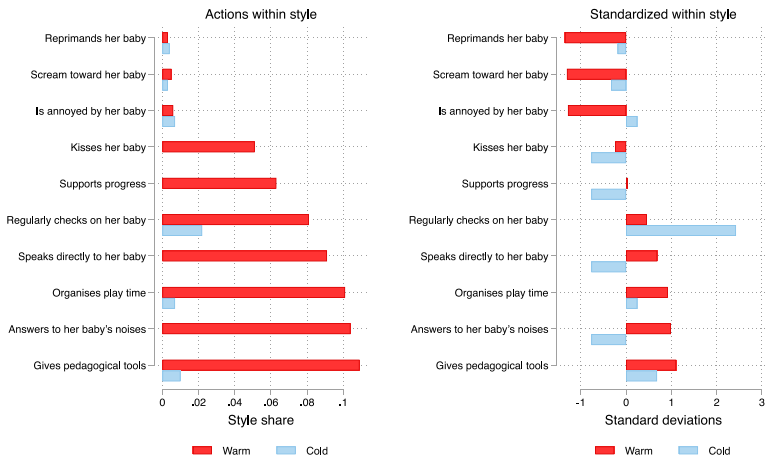
### 3.1 Parenting styles

In our main specification we aim to detect two styles using the three waves together. We estimate the classification for the entire sample with each parent providing three observations. In Section 4 we investigate the stability and consistency of the classification by estimating alternative models separately by wave and maternal education. We also expand to three parenting styles, and estimate a unique parenting style per family by summing across the three waves before estimating the topic model.

In Fig. 1 we display the probabilities of actions by the two types (left panel) and the standardized importance of an action within a style (right panel). The displayed probabilities in the left panel do not sum to one. The remaining probabilities are *not* taking the respective actions. The red bars represent the warm style and the blue the cold style.

Let us focus on the left panel for an illustration of the topic model. Picture a mother following a warm parenting style. While the mother is being interviewed, she closes her eyes and repeatedly draws an action from the warm action distribution (red bars).<sup>9</sup> The most likely action she will draw is giving the baby a pedagogical toy. The likelihood of drawing this action is proportional to the size of the red bar. Now picture a mother following a cold parenting style. She closes her eyes and draws from her action pool.

<sup>9</sup> Unfortunately, we do not observe how often parents engage in certain actions for each visit which would provide richer information for the model to learn from.



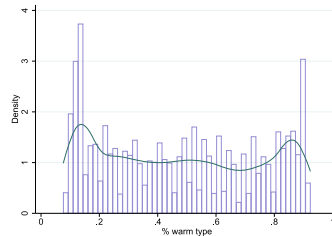
**Fig. 1** Distribution of actions by parenting style. The left panel describes the topic share of actions for each of the two parenting styles. The displayed probabilities in the left panel do not sum to one. The remaining probabilities are *not* taking the respective actions. The right panel shows the standardized importance of an action within a style by setting the mean to 0 and the standard deviation to 1

The most likely outcome is that she draws nothing since the blue bars are very small. Now if she actually does choose an action, it is most likely to be checking on the baby. In contrast to this example, parents generally do not follow pure styles, but a mix of the two. So a mother following a 50% warm style and 50% cold style will alternate between red and blue bars when she is deciding on her actions.

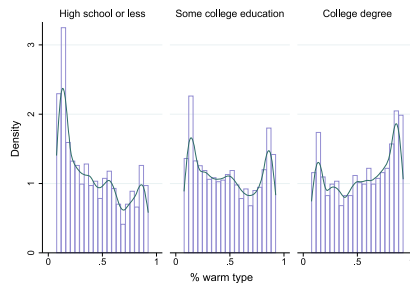
In the left panel we see that the main difference is the high amount of actions for the warm type, and the absence of actions for the cold type. The action that distinguishes the two types most in relative terms is supportive comments made by the parent to the child about its progress. While it is very common for the positive type to make supportive comments about the progress of the child, this is hardly the case for parents of the negative type, i.e. warm parents are 626 times more likely to do so. Similarly large differences exist for speaking to the child directly. The few actions the cold type engages in include reprimanding the child or expressing annoyance.

In the right panel we see the importance of a given action within a style, i.e. we subtract the overall mean of all actions from the probability of each action and divide by the standard deviation within that style. In this way it becomes very clear that warm mothers are much more likely to give pedagogical tools and react to noises than reprimand or scream at the baby. The cold parent, in contrast, checks on the child regularly if engaging with the child at all. The distribution of actions across types suggests that what distinguishes parents is the richness and warmth of action by one type versus the absence of interaction by the other, hence the labels warm and cold parents.

The LDA algorithm assigns to each parent a share of parenting style 1, the warm style (and with the remaining probability they follow style 2, the cold style). Figure 2 shows the distribution of the warm parenting style for the pooled sample in which each parent appears three times. While many parental styles are an intermediate mix of styles, we see a concentration of two masses: one with a low probability of engaging in a warm style (i.e. with a high probability of engaging in the cold style) and the opposite.



**Fig. 2** Distribution of parenting styles. The transparent bars represent the binned probabilities of the probability of engaging in a warm rather than a cold parenting style, while the solid line is the kernel density. The sample is the pooled sample in which each parent appears three times



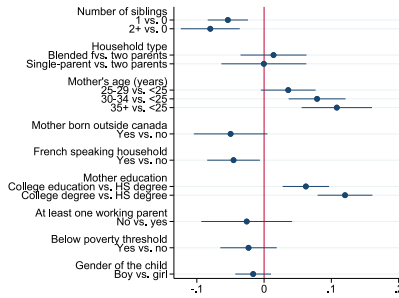
**Fig. 3** Distribution of styles by maternal education averaged across waves. The transparent bars represent the binned probabilities of the probability of engaging in a warm rather than a cold parenting style, while the solid line is the kernel density. The sample is the pooled sample in which each parent appears three times

### 3.2 Correlates and persistence of parenting styles

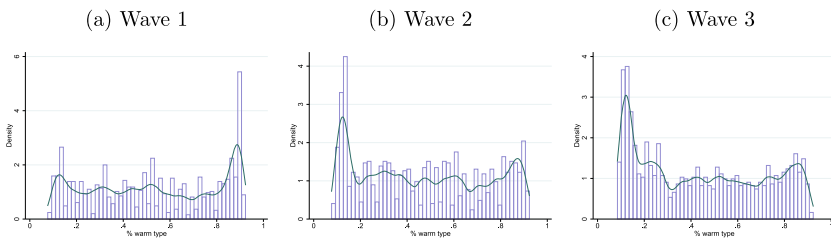
In Fig. 3 we show the distribution of warm types by maternal education. In the left panel, we see that mothers with high school or less tend to be of the cold type with an average share of warm style of 41.8%. In the middle panel, we see that for mothers with some college education the distribution appears closer to bi-modal with an average probability of warm style of 48.3%. Finally, in the right panel we see that amongst more educated mothers with a college degree, the average likelihood of being engaging in a warm parenting style increases to 53.9%. While the differences in these averages might not appear too large, the distributions are quite different, particularly in the share of mass at the lower and upper ends of the distribution.

The previous figure suggests that the likelihood of being a warm mother is increasing in education. In order to take a more systematic look at the relationship between type and individual characteristics, we regress the probability of being a warm type on age, education, poverty level, whether the parent is an immigrant, marital status, employment status, number of siblings, and the gender of the child. In Fig. 4 we see that parents with one additional child have a parenting style that is 5 percentage points colder, and with at least two more children it is 8 percentage points colder. The warmth of parenting styles appears to be increasing in maternal age; for instance, it is more than 11 percentage points greater for mothers older than 35 years compared to those younger than 25 when the first survey is conducted. Warmth also





**Fig. 4** Warm style probability and family characteristics. Coefficient plot with 95% confidence intervals from regressing warm parenting style on family characteristics using the pooled sample



**Fig. 5** Distribution of styles by wave. The transparent bars represent the binned probabilities of the probability of being a warm type rather than a cold type, while the solid line is the kernel density

increases with education and is 12 percentage points higher for those with a college degree than for those without any college education. We also find weaker evidence that warmth is lower for mothers born outside Canada and for French-speaking mothers. Notably, we find no difference between single and married mothers. While the fact that warm parenting tends to be less likely with more children in the household might suggest that exhaustion or constraints are an important driver, the fact that we find no difference between single and two-parent households casts doubts on this hypothesis.

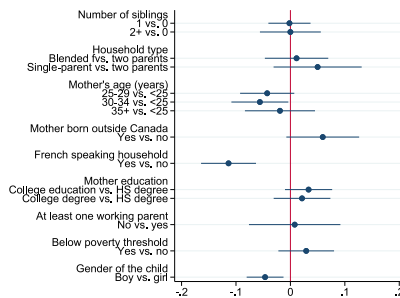
Next we look into the persistence of parenting styles across waves. In Fig. 5 we see that parenting styles flow from warm to cold over the course of the three survey waves. Estimating an autoregressive process in Table 3 reveals a persistence parameter of 0.31. When regressing parenting style on individual fixed effects one achieves an  $R^2$  of 0.52. This analysis suggests both significant persistence, but also substantial movement.

In order to investigate which characteristics can explain changes across waves, we regress warm parenting in wave 3 minus warm parenting in wave 1, i.e. the change in warm parenting, on family characteristics. We plot the coefficients in Fig. 6. The only clear correlations are a strong drop for French-speaking households by 11 percentage points, who already had a lower warmth at baseline, and for boys it declines by 5 percentage points. It is notable that, while there is no difference in the baseline level of parenting style by gender of the child, by the time of the third wave boys are less likely to receive a warm parenting style.

**Table 3** AR(1) estimation of warm style

	(1) Pooled	(2) Wave 2	(3) Wave 3
Warm style in $t - 1$	0.313*** (0.017)	0.261*** (0.025)	0.363*** (0.025)
Constant	0.302*** (0.010)	0.333*** (0.015)	0.275*** (0.013)
Observations	2,886	1,443	1,443
R-squared	0.100	0.070	0.131

*Note:* Estimation of warm parenting style as autoregressive process using OLS. The column headings indicate the estimation sample of the dependent variable. Robust standard errors clustered at the family level in parentheses. \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$

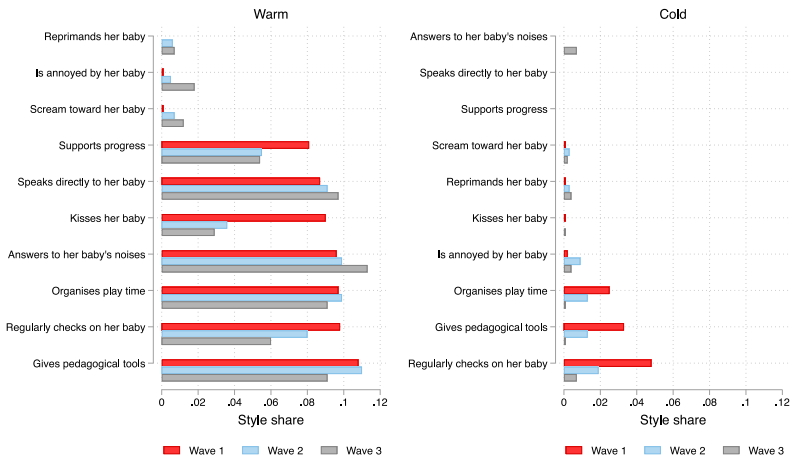


**Fig. 6** Increase in warm style probability from first to third wave and family characteristics. Coefficient plot with 95% confidence intervals from regressing warm parenting style in wave 3 minus warm parenting style in wave 1 on family characteristics using the classification of the disaggregated sample

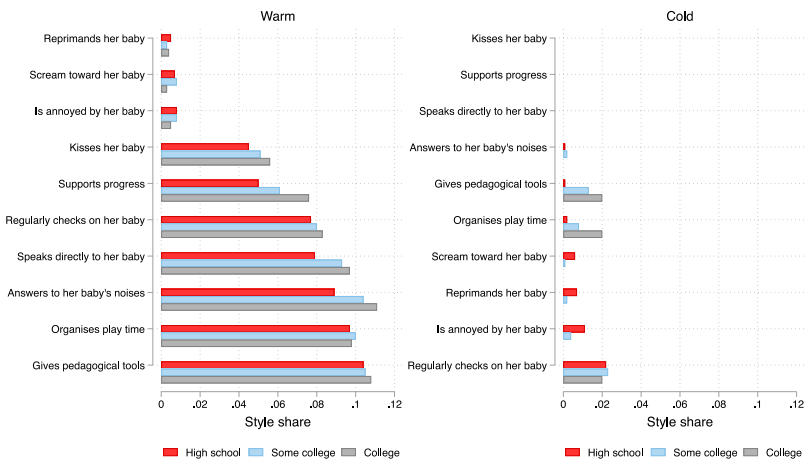
### 4 Alternative classifications of parenting styles

Some actions might be more pertinent to certain ages of the child, as is indicated by the shift in the distribution of actions in Table 2. Therefore, styles might change as well. In Fig. 7 we show the distribution of actions across styles when estimating the topic model on each wave separately. In general, we see similar patterns to the pooled estimation. On the left, we see that within the warm style supporting progress, kissing the baby, and regularly checking on the baby become less prevalent features of the warm style from wave to wave. On the right side, we see that the cold style features some control by regularly checking on the baby, giving toys, or organizing playtime in the first wave. However, by the third wave, the cold type stands out through extreme inaction. While estimating the classification for each survey wave separately provides the previously mentioned insights, it comes at the expense of comparability over time.

Next, we estimate topic models separately for each maternal level of education. In Fig. 8 we see in the left panel that the warm style is highly related across levels of education, with each action represented with a similar probability as in the benchmark estimation. However, in the right panel, we see some differences for the cold type. On the one hand, the cold style of the college-educated mother is characterized by some control and structure, featuring a minor level of some activities, for instance, regularly checking



**Fig. 7** Distribution of actions by style estimated separately by wave. The figure describes the topic share of actions for the warm (left) and cold (right) parenting styles estimated separately for each of the three waves



**Fig. 8** Distribution of actions by style estimated separately by maternal education. The figure describes the topic share of actions for the warm (left) and cold (right) parenting styles estimated separately by parental education

on the baby, organizing playtime, or giving pedagogical toys. On the other hand, the mothers with at most high school education mostly engage in negative actions, such as screaming, reprimanding, or expressing annoyance, if engaging in any activity at all.

As discussed in the Introduction, behavioral psychologists tend to define more than two parenting styles. We, therefore, use the same method as in the benchmark case but allow for three styles instead of two. In Appendix Figure 9 we show in the left panel the level of importance that each action forms within a style and in the right panel the standardized importance of actions within each of the three styles. While two of the styles are very similar to the previous estimation with two styles, a new style emerges which we label “high control” for which regularly checking on the baby is by far the most prevalent action amongst this parenting style.

We also test an alternative way of feeding the data into the topic model by summing the actions of each parent across the three waves, thereby limiting each parent to one single observation. In Appendix Figure 10 we see that, while the resulting warm style is very similar to the benchmark case, aggregating the data reveals a slightly more active cold style.

Finally, we test whether priors play an important role for the stability of the resulting parenting styles. In Appendix Figure 11 we see that changing a prior has close to no effect. These alternative estimations highlight some of the trade-offs involved when using a topic model to summarize parenting styles. Estimating the LDA on separate samples reveals systematic differences, allowing for comparisons of the prevalence of actions across and within styles. However, this can come at the expense of reducing the comparability of assigned parenting styles across the different samples, when looking at dynamics over time, or relating styles outcomes, as, for instance, warmth in wave 1 then might not mean the same thing as warmth in wave 2. In the next section, we will relate parenting styles to cognitive and non-cognitive outcomes of the children.

## 5 Relating parenting types to children's outcomes

In the following we investigate whether children exposed to different parenting styles have better cognitive and non-cognitive outcomes at age 6, i.e. almost 4 years after the final survey wave for which we measure parenting styles. We look into the relationships between styles and skills for alternative ways of estimating the LDA, in particular the pooled LDA, which is our benchmark, and the LDA estimated on the aggregated data. Moreover, we compare the correlations between skills and parenting styles using LDA to more commonly used methods such as principal component analysis and k-means clustering.

To test the relationship between parental type and the accumulation of children's cognitive skills, we use the results from six cognitive tests conducted at age 6, so 4 years after our final wave.<sup>10</sup> Here the sample size reduces to 811 children who took the tests. We extract the first factor across the six tests and standardize the score obtained with a mean of 0 and a standard deviation of 1.

In the first column of Table 4 we regress the cognitive score measure at age 6 on parental warmth in each of the three waves obtained from the pooled LDA, which is our benchmark summary of styles. We find that the parenting style at the youngest age, i.e. when the baby is 5 months old, exhibits the highest correlation with the outcomes measured later in childhood. Moving from the coldest to the warmest parenting style is associated with an increase of 0.34 standard deviations in cognitive test scores.<sup>11</sup> In column (2) we add a range of controls and find that this coefficient reduces to 0.23 but

<sup>10</sup> Specifically, we use the scores from the following six tests: test on numbers, ROST test, test on words (EVIP), memory test (VCR), intersection game (FIT), and block design test (WISC-III).

<sup>11</sup> While a direct comparison to an intervention is not the aim of this paper, we provide some information for the sake of interpretation of the magnitude: Duflo et al. (2011) find that tracking students in school by prior achievement increased test scores by between 0.14–0.19 standard deviations, and Attanasio et al. (2020) find that a randomized early childhood intervention targeting parental investments disadvantaged of children aged 12 to 24 months led to an increase in cognitive skills of 0.25 standard deviations. These effects were measured about 1.5 years after the interventions ended and are considered to be large in magnitude.

**Table 4** Regression of cognitive skills on parenting styles

	Pooled LDA			Aggregated LDA			PCA	KMEANS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Warm <i>t</i> = 1	0.344** (0.135)	0.227* (0.133)						
Warm <i>t</i> = 2	0.197 (0.149)	0.002 (0.146)						
Warm <i>t</i> = 3	0.168 (0.135)	0.070 (0.132)						
Warm (mean)			0.705*** (0.183)	0.292 (0.178)				
Warm					0.539*** (0.134)	0.244* (0.129)		
Score comp. 1							0.029* (0.016)	
Cluster ID								0.147** (0.067)
Constant	-0.354*** (0.097)	-0.368*** (0.138)	-0.344*** (0.097)	-0.363*** (0.137)	-0.270*** (0.076)	-0.340*** (0.125)	-0.222** (0.106)	-0.446*** (0.149)
Observations	811	811	811	811	811	811	811	811
R-squared	0.019	0.157	0.018	0.156	0.020	0.157	0.156	0.158
SES controls	NO	YES	NO	YES	NO	YES	YES	YES
Joint signif.	0.00	0.26						

*Note:* Measures of parenting styles are summarized as specified in the column headings. 'Pooled LDA' is the benchmark method, 'aggregated LDA' is by summing parental actions over the three waves before running the LDA, 'PCA' is principal component analysis, and 'KMEANS' is k-means clustering. The dependent variable is computed by taking the first factor of six measures of cognitive ability at age 6 years old. The score is standardized with a mean zero and a standard deviation of 1. SES controls include family composition (number of siblings, household type), maternal characteristics (age, whether born outside Canada, educational attainment), parental working status, language spoken at home, and whether family is below poverty threshold. They are described in more details in Table 1. The joint significance is the p-value from an F-test of joint significance for the three style variables. Robust standard errors in parentheses. \*\*\**p* < 0.01, \*\**p* < 0.05, \**p* < 0.1

remains significant at the 10% level. While we have no exogenous variation and therefore cannot claim causality, this correlation provides suggestive evidence to the debate about when critical periods of human capital development might be.

In column (3) we show the results when taking the average of the three warmth levels across the three waves and find a large and significant coefficient of 0.71. However, once we add controls in column (4) the coefficient reduces to 0.29 and is slightly too noisy to be significant at conventional levels. In column (5) we use the parenting style estimated using the aggregated data and again find a large significant coefficient of 0.54, which in column (6) when adding controls reduces to 0.24, but is precise enough to remain significant at the 10% level.

As a comparison in column (7), we include the first score from a principal component analysis and in column (8) we use a dummy derived using k-means clustering. According to the  $R^2$ , the strength of the association between parenting style and cognitive skills is similar across the three models.<sup>12</sup> However, one benefit of the LDA classification lies in its interpretation. Column (6) indicates that children from pure warm type parents have 0.24 standard deviation higher cognitive skills than children from pure cold type parents after including a range of controls. The principal component method does not allow us to deliver such an intuitive interpretation as it is less clear what moving from a low to a high score entails. The interpretation of the coefficient from the k-means clustering method in column (8) is more straightforward. It indicates that children of parents with a warm parenting style have 0.15 standard deviation higher cognitive skills than those with parents with a cold style. The effect is smaller with the cluster algorithm than with the LDA algorithm since it classifies parents between two styles, but does not allow parents to be of a mixed type as the LDA does.

In Table 5 we repeat the same exercise for non-cognitive scores, with a greater score representing fewer behavioral problems.<sup>13</sup> We find relatively similar albeit weaker associations than for cognitive skills. Again warmth at 5 months exhibits the strongest correlation without controls in column (1), but not anymore with controls in column (2). Mean warmth from the pooled LDA displays a coefficient of 0.33 with controls in column (4), and warmth from the aggregated data 0.25 with controls in column (6), both significant at the 10% level.

## 6 Conclusion

Human capital accumulation is one of the most important fundamentals of productivity and innovation. However, summarizing inputs into the human capital production functions is riddled with complications, including the potentially high dimensionality and non-linear relationships between parental actions. In this paper, we provide a new

---

<sup>12</sup> When excluding controls, the  $R^2$  for explanatory power of the parenting style measured using the aggregated LDA for cognitive skills is 5% higher than for the principal component method and 11% higher than for k-means clustering.

<sup>13</sup> We use the standardized value of the first factor of six measures of non-cognitive skills, namely not suffering from hyperactivity, inattention, anxiety, emotional disorders, aggressiveness, and social withdrawal (as reported by the mother).

**Table 5** Regression of non-cognitive skills on parenting styles

	Pooled LDA			Aggregated LDA			PCA	KMEANS
	(1)	(2)	(3)	(4)	(5)	(6)		
Warm <i>t</i> = 1	0.224* (0.132)	0.119 (0.132)						
Warm <i>t</i> = 2	0.043 (0.147)	0.032 (0.145)						
Warm <i>t</i> = 3	0.195 (0.142)	0.177 (0.141)						
Warm (mean)			0.456** (0.181)	0.328* (0.188)				
Warm					0.341*** (0.131)	0.250* (0.136)		
Score comp. 1							0.027 (0.017)	
Cluster ID								0.138* (0.071)
Constant	-0.228** (0.097)	-0.424*** (0.141)	-0.223** (0.096)	-0.425*** (0.141)	-0.171** (0.076)	-0.389*** (0.128)	-0.269** (0.114)	-0.480*** (0.152)
Observations	811	811	811	811	811	811	811	811
R-squared	0.008	0.077	0.008	0.076	0.008	0.077	0.076	0.077
SES Controls	NO	YES	NO	YES	NO	YES	YES	YES
Joint significance	0.07	0.34						

*Note:* Measures of parenting styles are summarized as specified in the column headings. 'Pooled LDA' is the benchmark method, 'aggregated LDA' is by summing parental actions over the three waves before running the LDA, 'PCA' is principal component analysis, and 'KMEANS' is k-means clustering. The dependent variable is computed by extracting the first factor of six measures of non-cognitive ability at age 6 years old. The score is standardized with a mean zero and a standard deviation of 1. SES controls include family composition (number of siblings, household type), maternal characteristics (age, whether born outside Canada, educational attainment), parental working status, language spoken at home, and whether family is below poverty threshold. They are described in more details in Table 1. The joint significance is the p-value from an F-test. Robust standard errors clustered at the family level in parentheses. \*\*\**p* < 0.01, \*\**p* < 0.05, \**p* < 0.1

way to summarize parental styles adopted from computational linguistics. We use an unsupervised machine learning model, the latent Dirichlet allocation, to classify parenting styles into two types. The resulting types can be interpreted as warm parents who encourage their children and express their affection, versus cold parents who do not interact much with their children and are more likely to punish them when they do so.

We show that these two styles relate systematically to parental characteristics, i.e. mothers with higher education and older mothers tend to be more likely to engage in warm parenting. Over time warm parenting declines, in particular when the child is a boy. Moreover, we show that the warmer a parenting style, the higher levels of cognitive and non-cognitive skills achieved years later. While we cannot establish a causal relationship between parenting styles and outcomes due to the nature of the data, we are optimistic that future studies including natural experiments or randomized control trials can make use of the proposed methodology to classify parenting styles based on rich information of their actions without having to artificially reduce the number of actions or investments based on ex-ante priors or ex-post information. The proposed approach can be done with an extremely large set of actions or even detailed time use data in a relatively agnostic way, potentially allowing the identification of more styles.

**Acknowledgements** CR would like to thank the FRQSC for financial support (grant number 2020-NP-267422).

**Conflict of interest** The authors declare no competing interests.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

## 7 Appendix

### 7.1 Latent Dirichlet allocation

Adapting the technical terms from Blei et al. (2003) for text and applying them to our objective, the corpus of behavioral actions  $D$  is composed of parents  $w$  of actions. A behavioral type is a probability distribution over all actions. The assumed underlying process with which types generate actions is by drawing  $\theta$  from a Dirichlet distribution with hyperparameter  $\alpha$ . Then for each action  $n$  of all actions  $N$ , one chooses a type from  $z_n$ . After that an action  $w_n$  is chosen for the corresponding type  $z_n$  from a Dirichlet distribution with hyperparameter  $\beta$ .



Written formally, the generative process of actions is expressed as the following joint distribution

$$p(\beta, \theta, z, w_d) = \prod_{i=1}^k p(\beta_i) \prod_{d=1}^D p(\theta_d) \left( \prod_{n=1}^N p(z_{d,n} | \theta_d) p(w_{d,n} | \beta, z_{d,n}) \right).$$

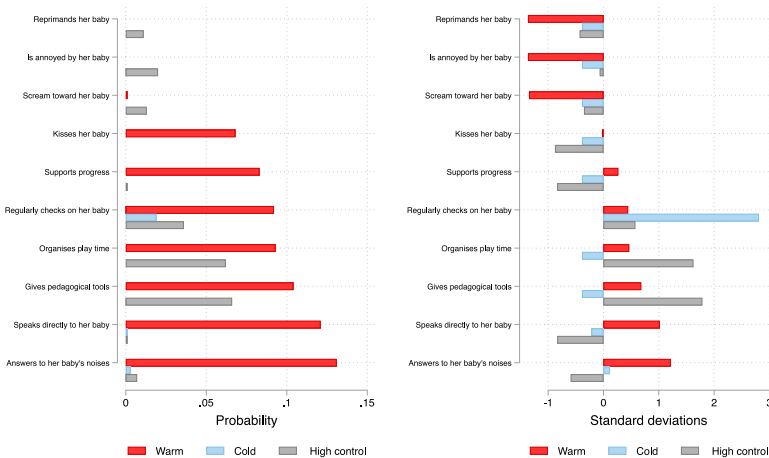
Given the corpus of actions, the task of the algorithm is to infer the type-specific action distribution and the parent specific type distribution. So the posterior distribution of the latent variables is given by

$$p(\beta, \theta, z | w_d) = \frac{p(\beta, \theta, z, w_d)}{p(w_d)}.$$

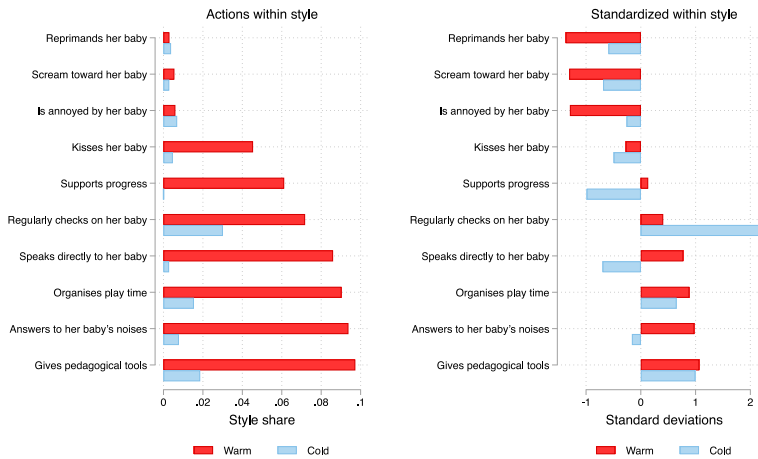
In order to infer the marginal distribution  $p(w_d)$ , which can be done through approximation using Gibbs sampling, or Variational Kalman Filtering and Variational Wavelet Regression, we rely on the Stata implementation developed by Schwarz (2018). Schwarz (2018) use the inference algorithm developed by Hoffman et al. (2010) and implemented by Pedregosa et al. (2011). As is the case in Draca and Schwarz (2021), the assumption of the independence of responses does not strictly hold in our approach. If an action has been recorded, the same action is not recorded again for the same person. They discuss in detail why the inference of LDA is nonetheless still valid.

### 7.2 Additional figure

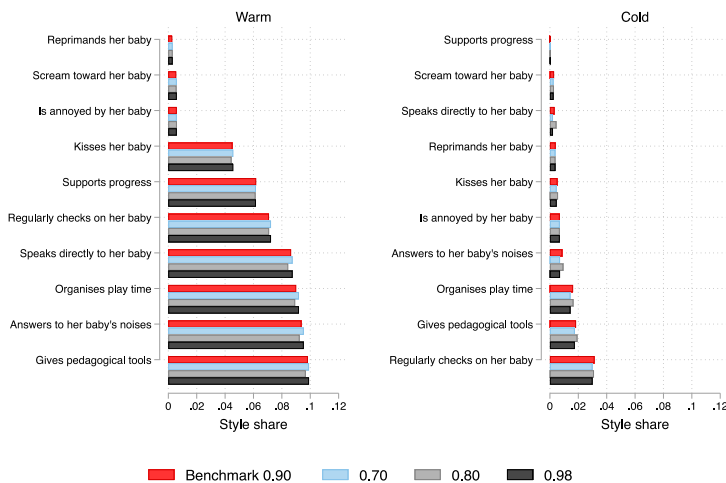
Figures 9, 10 and 11



**Fig. 9** Distribution of actions with three styles. The left panel describes the topic share of actions for each of the two parenting styles. The right panel shows the standardized importance of an action within a style by setting the mean to zero and the standard deviation to 1



**Fig. 10** Distribution of actions by parenting style when aggregating data. The left panel describes the topic share of actions for each of the two parenting styles. The right panel shows the standardized importance of an action within a style by setting the mean to zero and the standard deviation to 1. While in the benchmark each parent/survey wave provides one observation, here each parent provides one observation by summing all their actions



**Fig. 11** Distribution of actions with different priors. The figure describes the topic share of actions for the warm (left) and cold (right) parenting styles estimated using different priors. The legend specifies the hyperparameter  $\alpha$ , i.e. the parent specific style prior

**References**

Agostinelli, F., Doepke, M., Sorrenti, G., & Zilibotti, F. (2020). It takes a village: the economics of parenting with neighborhood and peer effects. Working Paper 27050. National Bureau of Economic Research.

Attanasio, O., Meghir, C., & Nix, E. (2020). Human capital development and parental investment in India. *Review of Economic Studies*, 87(6), 2511—2541.

- Attanasio, O. P. (2015). The determinants of human capital formation during the early years of life: Theory, measurement, and policies. *Journal of the European Economic Association*, 13(6), 949–997.
- Attanasio, O., Cattan, S., Fitzsimons, E., Meghir, C., & Rubio-Codina, M. (2020). Estimating the production function for human capital: results from a randomized controlled trial in Colombia. *American Economic Review*, 110(1), 48–85.
- Bandiera, O., Stephen, H., Andrea, P., & Raffaella, S. (2020). CEO behavior and firm performance. *Journal of Political Economy*, 128(4), 1325–1369.
- Baumrind, D. (1967). Child care practices anteceding three patterns of preschool behavior. *Genetic Psychology Monographs*, 75(1), 43–88.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3(Jan), 993–1022.
- Brenøe, A., & Epper, T. (2022). Parenting values moderate the intergenerational transmission of time preferences. *European Economic Review*, 148, 104208.
- Burton, P., Phipps, S., & Curtis, L. (2002). All in the family: A simultaneous model of parenting style and child conduct. *American Economic Review*, 92(2), 368–372.
- Cobb-Clark, D. A., Salamanca, N., & Zhu, A. (2019). Parenting style as an investment in human development. *Journal of Population Economics*, 32(4), 1315–1352.
- Cunha, F. (2015). Subjective rationality, parenting styles, and investments in children. In *Families in an Era of Increasing Inequality*. 83–94. Springer.
- Cunha, Flávio, Heckman, J. J., & Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3), 883–931.
- Del Boca, D., Flinn, C., & Wiswall, M. (2014). Household choice and child development. *The Review of Economic Studies*, 81(1), 137–185.
- Del Boca, D., Flinn, C. J., Verriest, E., & Wiswall, M. J. (2019). Actors in the Child Development Process. Working Paper 25596. National Bureau of Economic Research.
- Doepke, M., & Zilibotti, F. (2017). Parenting with style: Altruism and paternalism in intergenerational preference transmission. *Econometrica*, 85(5), 1331–1371.
- Doepke, M., & Zilibotti, F. (2019). *Love, money, and parenting: How economics explains the way we raise our kids*. Princeton University Press.
- Doepke, M., Sorrenti, G., & Zilibotti, F. (2019). The economics of parenting. *Annual Review of Economics*, 11.
- Draca, M., & Schwarz, C. (2021). How polarized are citizens? Measuring ideology from the ground-up. The Warwick Economics Research Paper Series (TWERPS) 1218.
- Duflo, E., Dupas, P., & Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review*, 101(5), 1739–74.
- Elliott, S., & Reid, M. (2019). Low-income Black mothers parenting adolescents in the mass incarceration era: The long reach of criminalization. *American Sociological Review*, 84(2), 197–219.
- Falk, A., Kosse, F., Pinger, P., Schildberg-Hörisch, H., & Deckers, T. (2021). Socio-economic status and inequalities in children's IQ and economic preferences. *Journal of Political Economy*, 129(9).
- Fiorini, M., & Keane, M. P. (2014). How the allocation of children's time affects cognitive and non-cognitive development. *Journal of Labor Economics*, 32(4), 787–836.
- Haidt, J., & Lukianoff, G. (2018). *The coddling of the American mind: How good intentions and bad ideas are setting up a generation for failure*. Penguin UK.
- Hoffman, M., Bach, F. R., & Blei, D. M. (2010). Online learning for latent dirichlet allocation. 856–864.
- Kiessling, L. (2021). How do parents perceive the returns to parenting styles and neighborhoods? *European Economic Review*, 139, 103906.
- McCoby, E., & Martin, J. (1983). Socialization in the context of the family: Parent-child interaction. *Handbook of child psychology*, 4, 1–101.
- Pedregosa, F., Varoquaux, Gaël, Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., & Dubourg, V., et al. (2011). Scikit-learn: Machine learning in Python. *The Journal of machine Learning research*, 12, 2825–2830.
- Schwarz, C. (2018). `ldagibbs`: A command for topic modeling in Stata using latent Dirichlet allocation. *The Stata Journal*, 18(1), 101–117.
- Zumbuehl, M., Dohmen, T., & Pfann, G. A. (2021). Parental involvement and the intergenerational transmission of economic preferences, attitude and personality traits. *Economic Journal*, 131(638), 2642–2670.