Routledge
Taylor & Francis Group

Educational Review

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/cedr20

Factors affecting teacher job satisfaction: a causal inference machine learning approach using data from TALIS 2018

Nathan McJames, Andrew Parnell & Ann O'Shea

To cite this article: Nathan McJames, Andrew Parnell & Ann O'Shea (12 May 2023): Factors affecting teacher job satisfaction: a causal inference machine learning approach using data from TALIS 2018, Educational Review, DOI: 10.1080/00131911.2023.2200594

To link to this article: https://doi.org/10.1080/00131911.2023.2200594

Published online: 12 May 2023.

EDUCATIONAL REVIEW
https://doi.org/10.1080/00131911.2023.2200594

Factors affecting teacher job satisfaction: a causal inference machine learning approach using data from TALIS 2018

Nathan McJames [a,b], Andrew Parnell[a,b] and Ann O'Shea[b]

[a]Hamilton Institute, Maynooth University, Maynooth, Ireland; [b]Department of Mathematics and Statistics, Maynooth University, Maynooth, Ireland

ABSTRACT
Teacher shortages and attrition are problems of international concern. One of the most frequent reasons for teachers leaving the profession is a lack of job satisfaction. Accordingly, in this study we have adopted a causal inference machine learning approach to identify practical interventions for improving overall levels of job satisfaction. We apply our methodology to the English subset of the data from TALIS 2018. Of the treatments we investigate, participation in continual professional development and induction activities are found to have the most positive effect. The negative impact of part-time contracts is also demonstrated.

ARTICLE HISTORY
Received 8 April 2022
Accepted 3 April 2023

KEYWORDS
Teacher job satisfaction; teacher retention; causal inference; machine learning; TALIS

Introduction

Background

Teacher supply and demand is an important challenge faced by many countries around the world (UNESCO, 2015). The scale of this problem is partly reflected in the teacher shortages currently facing many countries including England (Hilton, 2017), Ireland (O'Doherty & Harford, 2018), the United States (Wiggan et al., 2021), and many others. The scale of the challenge currently facing England is made clear by a recent House of Commons report which reveals that the 2019 five-year retention rate was at its lowest level since 1997, with 32.6% of teachers entering the profession in 2014 no longer teaching in classrooms five years later (Long & Danechi, 2021). These sustained high levels of attrition have led to a situation where the total number of all qualified teachers in England working outside of the state funded sector in 2019 (350,000) was nearly as high as the number of teachers working inside it (454,000). This comes at a time when secondary school pupil numbers are expected to rise by 7% in England between 2020 and 2026, thus placing increasing pressure on already difficult recruitment and retention targets. Teacher shortages are often more pronounced in Science, Technology,

CONTACT Nathan McJames nathan.mcjames.2016@mumail.ie Hamilton Institute, Maynooth University, Maynooth, Co. Kildare, Ireland Department of Mathematics and Statistics, Maynooth University, Maynooth, Co. Kildare, Ireland

Engineering and Mathematics (STEM) subjects (Han & Hur, 2021). This is also noted in the House of Commons report, which showed the subject-specific vacancy rate of unfilled teaching posts for these subjects was higher than the average (Science 1.4%, Technology 1.7% and Maths 1.4%) (Long & Danechi, 2021).

Teacher shortages may arise as a result of insufficient numbers of new entrants to the profession or high levels of qualified teachers leaving their posts. In addition to the serious challenges associated with not having enough teachers, higher levels of teacher turnover have been shown to negatively affect student learning and also incur large economic costs (Levy et al., 2012; Sorensen & Ladd, 2020). Encouraging teachers to stay in their posts is therefore very important. Research has shown that job satisfaction is one of the key predictors of a teacher's intention to remain in the profession (Klassen & Chiu, 2010; Madigan & Kim, 2021; Wang et al., 2015). Consequently, it is vital to identify factors that can improve job satisfaction in order to boost retention rates of qualified teachers and to attract new entrants to start their career. In addition to the economic and staffing implications of job satisfaction which are the primary focus of this study, Toropova et al. (2021) point out that happier teachers tend to have happier students, and more satisfied teachers provide higher quality teaching to their students as well (Klusmann et al., 2008; Spilt et al., 2011). Job satisfaction has also been shown to predict teacher self-efficacy which is another significant area of study within the literature (Burić & Kim, 2021). These reasons combine to make teacher job satisfaction a crucially important outcome of interest.

We decided therefore to investigate the effect that a number of selected factors may have on teacher job satisfaction. Job satisfaction is a term for which no single definition exists, but a widely accepted version given by Locke (1969) describes job satisfaction as "the pleasurable emotional state resulting from the appraisal of one's job as achieving or facilitating the achievement of one's job values". Informally, job satisfaction can be thought of as an overall sense of contentment with one's career.

Our approach in this study can be best described as an application of causal inference machine learning algorithms. We employ these cutting-edge statistical models in order to identify specific, implementable steps that may be taken to enhance the job satisfaction of qualified teachers. This is a key advantage of our approach, because it allows school principals and other policy makers to determine specific steps that may be taken as part of a school strategy for improving job satisfaction. Our method makes use of Bayesian Additive Regression Trees (BART) (Chipman et al., 2010), a cutting-edge modelling tool which enables us to detect non-linear relationships and interactions which would not normally be found in a standard linear model. Additionally, this strategy allows us to control for a much larger number of background (confounding) variables than would normally be possible when using a linear model. Furthermore, we demonstrate how this approach can be used to identify subgroups of teachers who are most (or least) likely to benefit from the positive effects of a given treatment.

Our study uses data from the third cycle of the Teaching and Learning International Survey (TALIS) which took place in 2018 (OECD, 2019a). TALIS is the world's largest survey of teachers and principals and has taken place every five years since 2008. A fourth cycle is due to take place in 2024. Participating teachers and principals are asked to complete questionnaires on a wide variety of topics such as personal background; current teaching duties; their perception of the school climate and job satisfaction. TALIS

2018 is the largest of the surveys to-date with 48 countries participating, and includes data on approximately 260,000 teachers from 15,000 primary, and lower and upper secondary schools. For the purpose of this study, however, we will limit our investigation to the data from England. This subset of the entire dataset contains a representative sample of 2009 primary and 2376 lower secondary school teachers for a total sample size of 4385.

We decided to focus on the English subset of the TALIS data for a number of reasons. Firstly, the English subset of the data is able to provide us with a relatively large sample size of teachers from both primary and lower secondary schools. This is advantageous for machine learning models, as it enables BART to more easily detect relationships between variables in the data, and this is essential for reliably producing accurate results. A second important factor we considered is that England is currently facing a serious teacher recruitment and retention problem (Hilton, 2017). This important contextual factor makes England a more appropriate choice than a country not currently facing such difficulties. Furthermore, a number of initiatives such as the Early Career Framework (Daly et al., 2021; Department for Education, 2019) have recently been introduced in England, thus making our investigation of mentorship and induction schemes particularly relevant to the English context.

With these data, and using a causal inference machine learning approach, we attempt to answer the following research question: What are the specific and implementable factors that have the most positive (or negative) impact on teacher job satisfaction? The factors we consider include participation in induction schemes; high levels of participation in continual professional development; team teaching; observing other teachers; mentorship schemes; teaching in a public school; class size; out-of-field teaching and having a part-time contract. Our decision to include these factors in our investigation has been informed by previous studies which show they have a strong association with both teacher job satisfaction and retention. We now discuss these findings in more detail in a literature review, focusing on key aspects relevant to our research.

Literature review

Induction and mentoring programmes

Induction is a broad term used to describe different activities or supports put in place for teachers to assist them in adapting to the ethos or practices of a new school (Allen, 2005). Induction programmes are frequently designed with newly-qualified teachers in mind, but we will use the slightly more inclusive definition from the TALIS questionnaire, which broadens the scope of induction activities to include supports for experienced teachers who have recently begun teaching in a different school (OECD, 2018).

Mentoring describes the arrangement whereby a newly-qualified teacher is assigned a more experienced member of staff at their school, who will advise and assist them as they begin their career (Allen, 2005). Roles of a mentor can vary, as can frequencies of meetings between a mentor and their mentee. For consistency, we will once again use the more general definition provided in the TALIS questionnaire which allows mentoring to encompass any situation where a more experienced teacher supports a less experienced one, who need not be newly qualified (OECD, 2018).

New teachers are commonly faced with many challenges in the classroom after they qualify (Guarino et al., 2006). To mitigate the risk of newly qualified teachers encountering difficulties, induction and mentoring schemes are often provided to support them during this formative stage of their career. In fact, induction for new teachers is statutory in many countries, such as in England where this is the case in state schools, and a new scheme for early career teachers has recently been introduced called the Early Career Framework (Daly et al., 2021; Department for Education, 2019).

International evidence often points towards induction and mentoring schemes as having a positive effect on the job satisfaction of participating teachers. Regression analyses of teachers in the US subset of TALIS 2018, for example, have found a strong link between the presence of a mentor and considerably higher levels of job satisfaction (Renbarger & Davis, 2019). This finding is backed by a review of ten studies on the effect of mentoring, which reveals consistent evidence in support of the positive effects of mentoring on teacher retention (Ingersoll & Kralik, 2004). Other studies based on the US subset of TALIS 2018 have also identified induction activities as having a positive effect on job satisfaction (Reeves et al., 2022). Additionally, the provision of induction supports for newly qualified teachers in their first year of teaching has also been linked to lower levels of attrition (Ronfeldt & McQueen, 2017).

Despite their widespread use, the evidence supporting the use of induction and mentoring schemes for all teachers has sometimes been brought into question. A large-scale review of over 90 studies by Allen (2005) found only limited evidence that participation in induction and mentoring schemes leads to higher retention rates of qualified teachers. Indeed, survey studies of teachers undergoing statutory induction in the UK suggest that initial teacher education may be far more important for preparing new teachers for the challenges they will face in their first year of teaching (Hulme & Wood, 2022).

It is also true that while induction schemes or mentoring may be beneficial for job satisfaction and retention in the long term, not all teachers report enjoying induction or mentoring at the time. Some teachers undergoing induction in the UK report it as being a stressful experience due to the busyness of their schedule, and others report dreading meetings with mentors who provide them with criticism (Smethem, 2007).

The effect of being a mentor on job satisfaction has been the focus of relatively little research in comparison to the effect of having a mentor. Despite this, there are still studies which show that mentoring arrangements can be mutually beneficial to both the mentor and the mentee. Lunsford et al. (2018) for example, in a study of US teachers, found that those with either a mentor or a mentee are on average more satisfied than teachers who do not.

It is important to note that most of the above findings are based on observational data, and therefore the positive correlation between induction or mentoring and job satisfaction cannot be claimed to be causal in nature. A recent study with a longitudinal design which tracked a sample of newly-qualified teachers in the US over the first 5 years of their career therefore makes an important contribution (Gray & Taie, 2015). At each follow-up visit, teachers who were assigned a mentor during their first year in the classroom were more likely to still be teaching than those who did not receive this extra support, thus showing a temporal association between mentoring and retention.

Continual professional development

Continual professional development (CPD) can refer to a wide range of activities designed to assist teachers as they build upon and improve their professional skills (OECD, 2018).

Higher levels of participation in CPD have often been linked to improved teacher job satisfaction (Wang et al., 2020; Yoon & Kim, 2022). In a joint study of English and international data from TALIS 2013, Sims (2017) was able to show that this relationship holds in both the national and international context. With two separate analyses, they demonstrated that there is a non-causal positive correlation between CPD and job satisfaction, firstly using data for England only, then again for a combined dataset of more than 50,000 teachers from 38 different countries.

CPD has also been shown to be related to higher levels of teacher retention. A survey of 500 teachers based in England, for example, who had just completed a professional development course showed that teachers who were more engaged with the CPD course were more likely to respond that the course had a positive effect on their intention to remain teaching (Coldwell, 2017). This link was less strong for teachers who only engaged moderately or weakly with the course. Furthermore, Allen and Sims (2017), in a study of teachers at state-funded secondary schools in England, found that similar effects were still being felt two years after participation in a science subject-specific CPD course, and that participation had reduced department turnover rates by two percentage points. This finding is especially important given that STEM subject teachers are known to be at higher risk of attrition (Han & Hur, 2021).

Despite these benefits, one challenge often faced by teachers is that there may be barriers to their attendance at different CPD activities due to factors such as timetabling issues, cost of travelling to CPD events or a lack of suitable events being organised (Zhang et al., 2020). It is unsurprising then, that the presence of barriers to attending quality CPD activities has also been linked to lower levels of job satisfaction (Renbarger & Davis, 2019).

Teacher cooperation

Higher levels of cooperation between teachers and staff within schools have been identified as a strong correlate of job satisfaction in previous research (Lopes & Oliveira, 2020). Examples of factors contributing to high levels of cooperation within a school could include team teaching, observation of other teachers' classes or sharing of teaching materials and resources (OECD, 2018). In fact, analysis of international data from TALIS 2013 which includes teachers from England, has shown teacher cooperation to be the most significant predictor of job satisfaction when accounting for other working conditions and teacher characteristics (Sims, 2017). Similar trends have also been found in Swedish data from TIMSS 2015, where cooperation has been identified as one of the strongest predictors of job satisfaction (Toropova et al., 2021). Although often seen as positive, teamwork can also have negative effects. Interview studies with Norwegian teachers, for example, have found that teamwork can sometimes be a source of stress, and disagreements can arise when teachers are unable to choose who they collaborate with (Skaalvik & Skaalvik, 2015).

In addition to having a positive effect on job satisfaction, teacher cooperation has been linked to lower levels of teacher turnover in the US (Nguyen, 2021), where teachers reporting higher levels of cooperation were found to be less likely to want to leave their current school. However, the same higher levels of cooperation were not associated with lower probabilities of teachers wanting to leave the teaching profession entirely.

Other factors

In our data, a public school is defined as any school managed by a public education authority, government agency, municipality or governing board appointed by government or elected by public franchise (OECD, 2018). Previous studies have found that job satisfaction is typically higher in private schools than in public schools. This discrepancy, however, is often attributed to the differing levels of autonomy (Lopes & Oliveira, 2020), or positive relationships with management (Sönmezer & Eryaman, 2008) which may be present in these two types of schools. Therefore, one might not expect to see a significant difference in the job satisfaction of public and private school teachers when controlling for these variables. Despite this, studies which have attempted to control for important policy, individual and workplace level characteristics have still found significantly higher levels of job satisfaction in private schools (Small, 2020). The effect on job satisfaction of teaching in a public vs. a private school is therefore an open question.

While larger class sizes and larger student teacher ratios have often been shown not to have a large effect on student achievement (Li & Konstantopoulos, 2017; Woessmann & West, 2006), a clear connection between class size and job satisfaction has not been established. One interview study of 200 teachers in the US, for instance, found that class size was one of the top 3 reasons reported by teachers as justifications for their current levels of job satisfaction (Perrachione et al., 2008). Other studies, however, have found that class size is not a major driver of American or Japanese teacher job satisfaction when controlling for other working conditions (Reeves et al., 2017).

A second factor which is less commonly examined in relation to teacher job satisfaction is the practice of out-of-field teaching. Out-of-field teaching has been linked to lower student achievement in a number of studies (Dee & Cohodes, 2008; Hill & Dalton, 2013), but the literature available on the effects that out-of-field teaching has on job satisfaction is quite limited. Olmos (2010) and Provasnik and Dorfman (2005) found that out-of-field teachers in the US were more prone to attrition, though other studies have not found as substantial an effect (e.g. Shen, 1997).

Finally, one additional factor which has not been the subject of much research in relation to teacher job satisfaction is contract-type. Our search for studies relating factors associated with the terms of a teacher's employment and their job satisfaction returned few results. Furthermore, those studies which we did find were not focused primarily on terms of employment, but instead used it as one of a variety of control variables, and results have varied across researchers. One investigation of the effect of personal characteristics on teacher job satisfaction, for example, found teachers with permanent contracts to be less satisfied on average (Gil-Flores, 2017). In contrast, Capone and Petrillo (2020) found teachers with permanent contracts to have higher levels of job satisfaction and well-being. Other studies which have investigated the effects of part-time or full-time contracts have revealed no discernible changes in job satisfaction (e.g. Ferguson et al., 2012).

Methods

Data and pre-processing

This study uses English data from TALIS 2018 (OECD, 2019a) which provides us with a representative sample of 4385 primary and lower secondary school teachers (2009 primary, 2376 lower secondary). Each observation includes more than 30 scales describing various teacher and school characteristics such as self-efficacy, participation in CPD and perceived cooperation among staff. The individual survey responses upon which these scales are based are also provided, as well as personal and background details for each of the teachers such as gender; school level; qualification and years' experience. A full list of all variables used can be found in Appendix 3. A description of how we handled missing data in these variables can be found at the end of this section.

The main variable of interest in this study is teacher job satisfaction. Teacher job satisfaction in the TALIS data is based on the responses of teachers to eight items which gauge a teacher's overall contentment and happiness with their current working environment and profession. All eight questions share a common stem which reads "We would like to know how you generally feel about your job. How strongly do you agree or disagree with the following statements?". An example item for measuring satisfaction with the working environment is "I enjoy working at this school", and an example item for satisfaction with the profession is "The advantages of being a teacher clearly outweigh the disadvantages". Possible responses to these items lie on a 4-point Likert scale, with options ranging from strongly disagree (1) to strongly agree (4). The ordinal responses to these items have been converted into a continuous measure of job satisfaction by the organisers of the TALIS study using an approach called confirmatory factor analysis. This continuous variable is the outcome we will use in our study. Confirmatory factor analysis is a very widely used approach in the social sciences (e.g. McInerney et al., 2018; Saloviita & Pakarinen, 2021). The organisers of TALIS have also conducted a number of tests to ensure the reliability and validity of the constructed teacher job satisfaction scale (OECD, 2019b). The resulting job satisfaction scale (after combining primary and lower secondary school teachers) has a mean of 12.42, and a standard deviation of 2.28.

To ensure a representative sample is collected during the data collection stages of TALIS, a stratified two-stage probability sampling design is used within each country. Each teacher within the TALIS dataset is therefore assigned a number of weights for the purposes of rigorously calculating population parameters of interest and their associated standard errors. The sampling weights resulting from this design were fully accounted for in our analysis. This was accomplished by using the Balanced Repeated Replication procedure described in the TALIS technical report (OECD, 2019b). The resulting confidence intervals are shown in Figure 3.

Data from the survey can be missing for a number of reasons. Some teachers did not reach every question, and others did not answer personal questions such as those concerning their age. Of the variables we have used, 52 contained missing values, with on average 8% of the data missing. In order to maximise the data available for use, we have imputed these missing responses with the R package missRanger (Mayer, 2019). This procedure involves substituting missing values with responses based on an individual's answers to all of the other questions in the survey. This enables us to retain information that would otherwise be lost if missing cases were deleted, and is more accurate than other approaches which simply use the mean value for imputation (Stekhoven & Bühlmann, 2012). One drawback of this, however, is that the uncertainty related to the imputation of these missing values is not captured in our main analysis. This should be borne in mind when interpreting our main results later, as the true 95% confidence intervals are likely to be slightly broader.

Traditional approaches

Commonly used approaches in international large-scale assessments to investigate the relationship between a set of independent variables, X, and a dependent variable, Y, include ordinary least squares regression and other more sophisticated modelling approaches such as multilevel models. These approaches are very useful but have a number of limitations. Firstly, they assume a linear relationship between each independent variable and the outcome of interest. This can lead to biased parameter estimates in some cases and can lead one to believe that there is no relationship between two variables when in fact there is. For example, the relationship between teacher attrition and age has been found to be U shaped in a number of different studies (Boe et al., 1997; Guarino et al., 2006).

A second limitation is that due to the cross-sectional and observational nature of the survey data, it is not possible to make any causal claims. Furthermore, the direction of the relationship is not always possible to determine. Teacher self-efficacy, for example, has usually been assumed to be an antecedent of job satisfaction in much of the literature, but a recent study by Burić and Kim (2021) finds the causal direction may actually be the opposite.

Thirdly, linear models can become difficult to interpret when a large number of covariates have been included as explanatory variables. This means that it can be difficult to control for a large number of factors simultaneously when investigating the association of one variable of interest with another while still maintaining the required interpretability. Consequently, researchers often limit their analysis to a smaller subset of the available data. However, not controlling for some variables may bias parameter estimates.

In the third section, we have concentrated on factors which relate to measures that school principals or other policy makers could introduce immediately with the view to improving job satisfaction levels. By contrast, much of the existing literature on teacher job satisfaction uses scale scores of different psychological constructs which have been validated using approaches such as confirmatory factor analysis (e.g. McInerney et al., 2018; Saloviita & Pakarinen, 2021). Such approaches are certainly useful because, for example, they have demonstrated a link between higher levels of teacher self-efficacy and job satisfaction. Teachers' levels of self-efficacy are not easily changed however, and so these results do not provide a directly implementable process that can be used to improve job satisfaction or the outcome of interest.

Bayesian additive regression trees for causal analysis

With the above considerations in mind, this study aims to investigate the effect of a number of binary factors, which we call treatments, on teacher job satisfaction. Our approach will be to use the R package bartCause which is a causal inference machine learning package for the R programming language (Dorie & Hill, 2020; R Core Team, 2021). The bartCause package allows us to estimate causal effects, and has been demonstrated to be highly competitive in causal inference machine learning competitions (Dorie

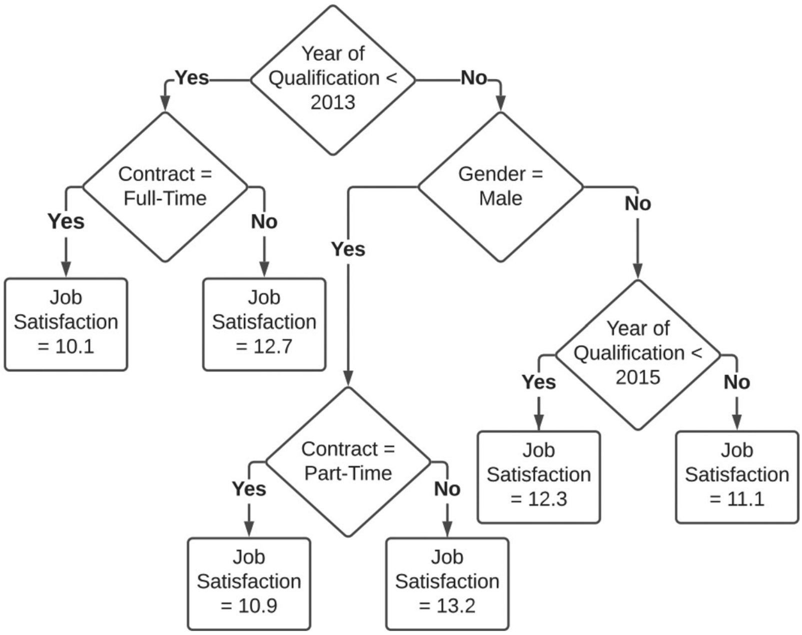

Figure 1. Example of a single decision tree for the TALIS data. Each teacher's information can be fed into the tree by following the decision rules. The terminal nodes provide the predictions for the job satisfaction of each teacher. In practice the BART model works by creating many different decision trees and summing the predictions together.

et al., 2019). The package owes its success to the impressive prediction capabilities of Bayesian Additive Regression Trees (BART), a Bayesian non-parametric modelling tool which is well suited to a wide variety of problems (Chipman et al., 2010). This flexibility has inspired a large number of BART extensions, such as a recent paper which shows how BART-based models can be adapted for use in a mediation analysis setting (Linero & Zhang, 2022).

BART is known as a sum of trees model which can flexibly and accurately predict an outcome of interest Y using a set of covariates X. It can be seen as an extension of regression modelling that automatically identifies interactions and non-linear relationships between the variables. In the case of a single tree model, BART makes predictions by establishing a set of decision rules which when followed, assign a prediction to each observation. See Figure 1 for an example of a decision tree.

Bayesian methods are becoming increasingly popular in educational research (König & van de Schoot, 2018). In particular, a recent study has used BART to estimate the causal effects of private tuition on student achievement (Suk et al., 2021). BART has also been used extensively in other fields outside of education, and is a popular choice for many quantitative researchers (e.g. Prado et al., 2021).

Treatment effect estimation

To evaluate whether a particular covariate is a causal predictor of the outcome variable we adopt the Neyman-Rubin causal model (Rubin, 1974; Sekhon, 2008; Splawa-Neyman et al., 1990). Central to the Neyman-Rubin causal model is the concept of potential outcomes which posits that there are two potential outcomes for each individual i, one that would be observed under treatment, $y_i(1)$, and one that would be observed under control, $y_i(0)$, (no treatment). The individual treatment effect would then be given by the difference between these potential outcomes: $\tau_i = y_i(1) - y_i(0)$. Observing individual i simultaneously under both treatment and control is impossible, however, and this is known as the fundamental problem of causal inference.

Estimation of τ is a difficult task, especially in the case of observational data. Challenges posed by observational data to estimating causal effects include the fact that individuals are not randomly assigned to the treatment and control groups, and that our observation of the data may not include all variables which have an influence on the outcome of interest or the non-random assignment mechanism. It is, however, possible to identify causal effects with a number of key assumptions (Kurz, 2021). These assumptions include:

(1) The stable unit treatment value assumption (SUTVA). It requires that the treatment status of an individual i does not affect the potential outcomes of any other individual j.
(2) The ignorability assumption. This requires that the potential outcomes of individual i must be independent of their treatment status conditioned on their observed covariates. In other words, we require there to be no confounding variables we have not observed.
(3) The overlap assumption. This requires that every individual must have a non-zero probability of being assigned to both treatment conditions.

Assuming the above assumptions hold, the bartCause package estimates treatment effects with BART by predicting the two potential outcomes for each individual, using their observed characteristics and an indicator of whether or not they received treatment as predictor variables. Hill (2011) showed that the average treatment effect (ATE), can then be estimated as:

$$\frac{1}{N}\sum_{i=1}^{N}\widehat{\tau_i} = \frac{1}{N}\sum_{i=1}^{N}\widehat{y_i}(1) - \widehat{y_i}(0)$$

Including propensity scores in causal models

Following the advice of Hahn et al. (2020), we will include an additional independent variable as a predictor in our model. This additional variable is known as the propensity score, and is defined as an individual's probability of being assigned to the treatment group. This probability can be estimated from an individual's characteristics such as their gender, year of qualification, degree type, etc. Logistic regression is a common choice for this task, but we have chosen to use BART instead to keep our approach as consistent as possible and retain the superior predictive approach.

The inclusion of the propensity score has been shown to improve the estimation of treatment effects (Hahn et al., 2020). Besides this practical advantage, it can also be interesting to look at different trends in the propensity scores for individual teachers. Analysing such trends allows us to identify, for example, which subgroups of teachers are particularly likely to belong to positive or negative treatment groups. This process can identify specific subgroups of teachers who need to be given extra support, or who would benefit from being assigned to a particular treatment group. We highlight some examples of this in our next section.

Choice of treatment variables

We calculate average treatment effects for each of the following treatment options (short names or abbreviations used in Figures are shown in brackets):

(1) Did the teacher take part in at least 4 CPD activities in the past year (CPD)?
(2) Did they take part in a formal/informal induction programme when they started teaching at their current school (Induction)?
(3) Do they take part in observing other teachers (Observing)?
(4) Do they take part in team teaching (Team Teaching)?
(5) Do they have a mentor (Has Mentor)?
(6) Are they a mentor to another teacher (Is Mentor)?
(7) Do they teach in a publicly managed school (Public School)? (Full definition in Appendix 1)
(8) Do they have 30 or more students in their class (30+ Students)?
(9) Are they an out-of-field teacher (Out-of-field)?
(10) Do they have a part-time contract (Part-Time)?

In each of the cases above, the ATE is estimated independently of the other treatments. The set of predictor variables included in X remains unchanged, as we control for the same covariates in every assignment option (with a few exceptions). For an exact definition of each of these treatments see Appendix 1. Appendix 3 identifies any variables which were removed from X for a specific assignment option. For example, it would not make sense to control the number of students in a class when investigating the effect of teaching a class with 30 or more students (Variable Code: TT3G38).

Figure 2 shows the control and treatment group sizes for the different factors that we have created and are investigating. The control group size for CPD is 1618, meaning that 37% (unweighted) of teachers in the sample did not take part in 4 or more CPD events over the course of the past year. The treatment group size for this assignment option is 2767, corresponding to a 63% participation rate in at least 4 CPD events. The other segments of the plot have similar interpretations.

As can be seen from Figure 2, 30% of teachers met our criteria for teaching out-of-field. A more in-depth analysis of these numbers reveals that 24% of secondary school teachers meet these criteria, and 37% of primary school teachers do. Further investigations also show that the subjects being taught out-of-field by teachers are different across the two school levels. We bring this point to the reader's attention to make clear that these teachers are all treated identically, and we do not make careful distinctions between reasons for teaching out-of-field. Furthermore, we do not distinguish between

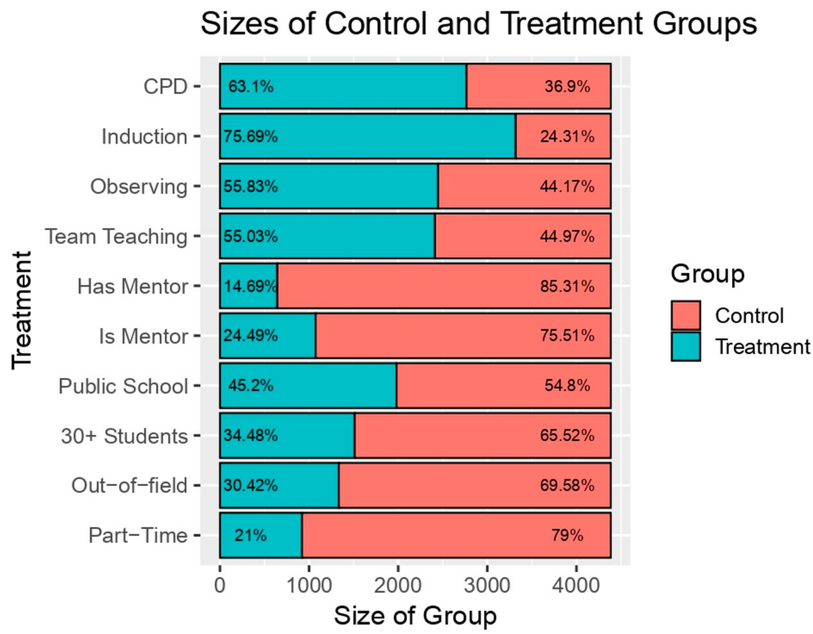

Figure 2. Percentage of teachers belonging to the control and treatment groups under investigation. There are different levels of balance across the groups.

primary vs. secondary school teachers for this treatment effect (or indeed any of the other treatment effects).

Results

This section describes the results from:

(1) Choosing a treatment assignment option to consider from the list in Section "Choice of treatment variables".
(2) Estimating the average treatment effect of this assignment option on job satisfaction.

For a visual representation of these results, see Figure 3 which indicates the final estimate and 95% confidence interval for each of the treatment effects. Diagnostic tests were also performed for all models fitted to the data to ensure convergence had been reached, and goodness of fit statistics were calculated to ensure satisfactory predictive performance.

Continual professional development

Our results identify participation in at least 4 CPD events over the course of a year as having a positive effect on teacher job satisfaction. The 95% confidence interval for this average treatment effect is [0.035, 0.309]. To give an idea of the magnitude of this treatment effect, consider that the teacher job satisfaction scale has a mean of 12.42, and a standard deviation of 2.28. Therefore, the centre-point of this confidence interval which

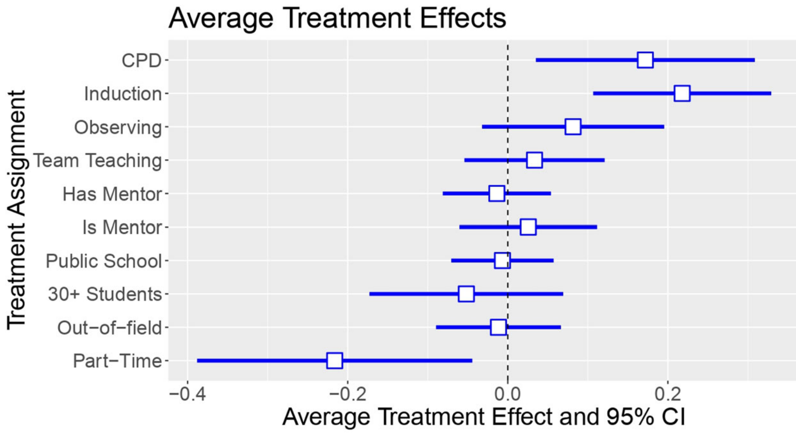

Figure 3. Plot of average treatment effects.

is at 0.172 would correspond to an increase in job satisfaction of 0.08 standard deviations, which is a small but positive improvement.

Induction and mentoring programmes

Our results show that taking part in induction when starting at a new school has a positive effect on job satisfaction. The 95% confidence interval for the average treatment effect on job satisfaction is [0.107, 0.329]. Therefore, taking part in an induction scheme is associated with a mean increase in job satisfaction of 0.218 meaning that induction schemes are the most beneficial of all of the treatment assignment options we have considered.

Mentoring, however, is not identified as having a strongly positive effect. As can be seen from the 95% confidence intervals in Figure 3, this is true for both mentors and mentees.

Observation and team teaching

Team teaching and observing the lessons of other teachers are both identified as having a positive effect on job satisfaction. The uncertainty in these estimates is quite large however, and this is reflected in the wide 95% confidence intervals shown in Figure 3 which both include zero within their range. Given the large confidence intervals, it may be that there are large effects of these variables, but the data here do not provide us with enough information to estimate them precisely. Alternatively, there may be subgroups for whom the causal effect is particularly high or low. This, however, would also be difficult to ascertain with a high degree of statistical confidence.

Other factors

Of the remaining factors we considered, the treatment assignment option with the largest effect in relation to job satisfaction is the possession of a part-time contract of less than 90% of a typical full-time contract's hours. This factor has the effect of reducing job

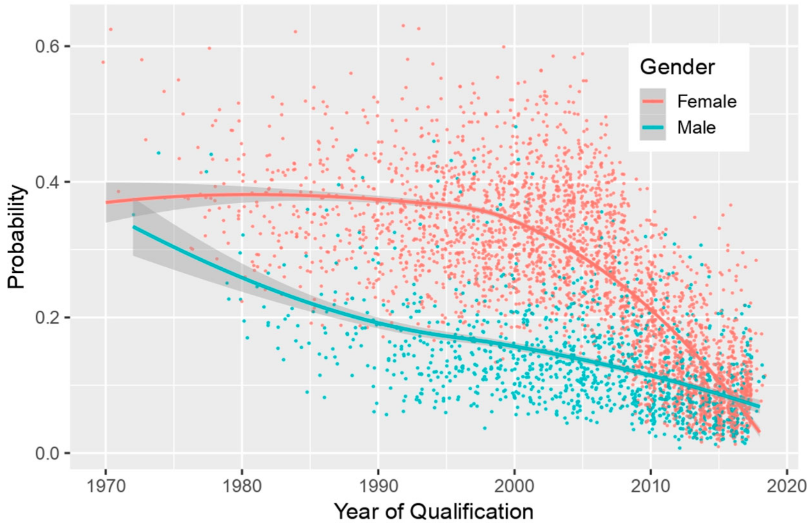

Figure 4. Probability of having a part-time contract. Female teachers have higher probabilities than male teachers, especially more experienced female teachers.

satisfaction on average by 0.216, 95% confidence interval [−0.388, −0.044]. The results from analysing the propensity scores for this factor show an interesting trend. Figure 4 shows that experienced female teachers generally have much higher propensity scores (probability of being assigned to the treatment group) than their male colleagues.

The other factors we have considered are out-of-field teaching, working in a public vs. a private school, and teaching a class with 30 or more students. According to our results, these factors are not associated with a strong effect on job satisfaction. We emphasise again that these factors may indeed be very important, but the precision with which the data allows us to estimate these effects is insufficient to make such claims with a high degree of statistical confidence in this case.

Discussion

We begin by discussing our main findings in more detail, and go on to highlight some key aspects of this study which make a new and important contribution to the literature on teacher job satisfaction. We finish this section by drawing the reader's attention to some limitations of this study, and by suggesting areas for future research.

Main findings

Continual professional development

Our results identify high levels of participation in CPD as having a positive effect on teacher job satisfaction. This is in agreement with multiple studies which have found a strong correlation between CPD and job satisfaction (e.g. Sims, 2017; Wang et al., 2020; Yoon & Kim, 2022). Crucially, our result supports these previous findings by verifying the strong positive effects of CPD using a causal inference approach, and thus we are able to infer results about causation and not just correlation. Furthermore, as job satisfaction is known to be an important predictor of teacher intentions to remain teaching (Madigan & Kim, 2021), our results also support recent findings from studies of teachers based in England which have linked CPD to higher levels of retention (e.g. Allen & Sims, 2017; Coldwell, 2017). In addition, we have ensured that our treatment effect estimates are as unbiased as possible, by removing the effect of possible confounding variables on our outcome of interest.

We noted that only 63% of teachers in the English dataset have reached this high level of CPD. Barriers to participation in CPD are known to be a key predictor of job satisfaction (Zhang et al., 2020). Our results therefore also provide strong support for this body of work by demonstrating the positive gains that can be made by removing such barriers and encouraging and enabling more teachers to engage in CPD events. For this reason, the emphasis on the importance of engagement with CPD in the Early Career Framework in England is very welcome (Department for Education, 2019). In addition, the inclusion of a 10% and 5% reduction in timetabled teaching hours for teachers in their first and second years, in order to enable them to fully avail of the supports and training offered during this time is likely to be crucial.

We highlight the fact that our investigation has only considered a binary version of CPD. In reality, however, levels of attendance at CPD belong on a spectrum, not just high/low. Furthermore, the benefits from CPD are likely to depend on many factors such as the quality and relevance of the training to a teacher's needs. These factors warrant further investigation but were beyond the scope of this study. Despite this, we do find clear evidence in favour of recommending CPD as a measure for improving job satisfaction.

Induction and mentoring programmes

Our finding that induction schemes have a very positive effect on job satisfaction agrees with prior work from Ronfeldt and McQueen (2017). Contrary to the review by Allen (2005), we did not detect high levels of heterogeneity in the treatment effect estimates of this assignment option. The recent introduction of the Early Career Framework in England which includes mandatory induction for new teachers is therefore an excellent step forward, but we argue that induction schemes should also be made available more generally for all new teachers at a school, regardless of number of years qualified or experience in the classroom.

Unlike some previous studies (e.g. Ingersoll & Strong, 2011; Renbarger & Davis, 2019), our results do not identify the presence of a mentor as being beneficial for job satisfaction. There are a number of plausible reasons for this. Firstly, there may be some unobserved or unaccounted for confounding variables common to schools with mentorship schemes which bias the estimates of these analyses. Secondly, we did not consider other aspects related to mentoring, such as the subject area of the mentor. Research has shown that a mentee is more likely to benefit from a mentoring arrangement if their mentor is a teacher from the same grade level (Parker et al., 2009). Other factors such as the mentoring quality and the frequency of meetings can also be important (Richter et al., 2013). The provision of training for mentors taking part in the Early Career Framework to improve mentoring quality is therefore commended.

An indicator of whether or not a teacher is currently a mentor to another member of staff was also included as a component in our analysis. Similarly, we did not find that this treatment was associated with an appreciable increase or decrease in job satisfaction. Again, this could be a result of our binary view of mentoring relationships, in which we only consider the presence or absence of a mentee, and fail to account for other aspects such as the quality of the mentoring relationship, which has been demonstrated to be an important predictor of job satisfaction (Lunsford et al., 2018).

Observation and team teaching

The fact that we have not found a clear link between team teaching or observation with job satisfaction may initially appear to be strange. The literature reviewed consistently pointed towards higher levels of teamwork and cooperation as having a positive effect on teacher job satisfaction. Therefore, we might have expected to see this reflected in our results also.

One plausible explanation for this is that higher levels of teamwork and cooperation within a school are difficult to attribute to a small number of specific practices such as team teaching and observation. Higher levels of teamwork and cooperation within a school are characterised by many different aspects such as sharing resources with colleagues and collaborating together on different projects etc. As a result, it is difficult to capture the true impact of higher levels of teamwork and cooperation as a whole by only considering two of a much larger number of indicators. Therefore, the absence of a large effect size here does not necessarily mean that team teaching and observation are not useful practices. Rather, the results indicate that only implementing one or two of these factors is unlikely to yield significant improvements in job satisfaction, and efforts should instead be focused on improving teamwork and cooperation as a whole. This is made clear by the very small treatment effect sizes that result from us considering two of these practices in isolation.

Other factors

We investigated whether working at a publicly owned and managed school affects job satisfaction. The results from our approach do not identify a significant causal effect for this treatment assignment. This result is in line with work by Dahler-Larsen and Foged (2018) who attribute the difference in job satisfaction between public and private schools to differences in organisational characteristics, as opposed to the ownership of the school.

In line with research by Reeves et al. (2017), our results show that teaching a class with 30 or more students does not have a large effect on job satisfaction. We should note, however, that our finding is based on a cut-off point of 30 students. This value was chosen to ensure an approximately even split of teachers in the treatment and control groups. It is possible, however, that a different value would yield different results, and teachers at the more extreme end of the distribution with greater than 35 students may experience a more negative effect from this treatment.

Given the lack of research linking out-of-field teaching to job satisfaction we thought it was important to include this as a factor in our study. The magnitude of the treatment effect that we have obtained in our results for this factor is very small, but out-of-field teaching is a complex phenomenon (Hobbs & Törner, 2019), so it is reasonable to expect that the effects of teaching out-of-field may be dependent on a number of contextual factors such as how dissimilar the subject being taught is to one's area of expertise. A more detailed investigation of the effects of out-of-field teaching on job satisfaction is therefore warranted.

As in the study by Ferguson et al. (2012), the contract-type used in our study refers to full-time or part-time contracts. We have chosen this as it will allow us to have more evenly balanced control and treatment groups. Our results show that teachers on a part-time contract are less satisfied with their career than their full-time colleagues. Also, an analysis of the propensity scores for this treatment effect shows that experienced female teachers have much higher probabilities of being on part-time contracts than their male counterparts. Future research should investigate the reasons for this, and supports that might be put in place for teachers with childcare responsibilities.

In summary, of the factors that we have investigated, we have found that levels of participation in CPD and induction schemes have the strongest positive influence on job satisfaction. Conversely, we have also found that possessing a part-time contract can have a negative effect on job satisfaction. In the case of the other treatments we have studied, despite the average treatment effects often pointing in the direction we had expected, there was not enough certainty in these estimates to claim the presence of a clear causal effect.

Contribution of this paper

We believe this study makes three main contributions to the current literature on teacher job satisfaction. The first is that we have employed a causal inference machine learning approach, bringing the power of advanced statistical modelling techniques to an important problem in the world of education. One advantage of this approach is the ability to flexibly model job satisfaction without assuming a linear relationship between the predictor variables and the outcome of interest which is a common feature of most conventional statistical models. This approach is also well suited to detecting interactions between variables and allows us to include a much wider variety of covariates than would normally be possible when using linear models. This is absolutely crucial, because it enables us to model the response surface using the propensity score along with a large number of other variables, thus accounting for many potential sources of confounding which could otherwise bias treatment effect estimates.

Second, instead of identifying important characteristics related to a teacher's working environment such as cooperation, quality of school leadership or personal traits such as self-efficacy, we have established several specific and implementable measures that may be introduced in an attempt to improve job satisfaction. We summarise our findings with the following recommendations:

(1) Our results provide strong evidence that participation in an induction scheme when starting at a school can have a beneficial effect on teacher job satisfaction. We therefore recommend that schools not currently offering such schemes should endeavour to introduce them. We also recommend that schools currently offering induction schemes should encourage participation from all new staff, including experienced and novice teachers.

(2) We also find strong support for higher levels of participation in continual professional development having a positive effect on job satisfaction. Therefore we suggest that school authorities should make it a priority to identify and remove any barriers to staff attendance at CPD events, whilst also ensuring a regular calendar of relevant CPD activities are available for attendance.
(3) Our finding that part-time contracts are negatively impacting on the job satisfaction of affected teachers warrants a closer examination of how concerns about job security may be affecting teachers. The fact experienced female teachers are disproportionately more likely to be on a part-time contract also requires a review into the supports that may be put in place for teachers with young families.

Specific recommendations are important because although it may be known that certain factors such as stress are negatively correlated with job satisfaction (Klassen & Chiu, 2010), it is not always obvious how best to reduce stress levels among teachers, or if a set of proposed changes will have the desired effect. This study therefore avoids this pitfall by identifying factors such as induction schemes which can be beneficial for job satisfaction, while also identifying the negative effects of factors such as part-time contracts.

Finally, the propensity scores described in Section "Including propensity scores in causal models", although not the primary focus of this study, provide us with an interesting insight into the types of teachers more likely to belong to the treatment and control groups we have investigated. This can help us to identify certain subgroups of teachers who have not availed of positive treatments, and we can then ensure that these activities are made available to them. This can also help us to identify subgroups of teachers who are more likely to be exposed to the negative effects of a treatment, such as experienced female teachers who we found were significantly more likely to have a part-time contract.

Limitations and areas for future research

As discussed in the methodology, the causal inference approach that we have employed makes a number of important assumptions. Among these is the ignorability assumption, which requires that we have accounted for all potential sources of confounding when investigating a given treatment. Despite including a wide variety of control variables in our design matrix, X, it is certainly still possible that there may be some confounding variables not collected as part of the survey. Teachers with young children for example, may be more likely to work part-time, but there is no indication in the TALIS data whether teachers have young children. Future research could include a detailed assessment of the reasonableness of these assumptions in relation to TALIS by incorporating data from external sources, and using different diagnostic methods designed to assess these assumptions.

A second limitation of our approach is that some aspects of the working environment such as teamwork and cooperation are very difficult to capture with binary variables. Therefore, it may be less meaningful to investigate binary factors in relation to aspects such as this, because levels of teamwork and cooperation cannot be fully characterised by a simple dummy variable. Also, hours of CPD attended and the number of students in a class are both continuous variables. Therefore, their impact on job satisfaction cannot be fully appreciated by artificially converting them into a binary factor. Additional studies which use causal inference machine learning algorithms designed to handle continuous treatment variables may be better suited to this task.

Finally, as our results are based on data which only includes teachers from England, we cannot claim that the same treatment effects would be observed in other countries and cultures, where other factors may be more important for improving the job satisfaction of teachers. The application of a similar approach to ours, but to different countries in the TALIS data is therefore a promising area for future research.

Conclusion

Faced with increasing demand for qualified teachers in England and internationally, it is of the utmost importance to identify strategies for improving teacher job satisfaction. This can help to encourage higher retention rates of qualified teachers, and attract new entrants to start their career. Many studies which investigate factors associated with job satisfaction, however, instead of identifying specific and implementable measures for achieving this task, link higher levels of job satisfaction to positive working environments or higher levels of self-efficacy. Our study has tackled this issue by employing a causal inference machine learning approach to investigate the effect of a number of treatments on job satisfaction. We encourage school management teams and educational administrations to take note of our results which further support the provision of induction schemes for new teachers, and continual professional development for all staff. We also recommend an examination of how part-time contracts may be causing anxiety around job security and satisfaction for teachers. More generally, we advocate for further research into the specific steps that may be taken for improving job satisfaction through the use of causal inference methods.

Acknowledgements

The authors would like to thank Dr Nikolaos Fotou for sharing his knowledge and a number of resources with us. We are also grateful for the many helpful comments from the anonymous referees.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work has emanated from research conducted with the financial support of Science Foundation Ireland under grant number 18/CRT/6049. In addition, Andrew Parnell's work was supported by: a Science Foundation Ireland Career Development Award (17/CDA/4695) and SFI Research Centre award (12/RC/2289_P2). For the purpose of Open Access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. The preprint of this manuscript can be found at https://edarxiv.org/nasq9/ (McJames et al., 2022).

ORCID

Nathan McJames http://orcid.org/0000-0003-3349-0635

References

Allen, M. B. (2005). *Eight questions on teacher recruitment and retention: What does the research say?* Education Commission of the States (NJ3).

Allen, R., & Sims, S. (2017). *Improving science teacher retention: Do national STEM learning network professional development courses keep science teachers in the classroom*. Education Research in STEM Subjects: Main Collection.

Boe, E. E., Bobbitt, S. A., Cook, L. H., Whitener, S. D., & Weber, A. L. (1997). Why didst thou go? Predictors of retention, transfer, and attrition of special and general education teachers from a national perspective. *The Journal of Special Education*, *30*(4), 390–411. https://doi.org/10.1177/002246699703000403

Burić, I., & Kim, L. E. (2021). Job satisfaction predicts teacher self-efficacy and the association is invariant: Examinations using TALIS 2018 data and longitudinal Croatian data. *Teaching and Teacher Education*, *105*, Article 103406. https://doi.org/10.1016/j.tate.2021.103406

Capone, V., & Petrillo, G. (2020). Mental health in teachers: Relationships with job satisfaction, efficacy beliefs, burnout and depression. *Current Psychology*, *39*(5), 1757–1766. https://doi.org/10.1007/s12144-018-9878-7

Chipman, H. A., George, E. I., & McCulloch, R. E. (2010). BART: Bayesian additive regression trees. *The Annals of Applied Statistics*, *4*(1), 266–298. https://doi.org/10.1214/09-AOAS285

Coldwell, M. (2017). Exploring the influence of professional development on teacher careers: A path model approach. *Teaching and Teacher Education*, *61*, 189–198. https://doi.org/10.1016/j.tate.2016.10.015

Dahler-Larsen, P., & Foged, S. K. (2018). Job satisfaction in public and private schools: Competition is key. *Social Policy & Administration*, *52*(5), 1084–1105. https://doi.org/10.1111/spol.12324

Daly, C., Gandolfi, H., Pillinger, C., Glegg, P., Hardman, M., Stiasny, B., & Taylor, B. (2021). The early career framework–a guide for mentors and early career teachers. www.ucl.ac.uk/ioe-cttr

Dee, T. S., & Cohodes, S. R. (2008). Out-of-field teachers and student achievement: Evidence from matched-pairs comparisons. *Public Finance Review*, *36*(1), 7–32. https://doi.org/10.1177/1091142106289330

Department for Education. (2019). *Early career framework*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/978358/Early-Career_Framework_April_2021.pdf

Dorie, V., & Hill, J. (2020). Package 'bartCause'. https://cran.r-project.org/web/packages/bartCause/bartCause.pdf

Dorie, V., Hill, J., Shalit, U., Scott, M., & Cervone, D. (2019). Automated versus do-it-yourself methods for causal inference: Lessons learned from a data analysis competition. *Statistical Science*, *34*(1), 43–68. https://doi.org/10.1214/18-STS667

Ferguson, K., Frost, L., & Hall, D. (2012). Predicting teacher anxiety, depression, and job satisfaction. *Journal of Teaching and Learning*, *8*(1), 27–42. https://doi.org/10.22329/jtl.v8i1.2896

Gil-Flores, J. (2017). The role of personal characteristics and school characteristics in explaining teacher job satisfaction. *Revista de Psicodidáctica (English ed.)*, *22*(1), 16–22. https://doi.org/10.1387/RevPsicodidact.15501

Gray, L., & Taie, S. (2015). *Public school teacher attrition and mobility in the first five years: Results from the first through fifth waves of the 2007-08 beginning teacher longitudinal study*. First look. National Center for Education Statistics.

Guarino, C. M., Santibanez, L., & Daley, G. A. (2006). Teacher recruitment and retention: A review of the recent empirical literature. *Review of Educational Research*, *76*(2), 173–208. https://doi.org/10.3102/00346543076002173

Hahn, P. R., Murray, J. S., & Carvalho, C. M. (2020). Bayesian regression tree models for causal inference: Regularization, confounding, and heterogeneous effects (with discussion). *Bayesian Analysis*, *15*(3), 965–1056. https://doi.org/10.1214/19-BA1195

Han, D., & Hur, H. (2021). Managing turnover of STEM teacher workforce. *Education and Urban Society*, *54*(2), 205–222. https://doi.org/10.1177/00131245211053562

Hill, J. G., & Dalton, B. (2013). Student math achievement and out-of-field teaching. *Educational Researcher*, *42*(7), 403–405. https://doi.org/10.3102/0013189X13505682

Hill, J. L. (2011). Bayesian nonparametric modeling for causal inference. *Journal of Computational and Graphical Statistics*, *20*(1), 217–240. https://doi.org/10.1198/jcgs.2010.08162

Hilton, G. L. (2017). *Disappearing teachers: An exploration of a variety of views as to the causes of the problems affecting teacher recruitment and retention in England*. Bulgarian Comparative Education Society.

Hobbs, L., & Törner, G. (2019). Teaching out-of-field as a phenomenon and research problem. In L. Hobbs & G. Törner (Eds.), *Examining the phenomenon of "teaching out-of-field"* (pp. 3–20). Springer.

Hulme, M., & Wood, J. (2022). The importance of starting well: The influence of early career support on job satisfaction and career intentions in teaching. *Journal of Further and Higher Education*, *46*(4), 504–521. https://doi.org/10.1080/0309877X.2021.1985981

Ingersoll, R., & Kralik, J. M. (2004). *The impact of mentoring on teacher retention: What the research says*. Education Commission of the States. https://www.gse.upenn.edu/pdf/rmi/ECS-RMI-2004.pdf

Ingersoll, R. M., & Strong, M. (2011). The impact of induction and mentoring programs for beginning teachers: A critical review of the research. *Review of Educational Research*, *81*(2), 201–233. https://doi.org/10.3102/0034654311403323

Klassen, R. M., & Chiu, M. M. (2010). Effects on teachers' self-efficacy and job satisfaction: Teacher gender, years of experience, and job stress. *Journal of Educational Psychology*, *102*(3), 741–756. https://doi.org/10.1037/a0019237

Klusmann, U., Kunter, M., Trautwein, U., Lüdtke, O., & Baumert, J. (2008). Teachers' occupational well-being and quality of instruction: The important role of self-regulatory patterns. *Journal of Educational Psychology*, *100*(3), 702–715. https://doi.org/10.1037/0022-0663.100.3.702

König, C., & van de Schoot, R. (2018). Bayesian statistics in educational research: A look at the current state of affairs. *Educational Review*, *70*(4), 486–509. https://doi.org/10.1080/00131911.2017.1350636

Kurz, C. F. (2021). Augmented inverse probability weighting and the double robustness property. *Medical Decision Making*, *42*(2), 156–167. https://doi.org/10.1177/0272989X211027181

Levy, A. J., Joy, L., Ellis, P., Jablonski, E., & Karelitz, T. M. (2012). Estimating teacher turnover costs: A case study. *Journal of Education Finance*, *38*(2), 102–129. https://www.jstor.org/stable/23353968

Li, W., & Konstantopoulos, S. (2017). Does class-size reduction close the achievement gap? Evidence from TIMSS 2011. *School Effectiveness and School Improvement*, *28*(2), 292–313. https://doi.org/10.1080/09243453.2017.1280062

Linero, A. R., & Zhang, Q. (2022). Mediation analysis using Bayesian tree ensembles. *Psychological Methods*. Advance online publication. https://doi.org/10.1037/met0000504

Locke, E. A. (1969). What is job satisfaction? *Organizational Behavior and Human Performance*, *4*(4), 309–336. https://doi.org/10.1016/0030-5073(69)90013-0

Long, R., & Danechi, S. (2021). *Teacher recruitment and retention in England* (House of Commons Briefing Paper, Number 07222).

Lopes, J., & Oliveira, C. (2020). Teacher and school determinants of teacher job satisfaction: A multilevel analysis. *School Effectiveness and School Improvement*, *31*(4), 641–659. https://doi.org/10.1080/09243453.2020.1764593

Lunsford, L., Baker, V., & Pifer, M. (2018). Faculty mentoring faculty: Career stages, relationship quality, and job satisfaction. *International Journal of Mentoring and Coaching in Education*, *7*(2), 139–154. https://doi.org/10.1108/IJMCE-08-2017-0055

Madigan, D. J., & Kim, L. E. (2021). Towards an understanding of teacher attrition: A meta-analysis of burnout, job satisfaction, and teachers' intentions to quit. *Teaching and Teacher Education*, *105*, Article 103425. https://doi.org/10.1016/j.tate.2021.103425

Mayer, M. (2019). Package 'missRanger'. https://cran.r-project.org/web/packages/missRanger/missRanger.pdf

McInerney, D. M., Korpershoek, H., Wang, H., & Morin, A. J. (2018). Teachers' occupational attributes and their psychological wellbeing, job satisfaction, occupational self-concept and quitting intentions. *Teaching and Teacher Education*, *71*, 145–158. https://doi.org/10.1016/j.tate.2017.12.020

McJames, N., Parnell, A., & O'Shea, A. (2022). *Factors affecting teacher job satisfaction and retention: A causal inference machine learning approach using data from TALIS 2018*. EdArXiv.

Nguyen, T. D. (2021). Linking school organizational characteristics and teacher retention: Evidence from repeated cross-sectional national data. *Teaching and Teacher Education*, *97*, Article 103220. https://doi.org/10.1016/j.tate.2020.103220

O'Doherty, T., & Harford, J. (2018). Teacher recruitment: Reflections from Ireland on the current crisis in teacher supply. *European Journal of Teacher Education*, *41*(5), 654–669. https://doi.org/10.1080/02619768.2018.1532994

OECD. (2018). *TALIS 2018 teacher questionnaire*. https://www.oecd.org/education/school/TALIS-2018-MS-Teacher-Questionnaire-ENG.pdf

OECD. (2019a). *TALIS 2018 and TALIS starting strong 2018 user guide*. https://www.oecd.org/education/talis/TALIS_2018-TALIS_Starting_Strong_2018_User_Guide.pdf

OECD. (2019b). *TALIS 2018 technical report. Chapter 11. Validation of scales and construction of scale scores*. https://www.oecd.org/education/talis/TALIS_2018_Technical_Report.pdf

Olmos, F. (2010). *Square peg in a round hole: Out-of-field teaching and its impact on teacher attrition*. University of California, Irvine and California State University.

Parker, M. A., Ndoye, A., & Imig, S. R. (2009). Keeping our teachers! Investigating mentoring practices to support and retain novice educators. *Mentoring & Tutoring: Partnership in Learning*, *17*(4), 329–341. https://doi.org/10.1080/13611260903391500

Perrachione, B. A., Rosser, V. J., & Petersen, G. J. (2008). Why do they stay? Elementary teachers' perceptions of job satisfaction and retention. *Professional Educator*, *32*(2), 1–17. https://files.eric.ed.gov/fulltext/EJ862759.pdf

Prado, E. B., Parnell, A. C., McJames, N., O'Shea, A., & Moral, R. A. (2021). *Semiparametric Bayesian additive regression trees*. arXiv preprint arXiv:2108.07636.

Provasnik, S., & Dorfman, S. (2005). *Mobility in the teacher workforce findings from the condition of education 2005*. National Center for Education Statistics.

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Reeves, P. M., Pun, W. H., & Chung, K. S. (2017). Influence of teacher collaboration on job satisfaction and student achievement. *Teaching and Teacher Education*, *67*, 227–236. https://doi.org/10.1016/j.tate.2017.06.016

Reeves, T. D., Hamilton, V., & Onder, Y. (2022). Which teacher induction practices work? Linking forms of induction to teacher practices, self-efficacy, and job satisfaction. *Teaching and Teacher Education*, *109*, Article 103546. https://doi.org/10.1016/j.tate.2021.103546

Renbarger, R., & Davis, B. (2019). Mentors, self-efficacy, or professional development: Which mediate job satisfaction for new teachers? A regression examination. *Journal of Teacher Education and Educators*, *8*(1), 21–34. https://files.eric.ed.gov/fulltext/EJ1214939.pdf

Richter, D., Kunter, M., Lüdtke, O., Klusmann, U., Anders, Y., & Baumert, J. (2013). How different mentoring approaches affect beginning teachers' development in the first years of practice. *Teaching and Teacher Education*, *36*, 166–177. https://doi.org/10.1016/j.tate.2013.07.012

Ronfeldt, M., & McQueen, K. (2017). Does new teacher induction really improve retention? *Journal of Teacher Education*, *68*(4), 394–410. https://doi.org/10.1177/0022487117702583

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and non-randomized studies. *Journal of Educational Psychology*, *66*(5), 688–701. https://doi.org/10.1037/h0037350

Saloviita, T., & Pakarinen, E. (2021). Teacher burnout explained: Teacher-, student-, and organisation-level variables. *Teaching and Teacher Education*, *97*, Article 103221. https://doi.org/10.1016/j.tate.2020.103221

Sekhon, J. S. (2008). The neyman-rubin model of causal inference and estimation via matching methods. In *The Oxford handbook of political methodology* (Vol. 2, pp. 1–32).

Shen, J. (1997). Teacher retention and attrition in public schools: Evidence from SASS91. *The Journal of Educational Research*, *91*(2), 81–88. https://doi.org/10.1080/00220679709597525

Sims, S. (2017). *TALIS 2013: Working conditions, teacher job satisfaction and retention* (Statistical Working Paper). ERIC.

Skaalvik, E. M., & Skaalvik, S. (2015). Job satisfaction, stress and coping strategies in the teaching profession-what do teachers say? *International Education Studies*, *8*(3), 181–192. https://doi.org/10.5539/ies.v8n3p181

Small, C. (2020). *A comparison of public and private school teachers' job satisfaction when controlling for policy perspectives, individual, and workplace characteristics* [Doctoral thesis, Kennesaw State University].

Smethem, L. (2007). Retention and intention in teaching careers: Will the new generation stay? *Teachers and Teaching: Theory and Practice*, *13*(5), 465–480. https://doi.org/10.1080/13540600701561661

Sorensen, L. C., & Ladd, H. F. (2020). The hidden costs of teacher turnover. *AERA Open*, *6*(1), 1–24. https://doi.org/10.1177/2332858420905812

Sönmezer, M. G., & Eryaman, M. Y. (2008). A comparative analysis of job satisfaction levels of public and private school teachers. *Journal of Theory & Practice in Education (JTPE)*, *4*(2), 189–212. https://dergipark.org.tr/tr/download/article-file/63270

Spilt, J. L., Koomen, H. M., & Thijs, J. T. (2011). Teacher wellbeing: The importance of teacher–student relationships. *Educational Psychology Review*, *23*(4), 457–477. https://doi.org/10.1007/s10648-011-9170-y

Splawa-Neyman, J., Dabrowska, D. M., & Speed, T. (1990). On the application of probability theory to agricultural experiments. Essay on principles. Section 9. *Statistical Science*, *5*(4), 465–472. https://doi.org/10.1214/ss/1177012031

Stekhoven, D. J., & Bühlmann, P. (2012). Missforest—non-parametric missing value imputation for mixed-type data. *Bioinformatics*, *28*(1), 112–118. https://doi.org/10.1093/bioinformatics/btr597

Suk, Y., Kim, J.-S., & Kang, H. (2021). Hybridizing machine learning methods and finite mixture models for estimating heterogeneous treatment effects in latent classes. *Journal of Educational and Behavioral Statistics*, *46*(3), 323–347. https://doi.org/10.3102/1076998620951983

Toropova, A., Myrberg, E., & Johansson, S. (2021). Teacher job satisfaction: The importance of school working conditions and teacher characteristics. *Educational Review*, *73*(1), 71–97. https://doi.org/10.1080/00131911.2019.1705247

UNESCO. (2015). *The challenge of teacher shortage and quality: Have we succeeded in getting enough quality teachers into classrooms?*

Wang, H., Hall, N. C., & Rahimi, S. (2015). Self-efficacy and causal attributions in teachers: Effects on burnout, job satisfaction, illness, and quitting intentions. *Teaching and Teacher Education*, *47*, 120–130. https://doi.org/10.1016/j.tate.2014.12.005

Wang, K., Li, Y., Luo, W., & Zhang, S. (2020). Selected factors contributing to teacher job satisfaction: A quantitative investigation using 2013 TALIS data. *Leadership and Policy in Schools*, *19*(3), 512–532. https://doi.org/10.1080/15700763.2019.1586963

Wiggan, G., Smith, D., & Watson-Vandiver, M. J. (2021). The national teacher shortage, urban education and the cognitive sociology of labor. *The Urban Review*, *53*(1), 43–75. https://doi.org/10.1007/s11256-020-00565-z

Woessmann, L., & West, M. (2006). Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS. *European Economic Review*, *50*(3), 695–736. https://doi.org/10.1016/j.euroecorev.2004.11.005

Yoon, I., & Kim, M. (2022). Dynamic patterns of teachers' professional development participation and their relations with socio-demographic characteristics, teacher self-efficacy, and job satisfaction. *Teaching and Teacher Education*, *109*, Article 103565. https://doi.org/10.1016/j.tate.2021.103565

Zhang, S., Shi, Q., & Lin, E. (2020). Professional development needs, support, and barriers: TALIS US new and veteran teachers' perspectives. *Professional Development in Education*, *46*(3), 440–453. https://doi.org/10.1080/19415257.2019.1614967

Appendices

Appendix 1. Questions used to define treatment groups

Treatment	Question	Condition
CPD	During the last 12 months, did you participate in any of the following professional development activities?	Teachers who responded "yes" to any 4 of the 10 available options.
Induction	Did you take part in any induction activities?	Teachers who responded "yes" to either taking part in a formal or informal induction programme at their current school.
Observing	On average, how often do you do the following in this school?	Teachers who did not respond "never" to the option "Observe other teachers' classes and provide feedback."
Team Teaching	On average, how often do you do the following in this school?	Teachers who did not respond "never" to the option "Teach jointly as a team in the same class."
Has Mentor	Are you currently involved in any mentoring activities as part of a formal arrangement at this school?	Teachers who responded "yes" to having a mentor.
Is Mentor	Are you currently involved in any mentoring activities as part of a formal arrangement at this school?	Teachers who responded "yes" to being a mentor.
Public	Is this school publicly or privately managed?	Teachers with a principal who indicated their school is publicly managed.
30+ Students	How many students are currently enrolled in this class?	Teachers who answered 30 or more students.
Out-of-field	Were the following subject categories included in your formal education or training, and do you teach them during the current school year to any students in this school?	Teachers who indicated that at least one option given was not included in their education, but that they do currently teach it.
Part-Time	What is your current employment status as a teacher, in terms of working hours?	Teachers who indicated they do not have a full-time contract at their current school.

Appendix 2. Definitions of key terms given in TALIS questionnaire

Key term	Definition
CPD	In this section, "professional development" is defined as activities that aim to develop an individual's skills, knowledge, expertise and other characteristics as a teacher.
Induction	"Induction activities" are designed to support new teachers' introduction into the teaching profession and to support experienced teachers who are new to a school, and they are either organised in formal, structured programmes or informally arranged as separate activities.
Mentoring	"Mentoring" is defined as a support structure in schools where more experienced teachers support less experienced teachers. This structure might involve all teachers in the school or only new teachers. It does not include mentoring of student teachers doing teaching practice at this school.
Public School	This is a school managed by a public education authority, government agency, municipality or governing board appointed by government or elected by public franchise.
Private School	This is a school managed by a non-government organisation; e.g. a church, trade union, business or other private institution.

Appendix 3. List of potential confounders used

TALIS Variable Code	Description	Removed from X for treatment
IDCNTPOP	Primary/secondary school.	
TT3G01	Gender.	
TT3G03	Highest level of formal education completed.	
TT3G04	How did you receive your first teaching qualification?	
TT3G05	Year of qualification.	
TT3G08	Was teaching your first choice as a career?	
TT3G09	Permanent/fixed-term contract.	
TT3G10A	Working hours at this school.	Part-Time Contract.
TT3G10B	Working hours altogether.	Part-Time Contract.
TC3G12	School publicly/privately managed	Public School.
TT3G11A	Year(s) working as a teacher at this school.	
TT3G11B	Year(s) working as a teacher in total.	
TT3G11C	Year(s) working in other education roles.	
TT3G11D	Year(s) working in non-education roles.	
TT3G12	Do you currently work as a teacher at another school?	
TT3G14	Number of students in class with special needs.	
TT3G37	Subject taught.	
TT3G38	Number of students in class.	30+ Students.
TT3G39A	% of time spent on administrative tasks.	
TT3G39B	% of time spent keeping order in classroom.	
TT3G39C	% of time actually spent teaching.	
T3STBEH	Student behaviour stress.	
T3CLAIN	Clarity of instruction.	
T3CLASM	Classroom management.	
T3COGAC	Cognitive activation.	
T3COLES	Professional collaboration in lessons among teachers.	
T3EFFPD	Effective professional development.	
T3EXCH	Exchange and co-ordination among teachers.	
T3PDBAR	Professional development barriers.	
T3DISC	Teachers' perceived disciplinary climate.	
T3PERUT	Personal utility motivation to teach.	
T3PDIV	Needs for professional development for teaching for diversity.	
T3PDPED	Needs for professional development in subject matter and pedagogy.	
T3VALP	Perceptions of value and policy influence.	
T3SATAT	Satisfaction with target class autonomy.	
T3SECLS	Self-efficacy in classroom management.	
T3SEINS	Self-efficacy in instruction.	
T3SEENG	Self-efficacy in student engagement.	
T3SEFE	Self-related efficacy in multicultural classrooms.	
T3SOCUT	Social utility value.	
T3STAKE	Participation among stakeholders, teachers.	
T3TEAM	Team innovativeness.	
T3STUD	Teacher-student relations.	
T3WELS	Workplace well-being and stress.	
T3WLOAD	Workload stress.	
T3TPRA	Teaching practices, overall.	
T3COOP	Teacher co-operation.	
T3SELF	Teacher self-efficacy.	
T3DIVP	Diversity practices.	
T3JOBSA	Overall job satisfaction.	All.