



University of
Nottingham
UK | CHINA | MALAYSIA

Does difference help make a difference? Examining whether young trustees and female trustees affect charities' financial performance.

Economics Dissertation, ECON3001, 2020-21
Chris Hyland, 14328066

Abstract

Regulators, sector leaders, and campaign groups alike are increasingly urging UK charities to increase the age and gender diversity of their trustee boards. Claims proliferate that doing so will improve their performance and thus their impact. Prior empirical evidence for these claims is sparse and like much non-profit governance research it principally relies on surveys that are small, unrepresentative, and cross-sectional. For this paper, I construct a novel panel dataset of 10,696 large English and Welsh charitable companies from 2007-18. I do not find robust evidence of a link between trustee age or gender and charity financial performance, and I show that effects in line with previous research are not robust to the inclusion of charity fixed effects or the use of instrumental variables. Using a complementary propensity score matched difference-in-differences design, I find no statistically significant effect of recruiting a first trustee aged under 31 on total income, voluntary income, or the ratio of charitable expenditure to total income. More broadly, my dataset construction approach may allow other charity governance questions to be re-addressed with panel data methods.

Acknowledgements

I am grateful to Seung-Keun Martinez, Juergen Amann, Thomas Chen and Markus Eberhardt for helpful discussions and comments. I also thank Zainub Absar for proofreading in line with university policies.

Word count: 7,500

Supervisor: Seung-Keun Martinez

This Dissertation is presented in part fulfilment of the requirement for the completion of an undergraduate degree in the School of Economics, University of Nottingham. The work is the sole responsibility of the candidate. I do give permission for my dissertation to be made available to students in future years if selected as an example of good practice.

Table of contents

1. INTRODUCTION	3
2. LITERATURE REVIEW.....	4
2.1 THEORETICAL FRAMEWORKS.....	4
2.2 EMPIRICAL EVIDENCE: FOR-PROFIT FIRMS	4
2.3 EMPIRICAL EVIDENCE: NON-PROFITS	5
3. DATA AND VARIABLES	9
3.1 DATA.....	9
3.2 VARIABLE SELECTION.....	9
3.3 DESCRIPTIVE STATISTICS.....	10
4. METHODOLOGY.....	13
4.1 CORE MODEL.....	13
4.2 PROPENSITY SCORE MATCHED DIFFERENCE-IN-DIFFERENCES.....	15
5. RESULTS	16
5.1 CORE MODEL.....	16
5.2 PROPENSITY SCORE MATCHED DIFFERENCE-IN-DIFFERENCES.....	23
5.3 ROBUSTNESS AND EXTENSIONS	28
6. DISCUSSION.....	30
7. CONCLUSION	31
8. REFERENCES	33
9. APPENDIX.....	38

1. Introduction

20% of the adult population in England and Wales are aged 18-29 (ONS, 2019). However, only 2% of the trustees who govern and direct their 168,000 registered charities, usually as volunteers, fall in the same bracket (Charity Commission, 2010, 2018a). The Charity Commission (2010, p. 4) actively wishes “to encourage young trusteeship, and to help to create more diverse boards”, and the Charities Aid Foundation (2015, p. 3) describes the average age of trustees being 57 as “extremely worrying”. A ‘Young Trustees Movement’ (2019) aims to double the number of trustees aged under 31 by 2024.

These organisations claim young trustees strengthen decision-making by bringing more diverse experiences, increase public support by visibly increasing representation, and contribute insights and technological capabilities to engage broader donor and beneficiary groups (Charity Commission, 2010; Charities Aid Foundation, 2015). To the best of my knowledge, however, no prior empirical evidence from the UK, and none internationally attempting causal identification, tests these claims of a link between younger trustees and charity performance, on which official recommendations are being based. My dissertation fills this gap, focusing on financial performance. In a sector whose workforce is 65% female, only 34% of the trustees of large charities are women (Inclusive Boards, 2018). I therefore also examine links between female trustees and financial performance.

By studying English and Welsh charitable *companies*, the structure adopted by most large charities, I construct a novel panel dataset of 10,696 charities from 2007-18 that is substantially larger and richer than previously used in this literature. Using a pooled OLS specification I can replicate results from prior research, including findings that age-diverse non-profit boards are associated with greater donations, and gender-diverse with lower (Siciliano, 1996). These effects disappear when including charity fixed effects, and when instrumenting for age and gender to mitigate endogeneity. Neither have been done before in this literature, as far as I am aware. I instrument for female representation using the proportion of male trustees sitting on other boards with a female director, and for age with the proportion of trustees aged 31 or over sitting on other boards with a trustee aged under 31. I find no robust, statistically significant effect of the proportion of female trustees, the proportion of trustees aged under 31, the standard deviation of age or mean age on charities’ total income, voluntary income, or the ratio of charitable expenditure to total income.

Using a complementary propensity score matched difference-in-differences design, I compare the performance of charities which recruit their first ever trustee aged under 31 with that of those that have never had a young trustee. Across a variety of matching procedures, I find no consistent statistically significant effect of first young trustees on the same financial performance measures.

My research thus suggests that some assumptions underlying charity trustee diversity recommendations should be re-evaluated. It also demonstrates the importance of using panel data in the non-profit governance literature for more credibly causal identification and establishes an accessible and replicable means of obtaining it.

The paper proceeds as follows. *Section 2* surveys prior theoretical and empirical literature. *Section 3* details the data sources and key variables. *Section 4* introduces both central empirical methodologies, and *Section 5* their key results with an array of robustness checks. *Section 6* discusses these results and *Section 7* concludes.

2. Literature review

2.1 Theoretical frameworks

As directors of profit-making firms and trustees of non-profits alike are often older and male, theoretical approaches to the effect of young and female board members on performance usually concern age and gender *diversity*.

Carter et al. (2010) argue that the strongest theoretical link comes from resource dependence theory (Pfeffer & Salancik, 1978), which suggests boards facilitate connections with external bodies holding crucial resources. Non-profit literature calls this “boundary spanning”, with fundraising a common example (Callen et al., 2009). Hillman et al. (2002), suggest a taxonomy of board member types that provide different resources. This implies the diverse information sets and connections of younger or female trustees can source more valuable resources from more sources, improving financial performance (Carter et al., 2010). Increased diversity itself may also raise organisational legitimacy (Hillman et al., 2007).

Other theories are less equivocal. Younger trustees are likely to have less training and experience and so less human capital (Borjas, 2016), but might contribute unique forms of human capital by being more agile and energetic (Mishra & Jhunjhunwala, 2013). Social psychology theories suggest board diversity could both encourage innovation and slow decision-making (Carter et al., 2010).

Overall, theory inconclusively predicts the effects of young and female trustees on non-profits’ financial performance. Notably, Hermalin and Weisbach (2005) suggest that if organisations are in equilibrium, rationally optimising their boards, empirical relationships between board characteristics and performance are spurious. The empirical literature rarely addresses this. I propose that evidence that young people often lack accurate information about trusteeships, and that some older trustees resist younger members (Charity Commission, 2010), may imply current disequilibrium. I return to this when discussing my results.

2.2 Empirical evidence: for-profit firms

Empirical non-profit literature is “often a step behind” (Garcia-Rodriguez et al., 2020, p. 2) that studying firms, and the effect of age and gender diversity on non-profits’ performance is especially underexplored (Stone & Ostrower, 2007). I will therefore first summarise relevant firm-related literature.

Plentiful research tests the relationship between female directors and firm performance but, as *Table 1* demonstrates, results are highly inconsistent. Post & Byron’s (2015) meta-analysis of 140 studies finds a positive relationship on average between female directors and accounting-based performance measures like return on assets (ROA), but none with market-based assessments.

Endogeneity likely partly explains conflicting results. Both omitted variable bias, perhaps due to firm culture affecting gender diversity and performance, and simultaneity bias, as women may be more likely to apply to and/or be selected by high-performing firms (Smith et al., 2006), are likely. Various mitigation strategies have been implemented. Adams and Ferreira (2009) instrument for female directors using the proportion of male directors sitting on another board with female representation. They argue that this is a proxy for informal social connections that help women into power without independently affecting performance. Studying 1,939 US firms, they show that a positive female-performance link from pooled OLS regression disappears when using fixed effects, instrumental variable, or system GMM estimation. Natural experiments,

such as the Norwegian gender quotas exploited by Ahern and Dittmar (2012), provide the most credibly causal evidence but are rare.

A much smaller literature examines director ages and firm performance. As *Table 2* illustrates, these are also inconclusive. Talavera et al. (2018), studying 97 Chinese banks, is perhaps the most plausibly causal. Using fixed-effects estimation, instruments of the age diversity in surrounding provinces and similarly sized banks, and two-step system GMM estimation, they find a negative relationship between the coefficient of variation of director age and both ROA and return on equity.

2.3 Empirical evidence: non-profits

Table 3 outlines the much scarcer (Stone & Ostrower, 2007) non-profit literature. Causal identification is rarely addressed and, as per the broader non-profit governance literature, papers typically rely on cross-sectional surveys of particular regions or organisation types (Cornforth, 2012).

Siciliano (1996) investigates associations between board diversity and performance using a survey of 240 US YMCA non-profits. Partial correlations, controlling for revenues, show age diversity is associated with higher donations. However, it has no significant correlation with consultant-assessed social performance or fiscal performance (revenue ÷ expenses). Gender diversity is related to lower donations and greater social performance but not to fiscal performance. Although an influential early contribution, the single control makes omitted variable bias probable. However, to the best of my knowledge, it remains the *only* research linking board member age to non-profits' overall performance, except Perkins and Fields' (2010) study of 82 churches which includes no organisation-level controls and likewise is cross-sectional.

Harris (2014), with a larger cross-sectional sample and many more controls, makes some progress regarding gender diversity. They find a positive effect on the student retention and enrolment growth of US universities, but none on other performance indicators. However, the author admits it cannot be interpreted causally.

The research most similar to mine is Elmagrhi et al. (2018), who manually transcribe annual reports from 2010-14 of 50 of the UK's largest charities. They primarily use pooled OLS estimation, but to reduce endogeneity they also try fixed effects estimation, two-stage least squares estimation (2SLS), and a Heckman selection model. When performing 2SLS they do not use an instrument, running a first stage merely on the controls, achieving little. Their two-stage Heckman model corrects for trustee self-selection using the proportion of women in each charity's industry to predict female trustee recruitment. The model assumes errors from each stage have a bivariate normal distribution (Bascle, 2008), so using this as an instrument in 2SLS may have been preferable. They find that female trustees, particularly up to a 'critical mass' of three, are associated with reduced short-term and total debt. They fail to find a direct effect of the proportion of female trustees on ROA, but only test this using pooled OLS regression. Further, ROA is not a financial performance measure charities commonly use (Boateng et al., 2016). Trustee ages are not considered, presumably as they are not published.

TABLE 1 — KEY EMPIRICAL RESEARCH OF THE LINK BETWEEN FEMALE BOARD MEMBERS AND FOR-PROFIT FIRMS' PERFORMANCE

Author(s)	Year	Country	Time period	No. of boards	Econometrics	Female measure	Association with ROA	Association with Tobin's Q	Association with other
Ahern & Dittmar	2012	Norway	2001-2009	248	Event study/IV	%		-ive	Negative market response
Adams & Ferreira	2009	USA	1996-2003	1939	OLS/FE/IV/GMM	%	-ive	-ive	
Liu et al.	2014	China	1999-2011	≈2,000	OLS/FE/IV/GMM	%	+ive		
Pathan & Faff	2013	USA	1997-2011	212	OLS/FE/GMM	%	+ive	Zero	
Carter et al.	2010	USA	1998-2002	641	OLS/FE/3SLS	Number	Zero	Zero	
Conyon & He	2017	USA	2007-2014	3,634	OLS(Quantile)/IV	%/presence	+ive	+ive	
Farrell & Hersch	2005	USA	1990-1999	309	Event study	Addition			Zero market response
Campbell et al.	2007	Spain	1995-2000	68	OLS/FE	%/Blau/Shannon		+ive	
Fernández-Temprano & Tejerina-Gaite	2020	Spain	2005-2015	87	OLS/FE	%	Zero	Zero	
Shehata et al.	2017	UK	2005-2013	34,798	OLS/FE	%/presence/Blau/Shannon	-ive		
Darmadi	2011	Indonesia	2007	169	OLS	%/Blau/presence	-ive	-ive	
Post & Byron	2015	36 different	Meta	90,070	Meta	Various	+ive	Zero	
Carter et al.	2003	USA	1997	638	OLS	%/presence		+ive	
Bonn et al.	2004	Japan, Australia	1999	273	OLS	%	Zero		+ive MTB Australia
Ali et al.	2014	Australia	2012	288	OLS	Blau	Zero		
Mahadeo et al.	2012	Mauritius	2007	42	OLS	%	+ive		

Notes: Econometrics abbreviations: OLS = ordinary least squares regression; FE = fixed effects panel regression; IV = instrumental variables regression using two-stage least squares; 3SLS = three-stage least squares estimation; GMM = difference or system generalised method of moments estimation; meta = meta-analysis combining many studies. "Female measure" details the board gender diversity explanatory variables, where: % = percentage of females; presence = dummy variable indicating at least one female; number = the number of females; addition = the event of a female director being appointed. Blau and Shannon are diversity indices equal to $1 - \sum_{i=1}^K P_i^2$ and $-\sum_{i=1}^K P_i \ln(P_i)$, respectively, where K is the number of gender categories present in a firm and P_i the proportion of board members in each (Shehata et al., 2017). ROA = return on assets (net income ÷ total assets), Tobin's Q measures market value compared to asset replacement value, and MTB = market to book ratio (market capitalisation ÷ total book value).

TABLE 2 — KEY EMPIRICAL RESEARCH OF THE LINK BETWEEN BOARD MEMBER AGE AND FOR-PROFIT FIRMS' PERFORMANCE

Author(s)	Year	Country	Time period	No. of boards	Econometrics	Age measure	Association with ROA	Association with Tobin's Q	Association with other
<i>Panel A: Age diversity</i>									
Talavera et al.	2018	China	2009-2013	97	OLS/FE/IV/GMM	CV	-ive		
Shehata et al.	2017	UK	2005-2013	34,798	OLS/FE	CV	-ive		
Fernández-Temprano & Tejerina-Gaite	2020	Spain	2005-2015	87	OLS/FE	SD	+ive		+ive MTB
Kunze et al.	2011	Germany	2008	128	OLS/SEM	SD	-ive		
Mahadeo et al.	2012	Mauritius	2007	42	OLS	*	+ive		
Ali et al.	2014	Australia	2012	288	OLS	CV	-ive		
Ararat et al.	2015	Turkey	2006	95	OLS	Blau	+ive	Zero	
<i>Panel B: Average age</i>									
Carter et al.	2010	USA	1998-2002	641	OLS/FE/3SLS	Mean	Zero	Zero	
Faleye	2007	USA	1995-2002	2021	OLS/Fama-Macbeth	Mean		-ive	
Bonn et al.	2004	Japan, Australia	1999	273	OLS	Mean	Zero		-ive MTB Japan, zero Australia
<i>Panel C: Young age</i>									
Darmadi	2011	Indonesia	2007	169	OLS	% under 50	Zero	+ive	

Notes: Econometrics abbreviations: OLS = ordinary least squares regression; FE = fixed effects panel regression; IV = instrumental variables regression using two-stage least squares; 3SLS = three-stage least squares estimation; GMM = difference or system generalised method of moments estimation, SEM = structural equation modelling. "Age measure" details the board member age explanatory variable, where: SD = standard deviation of age; CV = coefficient of variation of age (standard deviation ÷ mean). Blau is a diversity index equal to $1 - \sum_{i=1}^K P_i^2$ where K is the number of age categories present in a firm, and P_i the proportion of board members in each (Shehata et al., 2017). ROA = return on assets (net income ÷ total assets), Tobin's Q measures market value compared to asset replacement value. * Mahadeo et al. (2012) use the number of age categories, of the six they specify, present in a board as their age explanatory variable.

TABLE 3 — KEY EMPIRICAL RESEARCH OF THE LINK BETWEEN TRUSTEE GENDER OR AGE DIVERSITY AND NON-PROFITS' PERFORMANCE

Author(s)	Year	Country	Non-profit type	Time period	No. of boards	Econometrics	Female measure	Age measure	Controls	Association of gender diversity	Association of age diversity
Elmagrhi et al.	2018	UK	Large charities	2010-2014	50	OLS/FE/IV/Heckman*	%/presence	N/A	Board size, audit firm size, presence of separate corporate governance committee, number of board meetings, liquidity, capital expenditure, industry dummies, year dummies	-ive: total/long-term/short-term debt to total assets ratios Zero: ROA	N/A
Harris	2014	USA	Colleges & universities	2008	554	OLS	%	N/A	% minority, % with financial expertise, % serve on other non-profit boards, % have worked in higher education, efficiency ratio, fundraising expenses, operating margin, organisation age, total assets, program service revenue, other revenue, dummy for conference sports programs, dummy for religious affiliation, dummy for survey type	+ive: student retention and enrolment growth; zero: total contributions, organisational rank, SAT scores	N/A
Perkins & Fields	2010	USA	Christian churches	Single year	82	OLS	SD	SD	Mean and standard deviation of tenure length, spiritual maturity, working relationships with senior pastor (LMX scale); mean age	Zero: Sunday school attendance, attendance growth, offering growth, ministry budget ratio	+ive: Sunday school attendance; -ive: attendance growth; zero: offering growth, ministry budget ratio
Buse et al.	2016	USA	Various	2012	1,456	SEM	Blau	Blau	Blau racial/ethnic diversity index, 8 measures of board diversity policies and practices, 8 measures of board inclusion behaviours	+ive: board diversity policies and practices, CEO-assessed internal and external governance practices	Zero: CEO assessments of internal and external governance practices; +ive: board diversity policies and practices
Siciliano	1996	USA	YMCA	1989	240	OLS	†	†	Revenue	-ive: donations; +ive: assessment mission fulfilment; zero: ratio revenue to expenses	+ive: donations; zero: assessment of mission fulfilment, ratio revenue to expenses
Garcia-Rodriguez et al.	2020	Spain	NGDOs	2011-2013	64	OLS	Blau	N/A	Board size, % insiders, duality, % founders, % board with experience in [variety of areas], mean educational level, total assets, organisation age, margin, revenue concentration, % public funding	Zero: multidimensional measure of financial vulnerability	N/A

Notes: Econometrics abbreviations: OLS = ordinary least squares regression; FE = fixed effects panel regression; IV = instrumental variables regression using two-stage least squares; Heckman = Heckman selection model; SEM = structural equation modelling. “Female measure” and “age measure” detail the trustee diversity explanatory variables, where: % = percentage of females; presence = dummy variable indicating at least one female; SD = standard deviation. Blau is a diversity index equal to $1 - \sum_{i=1}^K P_i^2$, where K is the number of gender or age categories present in a non-profit, and P_i the proportion of board members in each (Shehata et al., 2017). ROA = return on assets (net income ÷ total assets). * Elmagrhi et al. (2018) only use pooled OLS when testing the direct relationship between gender and ROA. † Siciliano (1996) use a custom age diversity measure which subtracts from 100 the highest % of members in any of the five age bands they specify on a board, then multiplies this by the total number of categories of age present. Gender diversity is calculated similarly.

3. Data and variables

3.1 Data

To overcome the data limitations typically constraining non-profit governance research, I identified that although the Charity Commission merely publishes current trustee names, Companies House publishes the names, titles, nationalities, and dates of birth of past and present *company directors*. Many charities, including 78% of the largest 4,000 by income (Charity Commission, 2020a), are charitable *companies*. For these, “trustees are the directors of the company” (Charity Commission, 2005a). I therefore match the charity register (Charity Commission, 2020a), which includes detailed finances since 2007 for charities with annual income over £500,000¹, with Companies House data. I have not encountered previous research doing this. This creates an unbalanced panel from 2007-18 of 10,696 charitable companies with annual income over £500,000: 79,417 charity-year observations.

I source some Companies House data from the Fame database (BVD, 2020). However, 1% of relevant birthdates are missing. Additionally, constructing my instrumental variables requires data on all directors at *other* companies that trustees in my sample are also on the boards of, which Fame cannot provide. I therefore built a Python wrapper for the Companies House (2020) API itself, based on that by Mistler (2018). My scripts ran for several days, due to rate-limiting and the hundreds of thousands of requests necessary. *Appendix 1* has full details.

I classify 85.9% of trustee genders using titles (Mr, Mrs, etc.), and the remainder based on first names using Genderize.io (Demografix ApS, 2020). Wais (2016, p. 36) finds Genderize.io to be “the best currently available approach.”

3.2 Variable selection

Given varying missions and strategies, “the non-profit sector has no accepted universal measure of success” (Crutchfield & Grant, 2012, p. 328). Boateng et al. (2016) investigate which performance measures 105 large British charities use in practice, finding charitable expenditure ÷ total income definitively the most common. Therefore, I use this ‘spending ratio’ as one financial performance measure. A greater ratio is generally perceived as favourable, though this can depend on a charity’s strategy (Sayer, 2018). Some charities secure lump-sum grants or endowments that are spent over several years, distorting this statistic, so I confirm my core results are insensitive to averaging it over 2, 3 or 4 years (*Appendix 2*).

I use total income, which Boateng et al. (2016) find is also widely used, as a second performance measure with a clearer interpretation. I adopt voluntary income as the third to allow comparisons with prior literature and because resource dependence theory perhaps implies reaching new donor groups is a particularly credible channel through which diverse boards might impact financial performance. I express both in natural logarithms due to their strong positive skew. 14.7% of voluntary income observations are zero, so I recode these as one before taking the logarithm. Just 12 raw observations have voluntary income of one, so this causes negligible distortion. My results are not sensitive to instead excluding observations with zero values

¹ From “part b” returns. I identified around fifty entry errors in this data while investigating outliers, which I corrected using annual reports or other Charity Commission records.

or using the inverse hyperbolic sine transformation which avoids recoding but is less readily interpreted (*Appendix 3*).

Following most of the literature (Darmadi, 2011), I use the proportion of trustees who are women to measure female presence. I test three age explanatory variables in turn. The proportion of trustees aged under 31 captures young trusteeship, standard deviation of age captures age diversity, and mean age the overall youth of a board. Selecting a 'youth' threshold is inevitably somewhat arbitrary. I choose under 31 to coincide with the Young Trustees Movement's (2020) targets.

I control for log board size, the proportion of non-British trustees, leverage, and log total assets as a charity size proxy. All are common in firm board diversity research (e.g. Adams & Ferreira, 2009; Campbell et al., 2007). Non-profit literature often includes a dummy for operating only in one area (e.g. McDonnell & Rutherford, 2019), whereas I have created a more granular set of dummies representing working in one local authority, throughout Greater London, UK-wide, or internationally. The Charity Commission does not report any unique classification per charity, so I create a mapping of company SIC codes to the NCVO (2020) variant of the ICNPO non-profit classifications to control for 16 charity types (*Appendix 4*). The NCVO have a dataset classifying every charity more rigorously than they are willing to share but could not compile in time for this paper.

3.3 Descriptive statistics

Table 4 summarises all variables employed, and *Table 5* contains the means of key variables by charity classification. Scout groups and youth clubs have the greatest proportion of young trustees, with a mean proportion of trustees aged under 31 of 7% against a 2% sample-wide mean. *Figure 1* shows the age profile of all trustees active in 2018. *Figure 2* visualises how the four explanatory variables have evolved 2007-18, suggesting young trustees and female trustees are increasingly common, though mean age has risen.

TABLE 4 — SUMMARY OF VARIABLES

Variable	Description	Source	Mean	Standard deviation	Max.	Min.
<i>Panel A: Dependent variables</i>						
Ln(Total income)	Natural logarithm of total annual income	Charity Register	14.53	1.13	20.78	13.12
Ln(Voluntary income)	Natural logarithm of total voluntary income. Voluntary income is “all income... that is, in substance, a gift made to it on a voluntary basis.” (Charity Commission, 2014) Zero values are recoded to one before transformation.	Charity Register	10.29	4.85	20.69	0.00
Spending ratio	Charitable expenditure ÷ total income. Charitable expenditure is “costs incurred by the charity in supplying goods or services to meet the needs of its beneficiaries” or “grants made to meet the needs of the charity’s beneficiaries.” (Charity Commission, 2020b)	Charity Register	0.85	0.37	53.75	0.00
<i>Panel B: Explanatory variables</i>						
Trustees <31	Proportion of active trustees aged under 31	Companies House API	0.02	0.07	1.00	0.00
SD trustee age	Standard deviation of age of active trustees. Set to zero if single trustee.	Companies House API	9.68	3.56	41.01	0.00
Mean trustee age	Mean age of active trustees	Companies House API	56.65	6.93	88.00	18.00
Proportion female	Proportion of active trustees who are female	Companies House API & Genderize.io	0.35	0.22	1.00	0.00
<i>Panel C: Board-level control variables</i>						
Foreign	Proportion of active trustees without exclusively British nationality	Companies House API	0.09	0.17	1.00	0.00
Ln(Board size)	Natural logarithm of the total active trustees	Companies House API	2.07	0.50	4.49	0.00
<i>Panel D: Charity-level control variables</i>						
Ln(Total assets)	Natural logarithm of total assets	Fame	14.57	1.62	22.21	6.04
Leverage	Long-term debt ÷ total assets	Fame	0.05	0.12	1.00	0.00
Locale	Set of dummy variables representing whether a charity’s largest scale of operation is: local, GLA (at least 10 authorities in the Greater London Area), UK-wide (at least 10 local authorities in England and/or Wales, or operating in Scotland or Northern Ireland), or international	Charity Register				
Classification	Set of dummy variables representing 16 charity classifications from the NCVO-adapted ICNPO (International Classification of Nonprofit Organisations) (NCVO, 2020). Based on custom mapping from SIC codes, see <i>Appendix 4</i> .	Fame				

Notes: Board variables are calculated on the final day of each charity’s financial year

TABLE 5 — DEPENDENT AND EXPLANATORY VARIABLE MEANS BY CHARITY CLASSIFICATION

Classification	Observations	Mean: Ln(Total income)	Mean: Ln(Voluntary income)	Mean: Spending ratio	Mean: Proportion < 31	Mean: SD age	Mean: Mean age	Mean: Proportion female
1. Culture and recreation	10,338	14.43	11.03	0.81	0.02	9.96	56.44	0.32
2. Education	18,090	14.81	9.01	0.89	0.02	9.12	55.81	0.34
2.2. Playgroups and nurseries	1,129	14.27	8.33	0.89	0.02	10.00	52.79	0.53
2.3. Research	1,913	14.79	9.40	0.85	0.01	9.03	57.00	0.29
3. Health	12,016	14.73	11.01	0.83	0.01	9.63	58.69	0.39
4. Social Services	22,851	14.33	10.57	0.85	0.02	10.04	57.03	0.38
4.1. Scout groups and youth clubs	78	14.74	8.43	0.93	0.07	13.20	54.52	0.33
5. Environment	978	14.36	10.88	0.80	0.01	9.78	57.77	0.30
6.1. Development	1,539	14.33	8.79	0.82	0.01	10.07	58.50	0.28
6.2. Housing	367	16.03	3.47	0.86	0.03	10.81	57.03	0.39
6.3. Employment and training	480	14.35	6.88	0.89	0.03	9.77	54.81	0.37
7. Law and advocacy	365	14.08	9.90	0.92	0.04	11.63	52.75	0.44
8.1. Grant-making foundations	260	14.15	11.28	0.75	0.01	10.20	58.86	0.35
8.2. Umbrella bodies	4,704	14.28	9.66	0.83	0.03	9.75	55.02	0.36
9. International	189	15.04	13.46	0.86	0.02	10.48	56.29	0.31
10. Religion	4,120	14.40	13.26	0.82	0.02	9.21	55.03	0.22
All	79,417	14.53	10.29	0.85	0.02	9.68	56.65	0.35

FIGURE 1 — HISTOGRAM OF TRUSTEE AGES, 2018

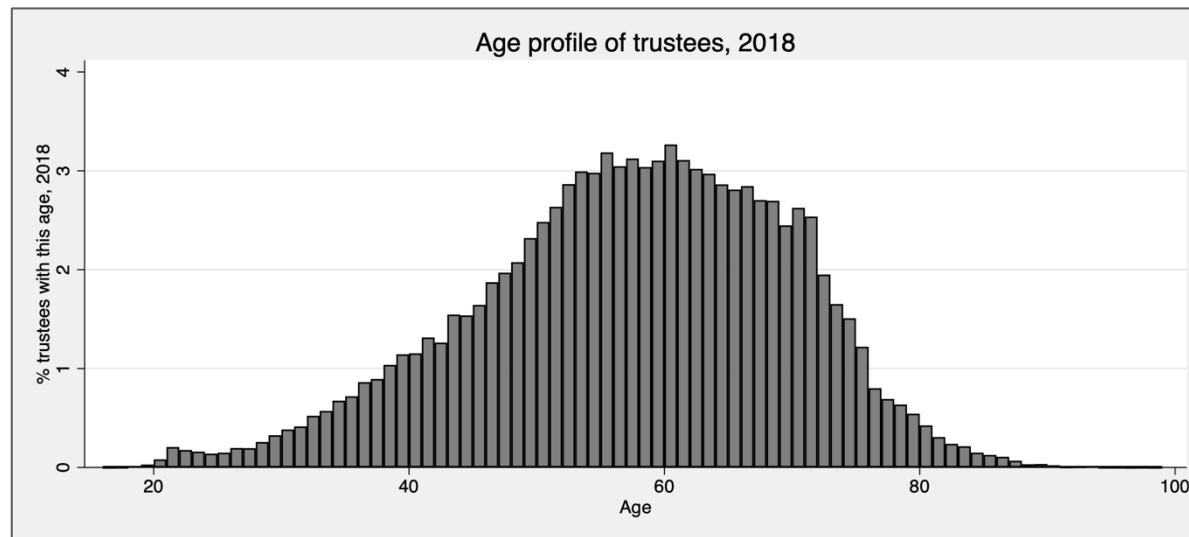
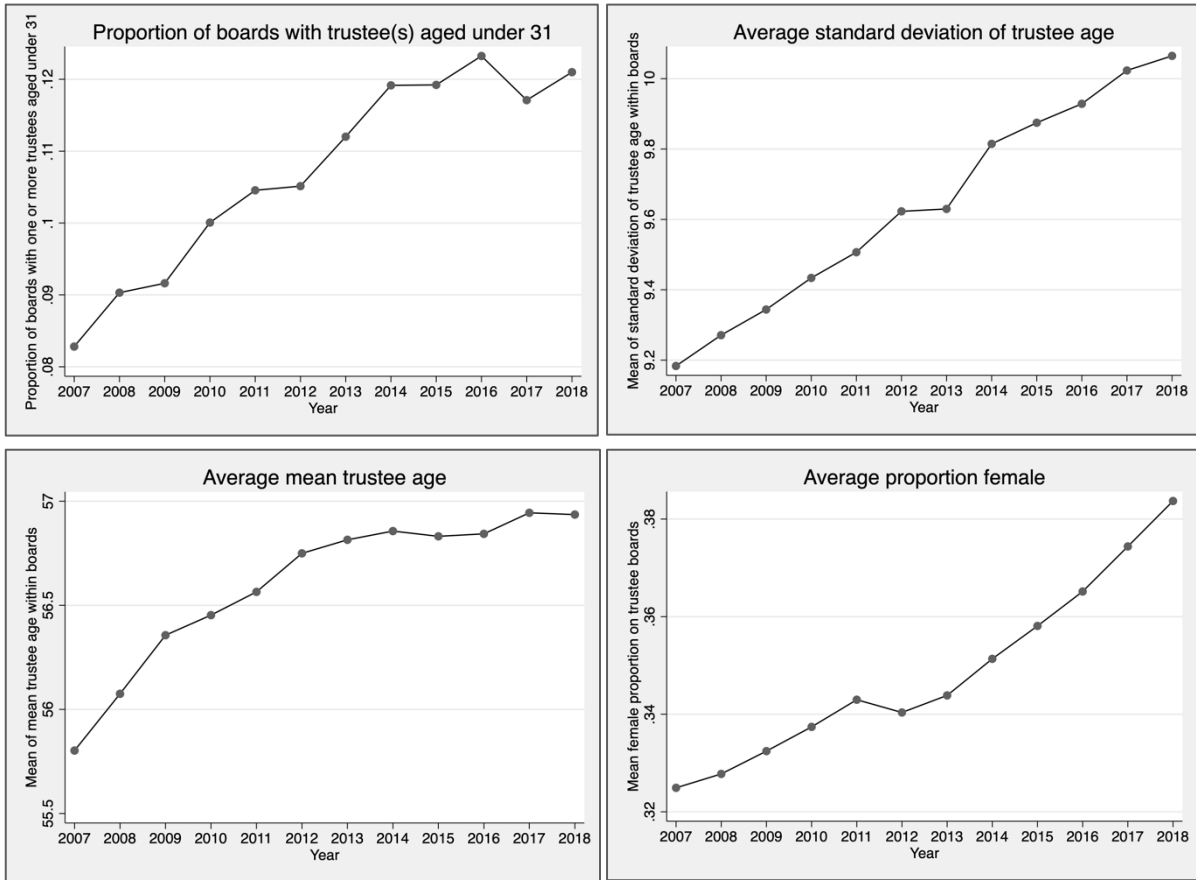


FIGURE 2 — TRUSTEE AGE AND GENDER TRENDS IN SAMPLED CHARITIES, 2007-18



4. Methodology

4.1 Core model

My core empirical specification is thus:

$$FinPerf_{it} = \beta_0 + \beta_1 AgeExplan_{i(t-1)} + \beta_2 PropFem_{i(t-1)} + \gamma X_{i(t-1)} + \lambda locale_i + \omega class_i + \tau_t + a_i + u_{it}$$

I test each of the three financial performance measures (*FinPerf*) with each of the three age-related explanatory variables (*AgeExplan*) in turn due to inevitable high multicollinearity, as per *Table 6*. The resulting nine combinations mean I must interpret statistical significance in the context of the multiple testing problem.

TABLE 6 — CORRELATION MATRIX OF AGE EXPLANATORY VARIABLES

	Proportion aged under 31	Standard deviation of age	Mean age
Proportion aged under 31	1		
Standard deviation of age	0.255	1	
Mean age	-0.442	-0.090	1

X is a vector of controls, while *locale* and *class* are sets of geographic scale and ICNPO classification dummies, excluding base groups. τ_t are year fixed effects, a_i charity fixed effects, and u_{it} idiosyncratic errors. In line with much existing literature (e.g. Carter et al., 2010; Fernández-Temprano & Tejerina-Gaite, 2020), all

independent variables are lagged by one year. Because they are defined at the end of each charity's financial year this is intuitively reasonable: when lagged they represent the situation at the *start* of the year over which financial performance is measured. Arguably (Reed, 2015), lagging also helps mitigate simultaneity bias (Liu et al., 2014). I show pooled OLS and random effects results but emphasise fixed effects estimation. This eliminates the confounding unobserved a_i , though also removes the time-invariant *locale* and *class* dummies.

As endogeneity likely remains, I instrument the proportion of female trustees using the proportion of male trustees sitting on another company board with at least one female director. Instruments are exogenous variables that affect the dependent variable solely ('exclusion restriction') through their effect on the endogenous regressor ('relevance'). This instrument was first proposed by Adams and Ferreira (2009) in the most cited paper in the board diversity literature, and it is now widely used² (e.g. Levi et al., 2014). They suggest it is a proxy for informal social connections that help women into power. Adams and Rangunathan (2017), adopting a similar instrument, argue these connections better enable male directors to identify appropriate female candidates, and improve inter-gender trust. These factors are likely *more* significant for charities, 81% of which recruit trustees by word of mouth and personal recommendation (Charity Commission, 2005b). Though reasonable justifications of instrument relevance, those Adams and Ferreira (2009) offer of exogeneity and the exclusion restriction are less concrete. They argue the instrument's correlation with an organisation's connectedness to other organisations is mitigated by organisation fixed effects and controlling for board size. They further show results are robust to controlling for total external board seats. I suggest that other, untestable, possibilities remain. Male board members' preponderance for tolerance and inclusion, as reflected in the gender balance on other boards on which they sit, could foster an organisational culture that directly affects financial performance, for example. Nevertheless, while its estimates should be treated cautiously, the instrument has sufficient plausibility to warrant inclusion. Similarly, I instrument the age explanatory variables using the proportion of trustees aged 31 or over who sit on another board with a member aged under 31. Experience working with younger directors may make them better able and more willing to recruit young trustees.

I explored other instruments. Following Liu et al. (2014) and Talavera et al. (2018) I instrumented using the mean of each gender and age characteristic in other charities of the same classification, in each year. Whether classifying using my ICNPO mapping or the underlying SIC codes, these had weak first stages. Others (e.g. Conyon & He, 2017) instrument using demographics of the areas that surround organisations. I implemented this using ONS (2020) population data, focussing on charities operating from one postcode. Given gender balance is often near parity and demographic change is gradual, these have weak first stages when instrumenting for gender or including charity fixed effects. Using *pooled OLS* estimation, proportion aged under 31 and mean age do have strong first stages, so for completeness these are reported in *Appendices 6-7*.

Given the difficulty of identifying truly exogenous instruments, the possibility of dynamic endogeneity, and the chance that control variables are also endogenous, Wintoki et al. (2012) encourage the use of dynamic panel GMM estimators in corporate governance research. I explored Arellano–Bover/Blundell–Bond

² Prior literature has not faced all-female boards, for which the instrument is undefined, but these are 1.9% of my sample. I set its value to 1 in these cases, but results are not sensitive to alternative specifications (*Appendix 5*). The same applies for instrumenting age, though only 48 observations have all-young boards.

(Roodman, 2009) two-step system GMM estimation, but could not find a specification that passed tests of overidentifying restrictions and appropriate serial correlation.

4.2 Propensity score matched difference-in-differences

I also conduct a propensity score matched difference-in-differences estimation of the average treatment effect (ATE) of a charity recruiting its first ever trustee aged under 31. This allows the identification of effects that may take years to materialise, as trustees build influence. It also mitigates any concerns that the ‘proportion aged under 31’ and ‘mean age’ variables in my core model change mechanically year-on-year as trustees age. In the vein of Gong and Girma (2020), I define a treatment group of charities ($T_i = 1$) which appoint their first trustee(s) aged under 31 during any year t in 2008-15 ($n=609$), and a control ($T_i = 0$) of those which at the end of 2018 had never had such a trustee ($n=2,838$). Time is defined relative to the treatment year, t , which corresponds to different calendar years for different charities³. Following Gong and Girma (2020), I set the reference year t for the control charities to be the median year each is in-sample. I label this the ‘median year’ approach.

The hypothetical ideal estimation, using the potential outcomes framework (Rubin, 1974), is:

$$ATE = E[Y_i(1) - Y_i(0)]$$

where $Y_i = FinPerf_{i(t+s)} - FinPerf_{it}$, and $Y_i(1)$ are *potential* outcomes if a charity did recruit a first young trustee, and $Y_i(0)$ if they did not. The ATE is the causal effect of recruiting a first young trustee on the change in financial performance between the treatment year and s years forward amongst all sampled charities that did not recruit a young trustee before 2008. I estimate for $s = 1, 2, 3$. It is not possible to determine the ATE directly, because both potential outcomes are never observed for a single charity. However, it can be estimated under the conditional mean independence assumption (CMIA), a weaker version of the conditional independence assumption (CIA). It states that mean potential outcomes are independent of treatment after conditioning on X_i , a vector of variables that determine selection into treatment. This is made more credible by the difference-in-differences design, where the outcome is the *change* in financial performance, which eliminates time-invariant differences in outcomes between treated and control charities (Smith & Todd, 2005). If the CMIA holds:

$$ATE = E[Y_i(1)|T_i = 1, X_i] - E[Y_i(0)|T_i = 0, X_i]$$

I choose X_i to comprise all independent and dependent variables from my core model in year $t - 1$, one year *before* treatment. Conditioning exactly on so many variables is not feasible, but Rosenbaum and Rubin (1983) demonstrate that adjusting for a scalar propensity score removes bias induced by all of them. The propensity score, $p(X_i) = Pr(T_i = 1|X_i)$, is the probability of being treated, conditional on X_i . Therefore:

$$ATE = E[Y_i(1)|T_i = 1, p(X_i)] - E[Y_i(0)|T_i = 0, p(X_i)]$$

I estimate propensity scores using a logit model. I then obtain ATE estimates using nearest neighbour matching with one neighbour and replacement, which matches each charity with the one of opposite

³ For example, year $t-1$ refers to 2007 for a charity that appoints its first trustee aged under 31 in 2008, and 2013 for one that does so in 2014.

treatment status with the closest propensity score. Using a single neighbour reduces finite sample biases caused by poor matches (Abadie & Imbens, 2016), and performs well in simulations (Austin, 2010). Then:

$$\widehat{ATE} = \frac{1}{N} \sum_{i=1}^N (2T_i - 1) (Y_i - Y_j)$$

where j is the unit matched to i , and N is the total number of units (Abadie & Imbens, 2016).

For this to be unbiased, the overlap assumption must also be satisfied: propensity scores must be bounded away from zero and one (Grilli & Rampichini, 2011). The SUTVA should hold too, which states potential outcomes do not depend on other charities' treatment statuses.

To explore the sensitivity of my results to the matching estimator, I also use the augmented inverse-probability weighting (AIPW) approach. This combines inverse probability weighting to model the treatment probability, where observations are weighted by the inverse of their propensity score, with a regression model of the outcome (Glynn & Quinn, 2010). The ATE estimate is consistent if *either* model is correct.

The CMIA is perhaps more credible if, in the nearest-neighbour specification, treated charities are matched with control charities whose t is set to be the *same calendar year*. I label this the 'same year' approach.

A parallel methodology cannot be followed for estimating the effect of a first *female* trustee, because just 17 charities would comprise the treatment group and 122 the control.

5. Results

5.1 Core model

Tables 7-9 report results from the core regressions, estimated using pooled OLS, random effects, fixed effects, and instrumental variable methods.

The pooled OLS estimates are the closest parallel to the prior cross-sectional non-profit literature. These suggest that a 10% increase in the proportion of trustees aged under 31, nearly a one-person increase on a mean-sized board of eight, is associated with a 3.2% increase in total income and an implausibly large 505% increase in voluntary income, all else equal⁴. Both are statistically significant at the 1% level. The sign and significance of the latter is consistent with Siciliano (1996). They construct the dependent variable as total donations ÷ total revenue, so I confirm this result holds when I do the same (*Appendix 8*). Pooled OLS estimates similarly suggest that a younger mean age is statistically significantly associated with higher total and voluntary income. A 10% decrease in mean age is linked to a 0.07% and 0.21% increase in each, respectively. Conversely, an increase in standard deviation of age is statistically significantly associated with a *fall* in total income, though the effect is small and statistically insignificant for voluntary income. There is no statistically significant relationship between young trustees or mean age and the spending ratio, though age diversity is associated with a very small increase in the spending ratio, statistically significant at the 5% level.

⁴ These and subsequent interpretations are calculated using the exponential function as coefficients are frequently sufficiently large to invalidate the approximation of $\ln(1 + x) \approx x$.

Pooled OLS results for female trustees are highly inconsistent, as per Harris (2014). A 10% increase in the proportion of female trustees is linked to a 4.9-5.88% fall in voluntary income, a 0.78-0.79 percentage point increase in the spending ratio, and has no statistically significant relationship with total income. The negative relationship with voluntary income is again consistent with Siciliano (1996).

However, statistical tests confirm pooled OLS estimation is likely inappropriate. Under the assumption that if a charity-specific time-invariant unobservable (a_i) is present it is uncorrelated with all independent variables in all time periods, as random effects models require, pooled OLS estimation is still consistent (Wooldridge, 2018). However, the composite error term ($v_{it} = a_i + u_{it}$) becomes serially correlated so standard errors and inferences drawn from them are incorrect. The Breusch and Pagan (1980) Lagrange multiplier test rejects the null hypothesis that the variance of a_i equals zero in every specification, suggesting the random effects model should be preferred to pooled OLS⁵. A Hausman test comparing random effects and fixed effects models, the latter permitting correlation between a_i and independent variables, finds coefficient estimates are systematically different in all specifications. This thus provides indicative evidence that both the pooled OLS and random effects estimates are not consistent.

Using fixed effects estimation, the previously highly statistically significant effects of the proportion of trustees aged under 31 and female trustees on voluntary income, in line with Siciliano (1996), become statistically insignificant. Indeed, no trustee age or gender coefficients for voluntary income or the spending ratio retain statistical significance. Three statistically significant coefficients are apparent in the total income model. The previously positive association of young trustees with total income becomes negative and is only statistically significant at the 10% level. It does not survive even the mildest attempts to correct for multiple hypothesis testing. The corresponding Benjamini Krieger Yekutieli (2006) sharpened-q value which accounts for the false discovery rate, considering the nine age-related coefficients obtained by fixed effects estimation⁶, is 0.52. The magnitude and statistical significance of the coefficient on standard deviation of age shrinks and again is not robust to this modest multiple testing correction. The magnitude of the female proportion coefficients in the total income model are similar to the pooled OLS estimates, and the standard errors lower. In two of the age specifications they are only statistically significant at the 10% level, however, with one significant at 5%. It appears that the statistically significant pooled OLS estimates are largely driven by unobserved charity-level time-invariant factors, which the fixed effects model eliminates, though there is weak evidence of a positive association of female trustees with total income.

Table 10 reports the first stages estimated when employing the director connections instruments. That they act in the expected direction in all cases, and are statistically significant at the 0.01% level, supports their relevance. It is curious that the age instrument is associated with lower female representation and vice versa⁷. The Cragg-Donald Wald F statistic and the robust Kleibergen-Paap rk Wald F statistic (Clemens, 2013) exceed the relevant Stock-Yogo (2005) critical value for 10% maximal relative bias, except when mean age

⁵ Wooldridge (2018) argues this test demonstrates little, for reasons including that it effectively detects *any* serial correlation in the composite error term, caused by a_i or not.

⁶ Values obtained using a script by Anderson (2007). Choice of a particular correction for multiple testing, and how many tests to account for, is highly subjective, so I do not employ it routinely.

⁷ Young trustees are not more likely to be male, as this might imply. In 2018, 50.0% of trustees in the sample aged under 31 were female, compared to 38.5% of those aged 31 and above.

(and female proportion, together) are instrumented. *Appendix 9* reports these statistics. Therefore, given the bias 2SLS estimation can introduce with weak instruments, I do not report IV estimates for the mean age models. As the models are exactly identified, I cannot conduct a Sargan-Hansen test.

When utilising the instruments, still including charity fixed effects, the age explanatory variables that retained some statistical significance in the fixed effects specification for total income become statistically insignificant. This reaffirms their interpretation as artefacts of multiple testing. The proportion of female trustees is positively associated with total income, statistically significant at the 5% level, in the model with proportion aged under 31. However, in the model with standard deviation of age the female coefficient is statistically insignificant. This result is therefore highly sensitive to model specification, and overall fails to provide compelling evidence of a gender effect on total income. The 'trustee connections' instruments thus do not find evidence of a consistent effect of trustee age or gender on financial performance.

Briefly addressing the control variables, the signs and statistical significance of the charity-level controls are similar across models. Total assets are positively associated with all three performance measures, in line with Harris (2014), and leverage is negatively associated with voluntary income. It might be that borrowing reduces the necessity of securing donations in the following year or improves capacity to generate income through charitable or trading activities. Trustee-level control estimates are a little less consistent. Larger boards have a positive relationship with total and voluntary income, but a negative association with the spending ratio when charity fixed effects are included. Non-British trustees seemingly have no relationship with financial performance, aside from perhaps a weak negative association with voluntary income.

TABLE 7 — CORE MODEL: TOTAL INCOME

Dependent variable: Age explanatory variable: Model:	Ln(Total income)										
	Proportion < 31				Standard deviation age				Mean age		
	Pooled OLS	Random Effects	Fixed Effects	Connections IV	Pooled OLS	Random Effects	Fixed Effects	Connections IV	Pooled OLS	Random Effects	Fixed Effects
Independent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
<i>Age:</i>											
Proportion < 31	0.277*** (0.106)	-0.045 (0.041)	-0.074* (0.043)	-0.367 (0.445)							
Standard deviation age					-0.009*** (0.002)	-0.003*** (0.001)	-0.002** (0.001)	-0.027 (0.034)			
Mean age									-0.007*** (0.001)	0.001 (0.001)	0.001 (0.001)
<i>Gender:</i>											
Proportion female	0.030 (0.035)	0.012 (0.016)	0.031* (0.017)	0.388** (0.179)	0.046 (0.035)	0.013 (0.016)	0.031* (0.017)	0.268 (0.293)	0.008 (0.035)	0.013 (0.016)	0.034** (0.017)
<i>Trustee controls:</i>											
Proportion non-British	0.111** (0.054)	-0.003 (0.029)	-0.000 (0.033)	-0.014 (0.036)	0.108** (0.053)	-0.004 (0.029)	-0.001 (0.033)	-0.009 (0.039)	0.095* (0.054)	-0.003 (0.029)	0.000 (0.033)
Ln(Board size)	0.193*** (0.018)	0.113*** (0.009)	0.096*** (0.010)	0.086*** (0.015)	0.206*** (0.018)	0.119*** (0.009)	0.099*** (0.010)	0.136* (0.075)	0.194*** (0.018)	0.113*** (0.009)	0.097*** (0.010)
<i>Charity controls:</i>											
Ln(Total assets)	0.487*** (0.007)	0.301*** (0.008)	0.221*** (0.011)	0.221*** (0.011)	0.483*** (0.007)	0.300*** (0.008)	0.221*** (0.011)	0.220*** (0.011)	0.490*** (0.007)	0.301*** (0.008)	0.221*** (0.011)
Leverage	-0.059 (0.062)	-0.014 (0.031)	-0.005 (0.032)	0.003 (0.033)	-0.055 (0.061)	-0.015 (0.031)	-0.005 (0.032)	-0.002 (0.033)	-0.077 (0.062)	-0.014 (0.031)	-0.005 (0.032)
Classification dummies	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No
Locale dummies	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No
Charity fixed effects	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes
Observations	66,277	66,277	66,277	65,442	66,277	66,277	66,277	65,442	66,277	66,277	66,277
Number of charities	8,865	8,865	8,865	8,030	8,865	8,865	8,865	8,030	8,865	8,865	8,865
R-squared	0.554	0.535	0.165	0.150	0.554	0.536	0.165	0.129	0.555	0.535	0.165

Notes: Heteroskedasticity-robust standard errors, adjusted for clustering at the charity level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All models include year dummies and a constant term. All independent variables are lagged by one year. The connections IV models, which also include charity-level fixed effects, use the proportion of male (aged 31 or over) trustees sitting on another board with a female (aged under 31) trustee as an instrument for 'proportion female' ('proportion < 31' and 'standard deviation age').

TABLE 8 — CORE MODEL: VOLUNTARY INCOME

Dependent variable: Age explanatory variable: Model:	Ln(Voluntary income)										
	Proportion < 31				Standard deviation age				Mean age		
	Pooled OLS (1)	Random Effects (2)	Fixed Effects (3)	Connections IV (4)	Pooled OLS (5)	Random Effects (6)	Fixed Effects (7)	Connections IV (8)	Pooled OLS (9)	Random Effects (10)	Fixed Effects (11)
<i>Age:</i>											
Proportion < 31	1.799*** (0.558)	0.568 (0.361)	0.194 (0.412)	4.847 (3.574)							
Standard deviation age					-0.002 (0.012)	0.000 (0.006)	0.001 (0.007)	0.355 (0.270)			
Mean age									-0.021*** (0.006)	-0.010** (0.004)	-0.005 (0.005)
<i>Gender:</i>											
Proportion female	-0.564*** (0.198)	-0.269** (0.133)	-0.122 (0.153)	-0.486 (1.504)	-0.512*** (0.198)	-0.259* (0.132)	-0.120 (0.152)	1.094 (2.431)	-0.606*** (0.199)	-0.296** (0.134)	-0.136 (0.154)
<i>Trustee controls:</i>											
Proportion non-British	0.406 (0.320)	-0.193 (0.201)	-0.415* (0.250)	-0.451* (0.271)	0.426 (0.320)	-0.186 (0.201)	-0.413* (0.250)	-0.513 (0.324)	0.369 (0.320)	-0.205 (0.201)	-0.419* (0.250)
Ln(Board size)	1.185*** (0.106)	0.501*** (0.063)	0.386*** (0.073)	0.340*** (0.114)	1.191*** (0.108)	0.507*** (0.063)	0.387*** (0.074)	-0.330 (0.590)	1.188*** (0.106)	0.498*** (0.063)	0.382*** (0.074)
<i>Charity controls:</i>											
Ln(Total assets)	0.188*** (0.035)	0.170*** (0.029)	0.177*** (0.049)	0.192*** (0.051)	0.180*** (0.035)	0.168*** (0.029)	0.177*** (0.049)	0.208*** (0.059)	0.195*** (0.035)	0.174*** (0.029)	0.178*** (0.049)
Leverage	-2.280*** (0.379)	-0.812*** (0.210)	-0.520** (0.228)	-0.569** (0.232)	-2.270*** (0.379)	-0.807*** (0.210)	-0.518** (0.228)	-0.508** (0.238)	-2.332*** (0.379)	-0.820*** (0.210)	-0.522** (0.228)
Classification dummies	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No
Locale dummies	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No
Charity fixed effects	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes
Observations	66,277	66,277	66,277	65,442	66,277	66,277	66,277	65,442	66,277	66,277	66,277
Number of charities	8,865	8,865	8,865	8,030	8,865	8,865	8,865	8,030	8,865	8,865	8,865
R-squared	0.115	0.108	0.003	-0.003	0.114	0.108	0.003	-0.094	0.115	0.108	0.003

Notes: Heteroskedasticity-robust standard errors, adjusted for clustering at the charity level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All models include year dummies and a constant term. All independent variables are lagged by one year. The connections IV models, which also include charity-level fixed effects, use the proportion of male (aged 31 or over) trustees sitting on another board with a female (aged under 31) trustee as an instrument for 'proportion female' ('proportion < 31' and 'standard deviation age').

TABLE 9 — CORE MODEL: SPENDING RATIO

Dependent variable: Age explanatory variable: Model:	Spending ratio										
	Proportion < 31				Standard deviation age				Mean age		
Independent variable	Pooled OLS (1)	Random Effects (2)	Fixed Effects (3)	Connections IV (4)	Pooled OLS (5)	Random Effects (6)	Fixed Effects (7)	Connections IV (8)	Pooled OLS (9)	Random Effects (10)	Fixed Effects (11)
<i>Age:</i>											
Proportion < 31	0.019 (0.071)	-0.012 (0.036)	-0.002 (0.029)	0.641 (0.432)							
Standard deviation age					0.001** (0.001)	0.001** (0.001)	0.000 (0.001)	0.047 (0.034)			
Mean age									-0.000 (0.000)	0.000 (0.000)	0.001 (0.000)
<i>Gender:</i>											
Proportion female	0.076*** (0.011)	0.051*** (0.010)	-0.002 (0.016)	0.058 (0.123)	0.075*** (0.011)	0.050*** (0.010)	-0.002 (0.016)	0.267 (0.261)	0.076*** (0.011)	0.052*** (0.010)	0.001 (0.016)
<i>Trustee controls:</i>											
Proportion non-British	0.025 (0.021)	0.021 (0.016)	0.018 (0.021)	0.008 (0.022)	0.026 (0.021)	0.022 (0.016)	0.018 (0.021)	-0.000 (0.030)	0.025 (0.021)	0.022 (0.016)	0.019 (0.021)
Ln(Board size)	0.007 (0.007)	0.005 (0.005)	-0.012** (0.006)	-0.023* (0.012)	0.005 (0.007)	0.003 (0.005)	-0.013** (0.006)	-0.111 (0.074)	0.007 (0.007)	0.005 (0.005)	-0.011* (0.006)
<i>Charity controls:</i>											
Ln(Total assets)	0.001 (0.004)	0.009*** (0.003)	0.073*** (0.005)	0.075*** (0.005)	0.002 (0.004)	0.010*** (0.003)	0.073*** (0.005)	0.077*** (0.006)	0.001 (0.004)	0.009*** (0.003)	0.072*** (0.005)
Leverage	-0.049 (0.031)	-0.043 (0.035)	-0.029 (0.047)	-0.035 (0.047)	-0.050 (0.031)	-0.044 (0.035)	-0.029 (0.047)	-0.027 (0.048)	-0.050 (0.031)	-0.043 (0.035)	-0.029 (0.047)
Classification dummies	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No
Locale dummies	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No
Charity fixed effects	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes
Observations	66,277	66,277	66,277	65,442	66,277	66,277	66,277	65,442	66,277	66,277	66,277
Number of charities	8,865	8,865	8,865	8,030	8,865	8,865	8,865	8,030	8,865	8,865	8,865
R-squared	0.011	0.009	0.012	0.005	0.011	0.009	0.012	-0.095	0.011	0.009	0.012

Notes: Heteroskedasticity-robust standard errors, adjusted for clustering at the charity level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All models include year dummies and a constant term. All independent variables are lagged by one year. The connections IV models, which also include charity-level fixed effects, use the proportion of male (aged 31 or over) trustees sitting on another board with a female (aged under 31) trustee as an instrument for 'proportion female' ('proportion < 31' and 'standard deviation age').

TABLE 10 — CONNECTIONS INSTRUMENTS: FIRST STAGES

Dependent variable:	Proportion < 31	Standard deviation age	Mean age	Proportion female
Model:				
Independent variable	(1)	(2)	(3)	(4)
<i>Instruments:</i>				
% males with female connections	-0.0108*** (0.00175)	-0.437*** (0.0980)	1.008*** (0.134)	0.0648*** (0.00509)
% >= 31 with < 31 connections	0.0516*** (0.00576)	0.807*** (0.192)	-0.886*** (0.227)	-0.0227** (0.00773)
<i>Trustee controls:</i>				
Proportion non-British	0.0125* (0.00530)	0.116 (0.351)	-1.470*** (0.439)	0.0514*** (0.0141)
Ln(Board size)	0.0124*** (0.00152)	1.860*** (0.0999)	-1.480*** (0.140)	0.0443*** (0.00456)
<i>Charity controls:</i>				
Ln(Total assets)	-0.00312*** (0.000819)	-0.0769 (0.0435)	0.286*** (0.0613)	-0.00205 (0.00204)
Leverage	0.00984* (0.00419)	0.0454 (0.226)	-0.793* (0.343)	-0.0182 (0.0104)
Observations	65,442	65,442	65,442	65,442
Number of charities	8,030	8,030	8,030	8,030

Notes: Heteroskedasticity-robust standard errors, adjusted for clustering at the charity level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All models include a constant term and year dummies. All variables are lagged by one year.

5.2 Propensity score matched difference-in-differences

Table 11 shows logit regressions used to estimate propensity scores for the median year and same year samples. Having a lower mean trustee age or higher trustee age diversity in the prior year increases the probability of recruiting a first young trustee, as might be expected. Propensity increases with the proportion of female trustees and board size, and falls with the proportion non-British and total assets.

Tables 12 and *13* show standardised differences and variance ratios for the median year and same year samples, respectively, demonstrating how well matching improves covariate balance. Matching very successfully improves balance in the median year sample, with all standardised differences being under the common benchmark of 0.1 (Austin, 2009) when using nearest neighbour matching or AIPW. Only one variance ratio is outside the oft-cited range of 0.5-2 (Stuart, 2010). It performs nearly as well using the same year sample, reducing the number of covariates with unacceptable standardised differences from 13 to one. There are four variance ratios outside the ideal range, though two are only marginally. *Figure 3* further demonstrates good propensity score overlap between treated and control groups, implying sufficient common support for reliable ATE estimation.

Table 14 contains the ATE estimates on the change in financial performance between the treatment year and one, two, and three years forward. The coefficient signs are not consistent, and only two effects are statistically significant. Using the median year sample, there is a negative effect of recruiting a first trustee aged under 31 on voluntary income in the subsequent year when using nearest neighbour matching, and a negative effect on the spending ratio three years forward when using nearest neighbour matching or AIPW. Neither are robust to matching treatment and control charities using the same reference calendar year. Nevertheless, because the latter is robust to both matching methods using the median year sample, I conduct a sensitivity test of whether it is likely to be driven by selection on unobservables. I use the recently developed concept of conditional c-dependence (Masten & Poirier, 2018). This relaxes the CIA by supposing:

$$|P(T_i = 1|Y_l = y_l, X_i = x_i) - P(T_i = 1|X_i = x_i)| \leq c$$

for all x_i , $l=0$ and 1 , $c \in [0,1]$.

If the CIA holds, c is zero. *Figure 4* illustrates how the bounds of the ATE in question vary with c . The breakdown point, the value of c beyond which the bounds include zero, is just 0.006. To interpret this result, *Figure 4* includes plots that indicate the maximum change in propensity scores induced by excluding each covariate in turn. Masten & Poirier (2018) argue this leave-one-out analysis provides a benchmark of the level of c-dependence unobservables might induce. The breakdown point is substantially lower than the c-dependence levels implied by all these benchmarks, confirming this effect is not robust.

Overall, propensity score matched difference-in-difference estimation finds no robust effect of recruiting a first young trustee on a charity's financial performance, even after several years.

TABLE 11 — LOGIT REGRESSIONS OF TREATMENT STATUS

Sample: Pre-treatment covariate	Median year (1)	Same year (2)
Mean age	-0.061*** (0.009)	-0.052*** (0.009)
Standard deviation age	0.034** (0.017)	0.039** (0.016)
Proportion female	0.626*** (0.237)	0.808*** (0.227)
Proportion non-British	-0.772** (0.345)	-0.562* (0.321)
Ln(Board size)	0.324*** (0.115)	0.332*** (0.109)
Ln(Total assets)	-0.228*** (0.047)	-0.187*** (0.044)
Leverage	-0.599 (0.406)	-0.574 (0.387)
Ln(Total income)	0.031 (0.068)	-0.008 (0.066)
Ln(Voluntary income)	-0.008 (0.010)	-0.006 (0.010)
Spending ratio	0.070 (0.153)	0.194* (0.114)
<i>ICNPO classification:</i>		
2. Education	-0.602*** (0.157)	-0.690*** (0.160)
2.2. Playgroups and nurseries	-0.399 (0.493)	-0.539 (0.475)
2.3. Research	-0.541 (0.338)	-0.636* (0.336)
3. Health	-0.399** (0.171)	-0.502*** (0.169)
4. Social Services	-0.030 (0.145)	-0.077 (0.144)
4.1. Scout groups and youth clubs	0.770 (1.554)	0.267 (1.411)
5. Environment	-1.771** (0.743)	-1.761** (0.738)
6.1. Development	-1.044** (0.450)	-1.056** (0.451)
6.2. Housing	0.679 (0.652)	0.545 (0.717)
6.3. Employment and training	0.036 (0.759)	-0.251 (0.695)
7. Law and advocacy	1.403 (1.065)	1.779* (0.914)
8.1. Grant-making foundations	-0.324 (0.831)	-0.497 (0.818)
8.2. Umbrella bodies	-0.245 (0.242)	-0.325 (0.239)
9. International	0.866 (0.624)	0.791 (0.663)
10. Religion	-0.183 (0.262)	-0.201 (0.256)
<i>Locale:</i>		
Greater London Area	-0.233 (0.247)	-0.168 (0.248)
UK-wide	-0.075 (0.114)	-0.020 (0.115)
International	-0.395** (0.165)	-0.362** (0.165)
Observations	3,447	20,784
Log likelihood	-1,507.717	-2,638.896

Notes: Heteroskedasticity-robust standard errors, adjusted for clustering at the charity level, in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Models include a constant. Dependent variable is a dummy equalling 1 if the charity recruits its first trustee aged under 31 in the following year, and 0 if the charity has never had a trustee aged under 31. Explanatory variables are measured the year before treatment. Omitted base ICNPO classification is "1. Culture and recreation", and omitted locale is "local". Median year sample includes control charities once using the median calendar year they are in-sample as their reference (t) year, same year includes control charities using every calendar year 2008-2015 in turn.

TABLE 12 — PRE-TREATMENT COVARIATE BALANCE STATISTICS: MEDIAN YEAR SAMPLE

Statistic: Sample:	Unmatched means		Standardised differences			Variance ratio		
	Treated	Controls	Raw	Matched: NN	Matched: AIPW	Raw	Matched: NN	Matched: AIPW
Pre-treatment covariate	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean age	57.135	58.934	-0.309†	0.022	0.086	1.048	1.106	1.236
Standard deviation age	9.065	8.615	0.151†	-0.003	-0.029	1.064	1.101	1.328
Proportion female	0.348	0.307	0.200†	-0.021	-0.013	1.099	1.104	1.050
Proportion non-British	0.066	0.089	-0.139†	-0.048	0.039	0.631	0.890	1.192
Ln(Board size)	2.107	2.076	0.066	-0.016	-0.065	0.817	0.988	1.226
Ln(Total Assets)	14.424	14.967	-0.359†	0.005	0.001	0.892	0.948	0.957
Leverage	0.042	0.055	-0.108†	0.046	0.026	0.635	0.939	0.966
Ln(Total income)	14.495	14.758	-0.243†	-0.008	-0.037	0.844	1.047	1.039
Ln(Voluntary income)	10.089	10.273	-0.037	-0.026	0.012	0.996	1.108	1.009
Spending ratio	0.856	0.845	0.040	0.001	-0.016	0.614	0.566	0.713
<i>ICNPO classification:</i>								
1. Culture and recreation	0.164	0.114	0.145†	-0.001	-0.021	1.359	0.998	0.952
2. Education	0.222	0.282	-0.138†	0.001	0.025	0.854	1.001	1.026
2.2. Playgroups and nurseries	0.011	0.009	0.027	-0.006	0.029	1.303	0.938	1.312
2.3. Research	0.020	0.030	-0.064	-0.020	-0.018	0.673	0.886	0.897
3. Health	0.144	0.179	-0.093	0.015	-0.018	0.844	1.027	0.969
4. Social Services	0.305	0.243	0.140†	-0.019	-0.021	1.154	0.978	0.976
4.1. Scout groups and youth clubs	0.002	0.000	0.041	0.000	-0.001	4.660†	1.000	0.973
5. Environment	0.003	0.016	-0.127†	-0.055	0.032	0.215†	0.595	1.293
6.1. Development	0.010	0.021	-0.094	-0.002	-0.009	0.464†	0.985	0.940
6.2. Housing	0.005	0.003	0.028	-0.005	-0.017	1.553	0.917	0.724
6.3. Employment and training	0.005	0.003	0.028	0.043	0.004	1.553	1.780	1.062
7. Law and advocacy	0.005	0.001	0.080	0.000	0.000	6.970†	1.000	0.993
8.1. Grant-making foundations	0.003	0.003	0.002	-0.048	-0.007	1.037	0.334†	0.880
8.2. Umbrella bodies	0.053	0.047	0.028	-0.013	0.012	1.124	0.947	1.053
9. International	0.007	0.002	0.061	-0.017	-0.011	2.655†	0.728	0.809
10. Religion	0.041	0.048	-0.032	0.068	0.033	0.870	1.310	1.148
<i>Locale:</i>								
Local	0.568	0.536	0.065	0.055	0.003	0.988	0.989	1.000
Greater London Area	0.038	0.039	-0.005	-0.032	-0.027	0.977	0.852	0.876
UK-wide	0.286	0.274	0.027	-0.039	0.001	1.028	0.961	1.001
International	0.108	0.152	-0.129†	-0.012	0.008	0.753	0.975	1.017

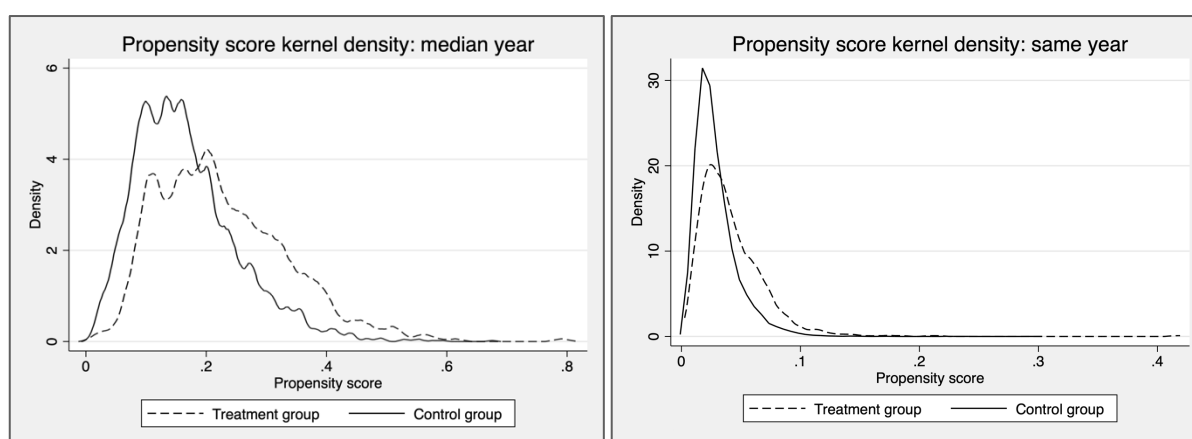
Notes: Matched using propensity scores generated from a logit model of propensity to recruit the first ever trustee aged under 31, using nearest neighbour matching with one neighbour (NN) and augmented inverse-probability weighting (AIPW). The reference (*t*) calendar year for control charities is the median year they are present in the sample. Standardised differences weight the difference in covariate means between treatment and control groups by the square root of the mean of their variances. Variance ratios are the mean ratio of the variance of each covariate between treated and control charities (Austin, 2009). † indicates standardised differences > 0.1, and variance ratios >2 or < 0.5.

TABLE 13 — PRE-TREATMENT COVARIATE BALANCE STATISTICS: SAME YEAR SAMPLE

Statistic: Sample:	Unmatched means		Standardised differences		Variance ratio	
	Treated (1)	Controls (2)	Raw (3)	Matched: NN (4)	Raw (7)	Matched: NN (8)
Pre-treatment covariate						
Mean age	57.135	58.621	-0.254†	-0.041	1.026	0.896
Standard deviation age	9.065	8.560	0.169†	0.060	1.054	1.147
Proportion female	0.348	0.303	0.218†	-0.092	1.094	0.982
Proportion non-British	0.066	0.084	-0.108†	0.000	0.661	1.019
Ln(Board size)	2.107	2.077	0.063	0.015	0.801	0.945
Ln(Total Assets)	14.424	14.938	-0.339†	-0.043	0.880	0.909
Leverage	0.042	0.055	-0.106†	-0.006	0.649	0.913
Ln(Total income)	14.495	14.761	-0.246†	-0.008	0.840	1.047
Ln(Voluntary income)	10.089	10.196	-0.021	-0.004	0.981	1.036
Spending ratio	0.856	0.836	0.072	0.051	0.651	0.890
<i>ICNPO classification:</i>						
1. Culture and recreation	0.164	0.113	0.149†	-0.096	1.373	0.772
2. Education	0.222	0.289	-0.155†	-0.031	0.841	0.970
2.2. Playgroups and nurseries	0.011	0.009	0.027	0.122†	1.309	2.723†
2.3. Research	0.020	0.031	-0.069	-0.018	0.653	0.902
3. Health	0.144	0.184	-0.106†	0.035	0.826	1.057
4. Social Services	0.305	0.235	0.159†	-0.005	1.183	0.994
4.1. Scout groups and youth clubs	0.002	0.000	0.035	-0.015	3.314†	0.500
5. Environment	0.003	0.014	-0.118†	-0.053	0.234†	0.603
6.1. Development	0.010	0.020	-0.083	0.084	0.502	1.658
6.2. Housing	0.005	0.003	0.023	-0.034	1.440	0.494†
6.3. Employment and training	0.005	0.004	0.020	-0.008	1.362	0.872
7. Law and advocacy	0.005	0.000	0.085	-0.021	9.910†	0.308†
8.1. Grant-making foundations	0.003	0.003	0.004	0.043	1.070	1.879
8.2. Umbrella bodies	0.053	0.047	0.024	0.017	1.107	1.074
9. International	0.007	0.002	0.062	-0.031	2.697†	0.483†
10. Religion	0.041	0.045	-0.022	0.039	0.909	1.177
<i>Locale:</i>						
Local	0.568	0.550	0.037	0.020	0.993	0.996
Greater London Area	0.038	0.038	-0.003	-0.009	0.988	0.959
UK-wide	0.286	0.267	0.042	0.015	1.045	1.015
International	0.108	0.145	-0.110†	-0.043	0.781	0.913

Notes: Matched using propensity scores generated from a logit model of propensity to recruit the first ever trustee aged under 31, using nearest neighbour matching with one neighbour (NN). Control charities are included using every calendar year 2008-2015 as their reference year (t) in turn. Standardised differences weight the difference in covariate means between treatment and control groups by the square root of the mean of their variances. Variance ratios are the mean ratio of the variance of each covariate between treated and control charities (Austin, 2009). † indicates standardised differences > 0.1, and variance ratios >2 or <0.5.

FIGURE 3 — PROPENSITY SCORE OVERLAP



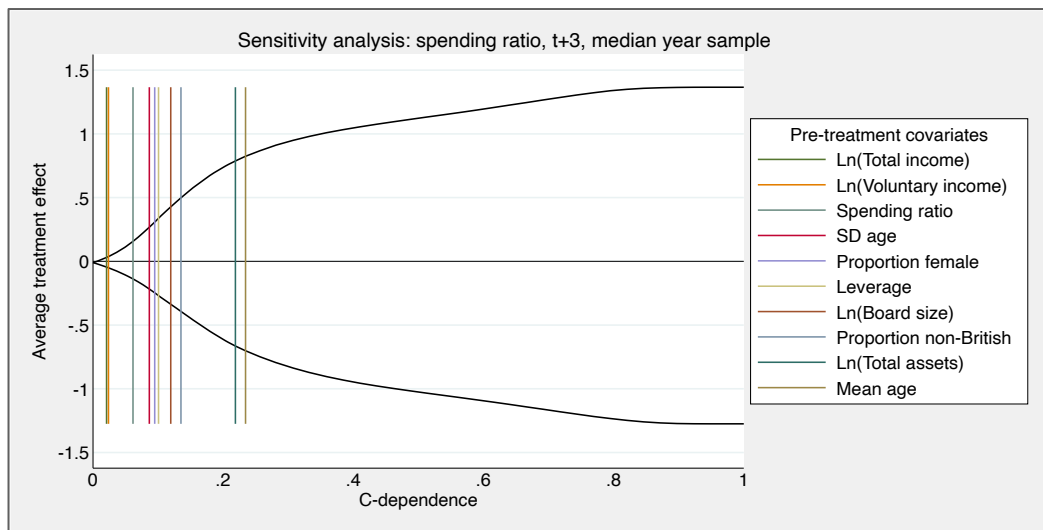
Notes: Median year sample sets the reference (t) calendar year for control charities as the median year they are present in the sample, same year includes control charities using every calendar year 2008-2015 as their reference in turn.

TABLE 14 — AVERAGE TREATMENT EFFECT ESTIMATES

Matching method:	Nearest neighbour (median year)			AIPW (median year)			Nearest neighbour (same year)		
	t+1	t+2	t+3	t+1	t+2	t+3	t+1	t+2	t+3
Outcome variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Ln(Total income)	-0.016 (0.013)	-0.004 (0.018)	0.005 (0.017)	-0.010 (0.015)	-0.002 (0.016)	0.010 (0.016)	-0.005 (0.013)	0.020 (0.016)	0.015 (0.017)
Ln(Voluntary income)	-0.232* (0.137)	0.002 (0.149)	-0.022 (0.164)	-0.124 (0.139)	-0.014 (0.165)	0.034 (0.178)	-0.149 (0.239)	-0.104 (0.220)	-0.121 (0.244)
Spending ratio	0.015 (0.011)	-0.008 (0.013)	-0.021* (0.013)	0.009 (0.012)	-0.013 (0.012)	-0.026** (0.012)	0.011 (0.014)	-0.007 (0.014)	0.003 (0.018)

Notes: Standard errors in brackets: robust Abadie-Imbens (2012) for nearest neighbour and median year sample, robust Abadie-Imbens (2011) for nearest neighbour and same year sample, and robust to clustering at the charity level for AIPW. Estimates are the average treatment effects of recruiting a first trustee aged under 31 on the change in each financial performance measure between the comparison period and period t . Estimates are generated using nearest neighbour matching with one neighbour and augmented inverse-probability weighting (AIPW). Median year sample includes control charities once using the median calendar year they are in-sample as their reference (t) year, same year includes control charities using every calendar year 2008-2015 in turn.

FIGURE 4 — SENSITIVITY ANALYSIS



Notes: Sensitivity analysis of the negative ATE of recruiting a first trustee aged under 31 on the change in the spending ratio between t and $t + 3$, using the median year sample. Black lines show the bounds of the ATE implied by different levels of conditional c-dependence (Masten & Poirier, 2018). Vertical lines show the maximum change in propensity score that would be induced by leaving out the pre-treatment covariate indicated, as benchmarks of the level of c-dependence leaving out each would imply.

5.3 Robustness and extensions

I now briefly discuss seven robustness checks and extensions. Results tables, in the Appendix, show fixed effects and IV estimates only. Appendix 10 presents the baseline results in the same manner, to aid comparisons.

5.3.1 Data characteristics

As charities are highly heterogeneous, some variables have a few extreme observations that could drive estimates. I already mitigate the extreme positive skew of some variables by expressing them in natural logarithms. Going further, Appendix 11 presents results when the dependent variables, total assets and leverage are winsorized at 1% and 99% of their sample distributions. Results are broadly unchanged.

14.1% of trustee genders are classified by first name. Although common in the literature (e.g. Ahern & Dittmar, 2012), this may not always be accurate, such as for unisex names, making the 'proportion female' variable imprecise and potentially disguising its true causal effect. Merely excluding charities with trustees whose genders cannot all be classified by their titles could increase sample selection bias. For example, health charities have more trustees titled "Dr". Instead, Appendix 12 shows results after dropping any charity that has any trustee gender in any year that Genderize.io reports is predicted with less than 95% certainty or based on under 50 reference gender-name pairs. All fixed effects and IV estimates of the effect of the proportion of female trustees on total income, except one, become statistically insignificant, further evidence they are not robust. This could also be due to the lowering of statistical power, as the sample size reduces by 28%.

A final dataset concern is that I only observe financial variables from the charity register when annual income exceeds £500,000, which could cause non-random sample entry and exit. Therefore, in Appendix 13 I use the

Fame database to exclude charities whose annual income falls below £500,000 in *any* year 2007-18. Total observations reduce by 18%, but the results are broadly consistent with the core estimates.

5.3.2 Subsamples

As I fail to find robust effects young or female trustees on financial performance when considering all charities, it is natural to consider whether effects do exist for particular charity types. In *Appendix 14* I report results when including only the 5,038 charities which are involved with helping children and young people. One might imagine that in these charities young trustees would be particularly well placed to suggest improvements to services or new ways of engaging beneficiary groups that might improve total and voluntary income. The estimates do not support this hypothesis, with coefficients and their statistical significance almost identical to the core specification.

Similarly, *Appendix 15* tabulates results when only studying the 426 in-sample charities whose purpose includes advancing human rights, or promoting religious or racial harmony, or equality and diversity. This includes “the elimination of discrimination on the grounds of age, sex or sexual orientation” (Charity Commission, 2013). The estimates do not suggest any differential effect of age or gender diversity, with one of the two statistically significant age and gender coefficients only being very weakly so, and the other being economically insignificant.

5.3.3 Model specification

It might be that the ‘trustee connections’ instruments merely proxy for the overall connectedness of boards, violating the exclusion restriction. In line with Adams and Ferreira (2009), in *Appendix 16* I thus include an additional control of the total number of external directorships that trustees hold. Results are generally robust, though it is notable that the magnitude of the effect of female proportion on total income in the proportion aged under 31 specification falls, and it is now statistically insignificant. This lends further credence to the original being unreliable.

Although I do not find any robust effect of young or female trustees when considered separately, *Appendix 17* includes an interaction term between proportion female and the age explanatory variables. All except one are statistically insignificant, and this result too becomes insignificant when using IV estimation.

The sampled charities have higher mean female representation than in many firm studies: 35%, compared to 17% in the UK SMEs studied by Shehata et al. (2017), for example. Indeed, 18% of observations have a majority-female board. It is thus reasonable to consider a direct measure of gender *diversity* in place of ‘proportion female’. In *Appendix 18* I thus use the Blau index which, here, gives the probability that two trustees selected at random from a board are of different genders (Rushton, 2008). The connections instrument for gender has a weak first stage when using this measure, so I only report fixed effects estimates. There are no statistically significant estimates of the effect of gender diversity on any financial performance measure.

6. Discussion

These results, which fail to find a robust effect of trustee age or gender on the financial performance of large UK charities, are compatible with several explanations.

Firstly, it may be that there is a consistent effect of young and female trustees on charity financial performance, but it is insufficiently strong to be detected by econometric analysis. My study, with 320 times as many observations as the most comparable prior work, Elmagrhi et al. (2018), presents something of a best case for identifying an effect if this is true. Hermalin and Weisbach (2005) argue board characteristics should be modelled as leading to board actions, which in turn affect performance. They write (p. 12) that when estimating the direct link “errors from both underlying equations are present, so the signal-to-noise ratio is low”. Research of firms has begun to take heed, with Adams and Ferreira (2009) examining the impact of directors’ gender on board inputs such as attendance, and on governance characteristics such as the relationship between CEO turnover and stock performance. Future non-profit research should explore this avenue. In 2018 the Charity Commission (2018b) began publishing the number of employees of each charity receiving total benefits in different bands above £60,000. Once a few years have been released, effects of trustee characteristics on top remuneration could be investigated, as an example.

Secondly, it might be that young and female trustees do affect financial performance, but not in a consistent manner. Charities are highly heterogeneous in their charitable purposes, service provision, beneficiary groups, organisational structures, and financing mechanisms (Epstein & McFarlan, 2011). Adams (2019) argues that there are very few corporate strategies that have been shown to uniformly change firm value, and board diversity should not be held to a higher standard. Carter (2010) expresses a similar view in terms of contingency theory, which posits that the optimal way to run an organisation depends on circumstances (Fiedler, 1981). I have investigated some circumstances in which trustee characteristics may be particularly relevant for performance, i.e. charities working with young people or promoting diversity, but future research should explore others. It may also be that the age and gender characteristics of trustees are already at an efficient equilibrium in most trustee boards. Ahern and Dittmar (2012) provide evidence this can be the case in firms, demonstrating that a Norwegian quota for 40% female directors led to immediate stock price declines and medium-term falls in market value for firms most impacted. This is not possible to test without a similar exogenous shock to trustee characteristics.

A third possibility is that there is no causal effect of young and female trustees, recruited and integrated as they currently are, but that there could be if their appointment were more widespread. For example, of the observations where a positive number of trustees aged under 31 are on a board, in 75.8% of cases only one is present. It thus might be the case that young and female trustees are recruited as tokens of diversity, or “window dressing” (Carter et al., 2003, p. 44), and so do not have the influence to affect performance. There is some qualitative support for this, with a Charity Commission (2010, p. 21) report finding from focus groups that “not being taken seriously” is a key challenge facing young trustees. Selection effects must also be acknowledged. Adams and Ragunathan (2017) present evidence that despite prior findings that women are, on average, more risk averse than men, female directors of Swedish companies are *less* risk averse than male directors. Further research regarding the characteristics of the female and young individuals who are currently trustees would be valuable in illustrating if performance effects could differ if more were recruited.

Finally, it may simply be the case that trustee gender and age have no causal effect on charity performance. Board members may have negligible impact on outcomes, or age and gender may be inconsequential in determining what impact they do have. Ahern and Dittmar's (2012) evidence suggests the former is not true for firms, and indeed they argue the lowering of the age and experience level of female directors recruited in response to the quotas may have caused the fall in performance. However, this evidence base is less established for non-profits.

7. Conclusion

This paper provides empirical evidence of the effect of the age and gender composition of trustee boards on the financial performance of charities. Such links are implied by UK regulators, sector leaders and campaign groups, but thus far have been largely untested. My dissertation finds no robust effect of the proportion of female trustees, proportion of trustees aged under 31, standard deviation of trustee age or mean trustee age on total income, voluntary income, or the ratio of charitable expenditure to total income. Some effects, consistent with prior research, are apparent in pooled OLS specifications, but are eliminated by the inclusion of charity fixed effects or the use of instrumental variables based on trustee connections to other directors. A propensity score matched difference-in-differences analysis confirms effects of recruiting a first young trustee are not evident even after they have had time to build influence on a board. These results are robust to different assumptions about my dataset and considering different subsamples and specifications. It thus makes several original contributions, as to the best of my knowledge there has been no previous research investigating any links between trustee age and non-profit performance that has not relied on cross-sectional regression, none examining a direct gender-performance link that has not relied on pooled OLS estimation, and none testing the effect of recruiting a first ever young trustee. It also pioneers a new combination of datasets that is richer and over three hundred times the size of that employed by the most similar prior research. This could help lower the barriers to entry for future non-profit governance research.

Nevertheless, my study has limitations. I only provide evidence from the UK, and just as the effects of firm diversity board diversity can vary with cultural and institutional context (Post & Byron, 2015), they may do so similarly for charities. My data only covers large charitable companies, not other types of charities nor smaller ones, though this currently appears the only possible route with publicly available data. Mitigating endogeneity convincingly is highly challenging in this literature, and I have discussed the assumptions underlying my instrumental variable strategy and why they might fail. It is also reasonable to imagine that variables other than gender and age are endogenous. Similarly, selection on unobservables is possible in my propensity score analysis. Attempts should continue to identify new instruments and natural experiments, in the for-profit and non-profit governance literatures, to achieve cleaner identification. It should also be noted that although I focus on financial performance, and financial means may reasonably translate into better achievement of charitable aims, ideally it is this social performance that would be evaluated directly. Because charities are so heterogeneous the nature of successful financial performance, particularly regarding spending ratios, can also vary greatly. Consistently measuring performance is a broader challenge in non-profit research (Crutchfield & Grant, 2012).

My results have several implications. They demonstrate the importance of non-profit research following the more mature for-profit literature in adopting panel data methods and identification strategies to demonstrate more plausibly causal relationships. It is important to note that I do not find any *negative* effect of young or female trustees on financial performance. However, those advocating trustee diversity based on uniformly positive performance effects, explicitly or implicitly, should perhaps reconsider their case. Many other arguments remain, including a moral case for equal opportunities and being representative of communities served, and developing the skills of future charity leaders. Policy, especially if mandates such as diversity quotas are being considered, should be written based on evidence rather than supposition.

Does difference help make a difference? Perhaps, but different approaches are needed to prove it.

8. References

- Abadie, A. & Imbens, G., 2011. Bias-corrected matching estimators for average treatment effects. *Journal of Business & Economic Statistics*, 29(1), pp. 1-11.
- Abadie, A. & Imbens, G., 2012. *Matching on the estimated propensity score: Harvard University and National Bureau of Economic Research*. [Online]
Available at: <https://www.nber.org/papers/w15301>
[Accessed 5 April 2021].
- Abadie, A. & Imbens, G., 2016. Matching on the estimated propensity score. *Econometrica*, 84(2), pp. 781-807.
- Adams, R., 2019. *Women on boards: The superheroes of tomorrow?*. [Online]
Available at: <https://www.youtube.com/watch?v=qrgBr52LcVs>
[Accessed 5 April 2021].
- Adams, R. & Ferreira, D., 2009. Women in the boardroom and their impact on governance and performance. *Journal of Financial Economics*, 94(2), pp. 291-309.
- Adams, R. & Raganathan, V., 2017. Lehman Sisters. SSRN: <https://ssrn.com/abstract=3046451>.
- Ahern, K. & Dittmar, A., 2012. The Changing of the Boards: The Impact on Firm Valuation of Mandated Female Board Representation. *Quarterly Journal of Economics*, 127(1), pp. 137-197.
- Ali, M., Ng, Y. L. & Kulik, C., 2014. Board Age and Gender Diversity: A Test of Competing Linear and Curvilinear Prediction. *Journal of Business Ethics*, Volume 125, pp. 497-512.
- Anderson, M., 2007. *Stata code to compute "sharpened" False Discovery Rate (FDR) adjusted q-values*. [Online]
Available at: http://are.berkeley.edu/~mlanderson/downloads/fdr_sharpened_qvalues.do.zip
[Accessed 5 April 2021].
- Ararat, M., Aksu, M. & Cetin, A., 2015. How Board Diversity Affects Firm Performance in Emerging Markets: Evidence on Channels in Controlled Firms. *Corporate Governance: An International Review*, 23(2), pp. 83-103.
- Austin, P., 2009. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in Medicine*, 28(25), pp. 3083-3107.
- Austin, P., 2010. Statistical Criteria for Selecting the Optimal Number of Untreated Subjects Matched to Each Treated Subject When Using Many-to-One Matching on the Propensity Score. *American Journal of Epidemiology*, 172(9), pp. 1092-1097.
- Bascle, G., 2008. Controlling for endogeneity with instrumental variables in strategic management research. *Strategic Organisation*, 6(3), pp. 285-327.
- Baum, C., Schaffer, M. & Stillman, S., 2010. *ivreg2: Stata module for extended instrumental variables/2SLS, GMM and AC/HAC, LIML and k-class regression*. [Online]
Available at: <http://ideas.repec.org/c/boc/bocode/s425401.html>
[Accessed 1 May 2021].
- Benjamini, Y., Krieger, A. M. & Yekutieli, D., 2006. Adaptive Linear Step-up Procedures that Control the False Discovery Rate. *Biometrika*, 93(3), pp. 491-507.
- Boateng, A., Akamavi, R. & Ndor, G., 2016. Measuring performance of non-profit organisations: evidence from large charities. *Business Ethics: A European Review*, 25(1), pp. 59-74.
- Bonn, I., Yoshikawa, T. & Phan, P., 2004. Effects of board structure on firm performance: A comparison of Japan and Australia. *Asian Business & Management*, 3(1), pp. 105-125.
- Borjas, G., 2016. *Labor Economics*. 7th ed. New York: McGraw-Hill Education.
- Breusch, T. S. & Pagan, A. R., 1980. The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics. *The Review of Economic Studies*, 47(1), pp. 239-253.
- Buse, K., Bernstein, R. S. & Bilimoria, D., 2016. The Influence of Board Diversity, Board Diversity Policies. *Journal of Business Ethics*, Volume 133, pp. 179-191.
- BVD, 2020. *Fame: the leading UK company research solution*. [Online]
Available at: <https://fame.bvdinfo.com/>
[Accessed 29 November 2020].
- Callen, J., Klein, A. & Tinkelman, D., 2009. The Contextual Impact of Nonprofit Board Composition and Structure on Organizational Performance: Agency and Resource Dependence Perspectives. *VOLUNTAS*, Volume 21, pp. 101-125.

- Campbell, K., Minguez-Vera & Antonio, 2007. Gender Diversity in the Boardroom and Firm Financial Performance. *Journal of Business Ethics*, Volume 83, pp. 435-451.
- Carter, D., D'Souza, F., Simkins, B. & Simpson, G., 2010. The Gender and Ethnic Diversity of US Boards and Board Committees and Firm Financial Performance. *Corporate Governance: An International Review*, 18(5), pp. 396-414.
- Carter, D., Simkins, B. & Simpson, G., 2003. Corporate Governance, Board Diversity, and Firm Value. *The Financial Review*, Volume 38, pp. 33-53.
- Charities Aid Foundation, 2015. *Young Trustees Guide: Developing the next generation of charity leaders*. [Online] Available at: https://www.cafonline.org/docs/default-source/about-us-publications/youngtrusteesreport_1682a_web_080915.pdf [Accessed 28 November 2020].
- Charity Commission, 2005a. *The Essential Trustee: What you need to know*. [Online] Available at: https://www.ihbc.org.uk/recent_papers/docs/Charity_Commission.pdf [Accessed 28 November 2020].
- Charity Commission, 2005b. *Start as you mean to go on: Trustee recruitment and induction report*. [Online] Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/284724/rs10text.pdf [Accessed 2 May 2021].
- Charity Commission, 2010. *A Breath of Fresh Air: young people as charity trustees*. [Online] Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/284702/rs23text.pdf [Accessed 28 November 2020].
- Charity Commission, 2013. *Guidance: Charitable Purposes*. [Online] Available at: <https://www.gov.uk/government/publications/charitable-purposes/charitable-purposes> [Accessed 5 May 2021].
- Charity Commission, 2014. *Charities Financial Reporting Standards: SORP (FRS 102), First Edition*. [Online] Available at: https://www.charitycorp.org/media/619101/frs102_complete.pdf [Accessed 5 April 2021].
- Charity Commission, 2018a. *Recent charity register statistics: Charity Commission*. [Online] Available at: <https://www.gov.uk/government/publications/charity-register-statistics/recent-charity-register-statistics-charity-commission> [Accessed November 28 2020].
- Charity Commission, 2018b. *New questions in the 2018 and 2019 charity annual return service*. [Online] Available at: <https://www.gov.uk/government/news/charity-annual-return-2018> [Accessed 10 April 2021].
- Charity Commission, 2020a. *Search the charity register*. [Online] Available at: <https://www.gov.uk/find-charity-information> [Accessed 29 November 2020].
- Charity Commission, 2020b. *The Register of Charities: Notes for users, Income*. [Online] Available at: http://apps.charitycommission.gov.uk/Showcharity/ShowCharity_Help_Page.aspx?ContentType=Chart_Help_Income&SelectedLanguage=English [Accessed 4 December 2020].
- Clemens, M., 2013. Blunt Instruments: Avoiding Common Pitfalls in Identifying the Causes of Economic Growth: Online-only appendix. *American Economic Journal: Macroeconomics*, 5(2), pp. 152-186.
- Companies House, 2020. *Companies House API*. [Online] Available at: <https://developer.company-information.service.gov.uk/api/docs/> [Accessed 29 November 2020].
- Companies House, 2021. *Companies House Developer Forum*. [Online] Available at: <https://forum.aws.chdev.org/> [Accessed 10 March 2021].

- Canyon, M. & He, L., 2017. Firm performance and boardroom gender diversity: A quantile regression approach. *Journal of Business Research*, Volume 79, pp. 198-211.
- Cornforth, C., 2012. Nonprofit Governance Research: Limitations of the Focus on Boards and Suggestions for New Directions. *Nonprofit and Voluntary Sector Quarterly*, 41(6), pp. 1116-1135.
- Crutchfield, L. & Grant, H., 2012. *Forces for Good: The Six Practices of High-Impact Nonprofits*. s.l.:John Wiley & Sons.
- Darmadi, S., 2011. Board diversity and firm performance: the Indonesian evidence. *Corporate Ownership and Control*, Volume 8.
- Demografix ApS, 2020. *Genderize.io: Determine the gender of a name*. [Online]
Available at: <https://genderize.io/>
[Accessed 29 November 2020].
- Elmagrhi, M. et al., 2018. Trustee board diversity, governance mechanisms, capital structure and performance in UK charities. *Corporate Governance*, 18(3), pp. 478-508.
- Epstein, M. & McFarlan, W., 2011. Measuring the efficiency and effectiveness of a nonprofit's performance. *Strategic finance*, 93(4), pp. 27-34.
- Faleye, O., 2007. Classified boards, firm value, and managerial entrenchment. *Journal of Financial Economics*, 83(2), pp. 501-529.
- Farrell, K. & Hersch, P., 2005. Additions to corporate boards: the effect of gender. *Journal of Corporate Finance*, 11(1-2), pp. 85-106.
- Fernández-Temprano, M. & Tejerina-Gaite, F., 2020. Types of director, board diversity and firm performance. *Corporate Governance*, 20(2), pp. 324-342.
- Fiedler, F., 1981. Leadership effectiveness. *American Behavioral Scientist*, 24(5), pp. 619-632.
- Garcia-Rodriguez, I., Romero-Merino, Elena, M. & Santamaria-Mariscal, M., 2020. The role of boards in the financial vulnerability of nonprofit organisations. *Financial Accountability & Management in Governments, Public Services and Charities*.
- Glynn, A. & Quinn, K., 2010. An Introduction to the Augmented Inverse Propensity Weighted Estimator. *Political Analysis*, 18(1), pp. 36-56.
- Gong, Y. & Girma, S., 2020. Top management gender diversity and performance: in search of threshold effects. *Cambridge Journal of Economics*, Volume beaa045.
- Grilli, L. & Rampichini, C., 2011. *Propensity scores for the estimation of average treatment effects in observational studies*. [Online]
Available at: <https://www.bristol.ac.uk/media-library/sites/cmm/migrated/documents/prop-scores.pdf>
[Accessed 10 April 2021].
- Harris, E., 2014. The Impact of Board Diversity and Expertise on Nonprofit Performance. *Nonprofit Management & Leadership*, 25(2), pp. 113-130.
- Hermalin, B. & Weisbach, M., 2005. Boards of directors as an endogenously determined institution: A survey of the economic literature. *Economic Policy Review*, 9(1), pp. 7-26.
- Hillman, A., Cannella, A. & Paetzold, R., 2002. The Resource Dependence Role of Corporate Directors: Strategic Adaptation of Board Composition in Response to Environmental Change. *Journal of Management Studies*, 37(2), pp. 235-256.
- Hillman, A., Shropshire, C. & Cannella, A., 2007. Organizational Predictors of Women on Corporate Boards. *The Academy of Management Journal*, 50(4), pp. 941-952.
- Inclusive Boards, 2018. *Charities: Inclusive Governance*. [Online]
Available at: <https://www.inclusiveboards.co.uk/wp-content/uploads/2019/10/Charities-Inclusive-Governance.pdf>
[Accessed 29 November 2020].
- Kunze, F., Boehm, S. & Bruch, H., 2011. Age diversity, age discrimination climate and performance consequences—a cross organizational study. *Journal of Organizational Behaviour*, 31(2), pp. 264-290.
- Larikov, A., 2021. *update.py*. [Online]
Available at: https://github.com/JuggleJobs/companies_house/blob/master/companies_house/update.py
[Accessed 5 April 2021].

- Levi, M., Li, K. & Zhang, F., 2014. Director gender and mergers and acquisitions. *Journal of Corporate Finance*, Volume 28, pp. 185-200.
- Liu, Y., Wei, Zuobao & Zie, F., 2014. Do women directors improve firm performance in China?. *Journal of Corporate Finance*, Volume 28, pp. 169-184.
- Mahadeo, Jyoti, Soobaroyen, T. & Hanuman, V., 2012. Board Composition and Financial Performance: Uncovering the Effects of Diversity in an Emerging Economy. *Journal of Business Ethics*, 105(3), pp. 375-388.
- Masten, M. & Poirier, A., 2018. Identification of Treatment Effects Under Conditional Partial Independence. *Econometrica*, 86(1), pp. 317-351.
- McDonnell, D. & Rutherford, A., 2019. Promoting charity accountability: understanding disclosure of serious incidents. *Accounting Forum*, 43(1), pp. 42-61.
- Mishra, R. K. & Jhunjhunwala, S., 2013. *Diversity and the Effective Corporate Board*. 1st ed. Oxford: Academic Press.
- Mistler, E., 2018. *companies-house 0.1.2: Dynamic API wrapper for Companies' House*. [Online] Available at: <https://pypi.org/project/companies-house/> [Accessed 5 April 2021].
- NCVO, 2020. *How do we classify subsectors?*. [Online] Available at: <https://almanac.fc.production.ncvocloud.net/about/subsector-classification/> [Accessed 5 April 2021].
- ONS, 2019. *Overview of the UK population: August 2019*. [Online] Available at: <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/articles/overviewoftheukpopulation/august2019> [Accessed 20 November 2020].
- ONS, 2020. *Population estimates by output areas, electoral, health and other geographies, England and Wales: mid-2019*. [Online] Available at: <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualsmallareapopulationestimates/latest> [Accessed 2 November 2020].
- Pathan, S. & Faff, R., 2013. Does board structure in banks really affect their performance?. *Journal of Banking & Finance*, 37(5), pp. 1573-1589.
- Perkins, C. & Fields, D., 2010. Top Management Team Diversity and Performance of Christian Churches. *Nonprofit and Voluntary Sector Quarterly*, 39(5), pp. 825-843.
- Pfeffer, J. & Salancik, G., 1978. *The external control of organizations: A resource dependence perspective*. New York: Harper & Row.
- Post, C. & Byron, K., 2015. Women on Boards and Firm Financial Performance: A Meta-Analysis. *Academy of Management Journal*, 58(5), pp. 1546-1571.
- Reed, W., 2015. On the Practice of Lagging Variables to Avoid Simultaneity. *Oxford Bulletin of Economics and Statistics*, 77(6), pp. 897-905.
- Roodman, D., 2009. How to do xtabond2: An introduction to difference and system GMM in Stata. *The Stata Journal*, 9(1), pp. 86-136.
- Rosenbaum, P. & Rubin, D., 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), pp. 41-55.
- Rubin, D., 1974. Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies. *Journal of Educational Psychology*, 66(5), pp. 688-701.
- Rushton, M., 2008. A note on the use and misuse of the racial diversity index. *The Policy Studies Journal*, 36(3), pp. 445-459.
- Sayer, K., 2018. The best metrics for a charity are outcomes for beneficiaries. *Third Sector*, 27 March.
- Shehata, N., Salhin, A. & El-Helaly, M., 2017. Board diversity and firm performance: evidence from the U.K. SMEs. *Applied Economics*, 49(48), pp. 4817-4832.
- Siciliano, 1996. The Relationship of Board Member Diversity to Organizational Performance. *Journal of Business Ethics*, 15(12), p. 1996.

- Smith, J. & Todd, P., 2005. Does matching overcome LaLonde's critique of nonexperimental estimators?. *Journal of Econometrics*, Volume 125, pp. 305-353.
- Smith, N., Smith, V. & Verner, M., 2006. Do women in top management affect firm performance? A panel study of 2,500 Danish firms. *International Journal of Productivity and Performance Management*, 55(7), pp. 569-593.
- Stock, J. & Yogo, M., 2005. Testing for Weak Instruments in Linear IV Regression. In: *Identification and Inference for Econometric Models*. New York: Cambridge University Press, pp. 80-108.
- Stone, M. & Ostrower, F., 2007. Acting in the Public Interest? Another Look at Research on Nonprofit Governance. *Nonprofit and Voluntary Sector Quarterly*, 36(3), pp. 416-438.
- Stuart, E., 2010. Matching methods for causal inference: A review and a look forward. *Statistical science : a review journal of the Institute of Mathematical Statistics*, 25(1), pp. 1-21.
- Talavera, O., Yin, S. & Zhang, M., 2018. Age diversity, directors' personal values, and bank performance. *International Review of Financial Analysis*, Volume 55, pp. 60-79.
- Wais, K., 2016. Gender Prediction Methods Based on First Names with genderizeR. *The R Journal*, 8(1), pp. 17-37.
- Wintoki, B., Linck, J. & Netter, J., 2012. Endogeneity and the dynamics of internal corporate governance. *Journal of Financial Economics*, 105(3), pp. 581-606.
- Wooldridge, J., 2018. *Introductory Econometrics: A Modern Approach*. 7th Edition ed. Boston: Cengage.
- Young Trustees Movement, 2019. *Press Release: The Young Trustees Movement: Will you pledge?*. [Online] Available at: <https://www.iwill.org.uk/wp-content/uploads/2019/10/The-Young-Trustees-Movement-PRESS-RELEASE-.pdf> [Accessed 28 November 2020].
- Young Trustees Movement, 2020. *Benefits*. [Online] Available at: <https://youngtrusteesmovement.org/benefits> [Accessed 28 November 2020].

9. Appendix

APPENDIX 1 — NOTES: COMPANIES HOUSE API

I accessed the Companies House (2020) API using the 'companies-house' Python script by Elias Mistler (2018), incorporating a revision from Alexander Larikov (2021) to account for a change in the API's reference URL (currently <https://developer-specs.company-information.service.gov.uk/companies-house-public-data-api/reference>). Drawing on a number of posts in the Companies House (2021) Developer Forum, I implemented rate-limiting using the "X-Ratelimit-Remain" and "X-Ratelimit-Reset" headers to ensure only 600 requests were sent per five-minute period. I further automated the paging through of requests containing more than 100 items, added error handling for 502 and 500 errors, and cached queries to disk as they were retrieved. I am happy to share this modified package on request.

I used the "list_company_officers" operation to retrieve the names, nationalities, months and years of birth, appointment dates, resignation dates and roles of all officers of the charities in my sample by their company number. I then used the unique officer identifiers (obtained from "items[].links.self") for officers active in my period of interest to obtain their titles using "list_officers_appointments". I classified trustees as male if they had the following titles: Mr, Sir, Lord, Baron, Brother, Count, Deacon, Duke, Earl, Father, Fr, His, Mister, Monsieur, Prince. I classified them female if their titles were: Mrs, Ms, Miss, Baroness, Countess, Viscountess, Duchess, Her, Lady, Princess, Dame, Sister.

To construct my instrumental variables, I additionally used "list_officers_appointments" to obtain the company numbers of other companies each trustee in my sample sat on in the relevant period. I then used "list_company_officers" to obtain the details of all other directors of these other companies, again using "list_officers_appointments" for each of these to obtain their titles for gender classification.

APPENDIX 2 — CORE MODEL: SMOOTHED SPENDING RATIO

Dependent variable:	Spending ratio (two years smoothing)					Spending ratio (three years smoothing)					Spending ratio (four years smoothing)				
Age explanatory variable:	Proportion < 31		SD age		Mean age	Proportion < 31		SD age		Mean age	Proportion < 31		SD age		Mean age
Model:	FE	IV	FE	IV	FE	FE	IV	FE	IV	FE	FE	IV	FE	IV	FE
Independent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
<i>Age:</i>															
Proportion < 31	-0.008 (0.026)	0.336 (0.344)				0.009 (0.024)	0.487 (0.309)				0.023 (0.026)	0.428 (0.302)			
SD age			0.000 (0.001)	0.025 (0.027)				0.000 (0.000)	0.039 (0.030)				0.000 (0.000)	0.034 (0.029)	
Mean age					0.001* (0.000)					0.000 (0.000)					0.000 (0.001)
<i>Gender:</i>															
Proportion female	0.012 (0.010)	0.175 (0.133)	0.012 (0.010)	0.320 (0.247)	0.015 (0.011)	0.011 (0.009)	0.144 (0.106)	0.011 (0.009)	0.395 (0.263)	0.012 (0.009)	0.019 (0.013)	0.057 (0.102)	0.020 (0.013)	0.305 (0.294)	0.020* (0.012)
<i>Trustee controls:</i>															
Proportion non-British	-0.026 (0.026)	-0.037 (0.029)	-0.026 (0.026)	-0.042 (0.032)	-0.025 (0.026)	-0.033 (0.030)	-0.044 (0.032)	-0.033 (0.030)	-0.039 (0.037)	-0.033 (0.031)	-0.064 (0.065)	-0.072 (0.066)	-0.064 (0.065)	-0.070 (0.068)	-0.063 (0.066)
Ln(Board size)	-0.006 (0.005)	-0.018** (0.009)	-0.007 (0.006)	-0.068 (0.062)	-0.005 (0.005)	-0.001 (0.005)	-0.012 (0.008)	-0.001 (0.006)	-0.092 (0.067)	0.000 (0.005)	0.001 (0.004)	-0.006 (0.007)	0.001 (0.004)	-0.075 (0.066)	0.001 (0.004)
<i>Charity controls:</i>															
Ln(Total assets)	0.073*** (0.005)	0.075*** (0.005)	0.073*** (0.005)	0.076*** (0.006)	0.073*** (0.005)	0.070*** (0.005)	0.073*** (0.005)	0.070*** (0.005)	0.076*** (0.007)	0.070*** (0.005)	0.066*** (0.005)	0.067*** (0.005)	0.066*** (0.005)	0.069*** (0.006)	0.066*** (0.005)
Leverage	-0.036 (0.046)	-0.035 (0.048)	-0.036 (0.046)	-0.027 (0.049)	-0.035 (0.046)	-0.078*** (0.014)	-0.079*** (0.015)	-0.078*** (0.014)	-0.060*** (0.022)	-0.078*** (0.014)	-0.114*** (0.037)	-0.116*** (0.037)	-0.113*** (0.037)	-0.099** (0.040)	-0.113*** (0.037)
Observations	56,843	56,256	56,843	56,256	56,843	48,650	48,103	48,650	48,103	48,650	41,219	40,662	41,219	40,662	41,219
Number of charities	7,995	7,408	7,995	7,408	7,995	7,404	6,857	7,404	6,857	7,404	6,863	6,306	6,863	6,306	6,863
R-squared	0.018	0.010	0.018	-0.050	0.018	0.022	0.008	0.022	-0.208	0.022	0.026	0.016	0.026	-0.202	0.026

Notes: Heteroskedasticity-robust standard errors, adjusted for clustering at the charity level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All models include year dummies and a constant term. All independent variables are lagged by one year. FE models include charity-level fixed effects, IV models additionally use the proportion of male (aged 31 or over) trustees sitting on another board with a female (aged under 31) trustee as an instrument for 'proportion female' ('proportion < 31' and 'standard deviation age'). Dependent variables are constructed using equally-weighted one-sided moving averages using the contemporary value and an additional one, two or three lead values.

APPENDIX 3 — CORE MODEL: ALTERNATIVE VOLUNTARY INCOME SPECIFICATIONS

Dependent variable:	Voluntary income (inverse hyperbolic sine transformed)					Ln(Voluntary income) (drop zero values)				
Age explanatory variable:	Proportion < 31		SD age		Mean age	Proportion < 31		SD age		Mean age
Model:	FE	IV	FE	IV	FE	FE	IV	FE	IV	FE
Independent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<i>Age:</i>										
Proportion < 31	0.211 (0.435)	5.129 (3.756)				-0.062 (0.165)	-0.020 (1.609)			
SD age			0.001 (0.007)	0.375 (0.284)				-0.004 (0.003)	-0.001 (0.120)	
Mean age					-0.005 (0.006)					0.004 (0.003)
<i>Gender:</i>										
Proportion female	-0.130 (0.161)	-0.545 (1.583)	-0.128 (0.161)	1.128 (2.559)	-0.146 (0.162)	-0.001 (0.070)	1.097 (0.711)	-0.001 (0.070)	1.089 (1.148)	0.012 (0.071)
<i>Trustee controls:</i>										
Proportion non-British	-0.438* (0.263)	-0.474* (0.284)	-0.436* (0.263)	-0.540 (0.340)	-0.443* (0.263)	0.024 (0.120)	-0.028 (0.130)	0.024 (0.120)	-0.027 (0.139)	0.028 (0.120)
Ln(Board size)	0.400*** (0.077)	0.353*** (0.120)	0.401*** (0.078)	-0.356 (0.620)	0.396*** (0.078)	0.171*** (0.034)	0.127** (0.054)	0.178*** (0.034)	0.129 (0.257)	0.176*** (0.035)
<i>Charity controls:</i>										
Ln(Total assets)	0.181*** (0.051)	0.197*** (0.053)	0.181*** (0.051)	0.214*** (0.062)	0.182*** (0.051)	0.137*** (0.029)	0.137*** (0.029)	0.137*** (0.029)	0.137*** (0.030)	0.136*** (0.029)
Leverage	-0.551** (0.241)	-0.603** (0.245)	-0.549** (0.241)	-0.539** (0.252)	-0.553** (0.241)	-0.141 (0.112)	-0.123 (0.113)	-0.141 (0.112)	-0.123 (0.114)	-0.137 (0.112)
Observations	66,277	65,442	66,277	65,442	66,277	56,468	55,532	56,468	55,532	56,468
Number of charities	8,865	8,030	8,865	8,030	8,865	8,280	7,344	8,280	7,344	8,280
R-squared	0.003	-0.003	0.003	-0.095	0.003	0.004	-0.006	0.004	-0.005	0.004

Notes: Heteroskedasticity-robust standard errors, adjusted for clustering at the charity level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. SD = standard deviation. All models include year dummies and a constant term. All independent variables are lagged by one year. FE models include charity-level fixed effects, IV models additionally use the proportion of male (aged 31 or over) trustees sitting on another board with a female (aged under 31) trustee as an instrument for 'proportion female' ('proportion < 31' and 'standard deviation age'). To mitigate zero values of voluntary income while improving positive skew, columns 1-5 transform it by the inverse hyperbolic sine function, and columns 6-10 drop zero value observations before taking the natural logarithm.

APPENDIX 4 — CHARITY CLASSIFICATIONS SIC MAPPING

Classifications (NCVO-adapted ICNPO)	SIC codes
	18129, 18130, 18140, 18201, 30120, 47781, 47990, 55209, 56301, 56302, 58110, 58141, 58142, 58190, 59111, 59112, 59113, 59120, 59131, 59132, 59133, 59140, 59200, 60100, 60200, 63910, 70210, 79120, 79909, 82301, 90010, 90020, 90030, 90040, 91011, 91012, 91020, 91030, 91040, 93110, 93120, 93130, 93199, 93210, 93290
1. Culture and recreation	
2. Education	81300, 85200, 85310, 85320, 85421, 85422, 85510, 85520, 85530, 85590, 85600
2.2. Playgroups and nurseries	85100, 88910
2.3. Research	71200, 72110, 72190, 72200, 74909
3. Health	32500, 86101, 86102, 86210, 86220, 86230, 86900, 87100, 87200, 87300, 87900, 96040
4. Social Services	31090, 47110, 47240, 47290, 55100, 55900, 56101, 56102, 56210, 56290, 63990, 64999, 68209, 82990, 84110, 84120, 84250, 88100, 88990, 95220, 96030, 96090
4.1. Scout groups and youth clubs	55202
5. Environment	1110, 1130, 1250, 1300, 1420, 1430, 1490, 1500, 1610, 1621, 1629, 2100, 2400, 3110, 3120, 3210, 3220, 9100, 38110, 38120, 38210, 38320, 39000, 42910, 47760, 74901, 75000
6.1. Development	16290, 20160, 30200, 32990, 33170, 33190, 35110, 36000, 41100, 41201, 41202, 43290, 43999, 45200, 45400, 49100, 49319, 49390, 50100, 50300, 51102, 52290, 61100, 62090, 63110, 63120, 64192, 64306, 64929, 66190, 68100, 68202, 68310, 68320, 69203, 70100, 70229, 71111, 71112, 71129, 73200, 74100, 74300, 74902, 77390, 78300, 81100, 82110, 82190, 82200, 82302, 82920, 84130, 84220, 84240, 98000
6.2. Housing	68201
6.3. Employment and training	78109, 78200, 85410
7. Law and advocacy	69102, 69109, 84230, 94200, 94920
8.1. Grant-making foundations	47190, 47510, 47599, 47610, 47791, 47799, 47820, 47910, 66300
8.2. Umbrella bodies	64209, 64301, 94110, 94120, 94990
9. International	84210, 99000
10. Religion	94910

APPENDIX 5 — CORE MODEL: ALTERNATIVE INSTRUMENT SPECIFICATIONS

Dependent variable:	Ln(Total income)				Ln(Voluntary income)				Spending ratio			
Age explanatory variable:	Proportion < 31		SD age		Proportion < 31		SD age		Proportion < 31		SD age	
Model:	IV all	IV drop	IV all	IV drop	IV all	IV drop	IV all	IV drop	IV all	IV drop	IV all	IV drop
Independent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
<i>Age:</i>												
Proportion < 31	-0.366 (0.525)	-0.464 (0.534)			5.960 (4.149)	5.147 (4.138)			0.762 (0.503)	0.762 (0.522)		
SD age			-0.021 (0.032)	-0.030 (0.036)			0.350 (0.252)	0.335 (0.280)			0.045 (0.031)	0.050 (0.037)
Mean age												
<i>Gender:</i>												
Proportion female	0.698** (0.354)	0.626* (0.346)	0.578 (0.480)	0.436 (0.516)	0.374 (2.860)	-0.461 (2.805)	2.321 (3.992)	1.640 (4.163)	0.129 (0.242)	0.100 (0.240)	0.378 (0.408)	0.411 (0.458)
<i>Trustee controls:</i>												
Proportion non-British	-0.029 (0.041)	-0.020 (0.041)	-0.025 (0.044)	-0.014 (0.045)	-0.505* (0.302)	-0.436 (0.296)	-0.572 (0.362)	-0.500 (0.354)	0.003 (0.025)	0.005 (0.025)	-0.005 (0.033)	-0.004 (0.035)
Ln(Board size)	0.073*** (0.022)	0.070*** (0.025)	0.114 (0.077)	0.129 (0.090)	0.290* (0.165)	0.322* (0.193)	-0.370 (0.606)	-0.326 (0.690)	-0.027 (0.017)	-0.027 (0.020)	-0.112 (0.074)	-0.123 (0.089)
<i>Charity controls:</i>												
Ln(Total assets)	0.221*** (0.011)	0.222*** (0.011)	0.220*** (0.011)	0.220*** (0.011)	0.197*** (0.052)	0.198*** (0.053)	0.209*** (0.059)	0.214*** (0.062)	0.075*** (0.006)	0.075*** (0.006)	0.077*** (0.006)	0.078*** (0.007)
Leverage	0.007 (0.034)	0.003 (0.034)	0.003 (0.034)	-0.003 (0.034)	-0.567** (0.234)	-0.585** (0.233)	-0.491** (0.243)	-0.520** (0.241)	-0.035 (0.047)	-0.032 (0.048)	-0.025 (0.048)	-0.023 (0.049)
Observations	65,442	64,348	65,442	64,348	65,442	64,348	65,442	64,348	65,442	64,348	65,442	64,348
Number of charities	8,030	7,914	8,030	7,914	8,030	7,914	8,030	7,914	8,030	7,914	8,030	7,914
R-squared	0.115	0.126	0.114	0.110	-0.006	-0.003	-0.099	-0.086	0.000	0.002	-0.094	-0.114

Notes: Heteroskedasticity-robust standard errors, adjusted for clustering at the charity level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. SD = standard deviation. All models include year dummies and a constant term. All independent variables are lagged by one year. All models use the proportion of male (aged 31 or over) trustees sitting on another board with a female (aged under 31) trustee as an instrument for 'proportion female' ('proportion < 31' and 'standard deviation age'). To deal with all female (all aged under 31) boards, the core model sets the instruments in these cases to 1. Here, 'IV all' columns use the proportion of all trustees, irrespective of their gender (age), sitting on another board with a female (aged under 31) trustee in those cases. 'IV drop' columns drop observations with all female or all young boards.

APPENDIX 6 — FIRST STAGES: LOCAL DEMOGRAPHICS INSTRUMENT

Dependent variable:	Proportion < 31	Standard deviation age	Mean age
Model:	Pooled OLS		
Independent variable	(1)	(2)	(3)
<i>Instruments:</i>			
Local proportion < 31	0.0668*** (0.00865)		
Local standard deviation age		-0.0481** (0.0161)	
Local mean age			0.169*** (0.0142)
<i>Gender:</i>			
Proportion female	0.0239*** (0.00361)	0.867*** (0.191)	-3.577*** (0.402)
<i>Trustee controls:</i>			
Proportion non-British	0.0111* (0.00565)	-0.601** (0.230)	-2.443*** (0.507)
Ln(Board size)	0.00184 (0.00261)	1.356*** (0.0999)	-0.131 (0.205)
<i>Charity controls:</i>			
Ln(Total assets)	-0.00331*** (0.000629)	-0.288*** (0.0271)	0.643*** (0.0567)
Leverage	0.00357 (0.00610)	0.333 (0.315)	-3.048*** (0.616)
Observations	47,282	47,282	47,282
Number of charities	6,637	6,637	6,637
Kleibergen-Paap rk LM statistic	59.72	8.91	142.71

Notes: Heteroskedasticity-robust standard errors, adjusted for clustering at the charity level, in parentheses.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. All models include a constant term and year, classification, and locale dummies. They do not include charity fixed effects. All variables are lagged by one year. "Local" variables are attributes of the population aged over 16 in the Lower Super Output Area in which each charity is based, in each year, using ONS (2020) population data. Sample includes only charities operating from a single postcode. The relevant Stock-Yogo (2005) critical value for 10% maximal relative bias is 16.38.

APPENDIX 7 — RESULTS: LOCAL DEMOGRAPHICS INSTRUMENT

Dependent variable:	Total income		Voluntary income		Spending ratio	
Age explanatory variable:	Proportion < 31	Mean age	Proportion < 31	Mean age	Proportion < 31	Mean age
Model:	Pooled OLS, local demographics IV					
Independent variable	(1)	(2)	(3)	(4)	(5)	(6)
<i>Age:</i>						
Proportion < 31	3.210*** (0.892)		1.718 (5.151)		0.889*** (0.282)	
Mean age		-0.034*** (0.009)		-0.050 (0.053)		-0.006** (0.003)
<i>Gender:</i>						
Proportion female	-0.109** (0.045)	-0.157*** (0.051)	-0.573** (0.258)	-0.725** (0.300)	0.045*** (0.015)	0.044** (0.017)
<i>Trustee controls:</i>						
Proportion non-British	0.141** (0.065)	0.098 (0.067)	0.585 (0.360)	0.479 (0.376)	0.015 (0.025)	0.011 (0.025)
Ln(Board size)	0.176*** (0.022)	0.179*** (0.021)	1.158*** (0.124)	1.154*** (0.123)	0.004 (0.008)	0.006 (0.008)
<i>Charity controls:</i>						
Ln(Total assets)	0.432*** (0.009)	0.444*** (0.010)	0.088* (0.046)	0.116** (0.055)	0.006 (0.005)	0.007 (0.006)
Leverage	-0.051 (0.068)	-0.146* (0.075)	-1.852*** (0.445)	-2.001*** (0.468)	-0.047 (0.041)	-0.064 (0.040)
Observations	47,282	47,282	47,282	47,282	47,282	47,282
Number of charities	6,637	6,637	6,637	6,637	6,637	6,637
R-squared	0.469	0.482	0.104	0.102	-0.016	0.001

Notes: Heteroskedasticity-robust standard errors, adjusted for clustering at the charity level, in parentheses.

*** p<0.01, ** p<0.05, * p<0.1. All models include a constant term and year, classification, and locale dummies. All independent variables are lagged by one year. Estimated using 2SLS without charity fixed effects, instrumenting for the age explanatory variables using the corresponding attributes of the population aged over 16 in the Lower Super Output Area in which each charity is based, in each year, using ONS (2020) population data. Sample includes only charities operating from a single postcode.

APPENDIX 8 — CORE MODEL: VOLUNTARY INCOME ÷ TOTAL INCOME

Age explanatory variable: Model:	Voluntary income ÷ total income										
	Proportion < 31				Standard deviation age				Mean age		
	Pooled OLS	Random Effects	Fixed Effects	Connections IV	Pooled OLS	Random Effects	Fixed Effects	Connections IV	Pooled OLS	Random Effects	Fixed Effects
Independent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
<i>Age:</i>											
Proportion < 31	0.071*	0.026	0.014	0.254							
	(0.040)	(0.020)	(0.022)	(0.159)							
Standard deviation age					-0.001*	-0.000	-0.000	0.019			
					(0.001)	(0.000)	(0.000)	(0.012)			
Mean age									-0.003***	-0.001***	-0.000
									(0.000)	(0.000)	(0.000)
<i>Gender:</i>											
Proportion female	-0.128***	-0.027***	-0.008	-0.037	-0.125***	-0.026***	-0.008	0.046	-0.141***	-0.030***	-0.009
	(0.014)	(0.008)	(0.009)	(0.064)	(0.014)	(0.008)	(0.009)	(0.103)	(0.014)	(0.008)	(0.009)
<i>Trustee controls:</i>											
Proportion non-British	0.077***	0.021	-0.001	-0.002	0.077***	0.022*	-0.001	-0.005	0.068***	0.020	-0.001
	(0.023)	(0.013)	(0.015)	(0.011)	(0.023)	(0.013)	(0.015)	(0.013)	(0.023)	(0.013)	(0.015)
Ln(Board size)	-0.001	-0.001	0.005	0.003	0.001	-0.000	0.005	-0.032	-0.001	-0.002	0.005
	(0.007)	(0.004)	(0.004)	(0.005)	(0.007)	(0.004)	(0.004)	(0.026)	(0.007)	(0.004)	(0.004)
<i>Charity controls:</i>											
Ln(Total assets)	-0.016***	-0.018***	-0.015***	-0.014***	-0.017***	-0.018***	-0.015***	-0.013***	-0.014***	-0.017***	-0.015***
	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)
Leverage	-0.203***	-0.054***	-0.038***	-0.041***	-0.202***	-0.053***	-0.038***	-0.037***	-0.212***	-0.054***	-0.038***
	(0.019)	(0.010)	(0.011)	(0.008)	(0.019)	(0.010)	(0.011)	(0.009)	(0.019)	(0.010)	(0.011)
Classification dummies	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No
Locale dummies	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No
Charity fixed effects	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes
Observations	66,277	66,277	66,277	65,442	66,277	66,277	66,277	65,442	66,277	66,277	66,277
Number of charities	8,865	8,865	8,865	8,030	8,865	8,865	8,865	8,030	8,865	8,865	8,865
R-squared	0.204	0.1945	0.015	0.010	0.204	0.1944	0.015	-0.068	0.207	0.1964	0.015

Notes: Heteroskedasticity-robust standard errors, adjusted for clustering at the charity level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All models include year dummies and a constant term. All independent variables are lagged by one year. The connections IV models, which also include charity-level fixed effects, use the proportion of male (aged 31 or over) trustees sitting on another board with a female (aged under 31) trustee as an instrument for 'proportion female' ('proportion < 31' and 'standard deviation age'). Dependent variable is constructed in line with Siciliano (1996), for comparability.

APPENDIX 9 — CONNECTIONS IV: WEAK INSTRUMENT TESTS

Endogenous variables being instrumented	Cragg-Donald Wald F statistic	Kleibergen-Papp Wald rk F statistic
Proportion < 31, proportion female	195.65	82.29
Standard deviation age, proportion female	17.13	10.21
Mean age, proportion female	5.02	3.48

Notes: Results from weak instruments tests when instrumenting for trustee age and gender using the proportion of male (aged 31 or over) trustees sitting on another board with a female (aged under 31) trustee as an instrument for 'proportion female' ('proportion < 31' and 'standard deviation age'). The Kleibergen-Papp Wald rk F statistic is robust to errors being non-i.i.d. (Baum et al., 2010). The Stock-Yogo (2005) critical value for two instruments, two endogenous regressors, and 10% maximal relative bias is 7.03.

APPENDIX 10 — REFERENCE: CORE RESULTS

Dependent variable:	Ln(Total income)					Ln(Voluntary income)					Spending ratio				
Age explanatory variable:	Proportion < 31		SD age		Mean age	Proportion < 31		SD age		Mean age	Proportion < 31		SD age		Mean age
Model:	FE	IV	FE	IV	FE	FE	IV	FE	IV	FE	FE	IV	FE	IV	FE
Independent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
<i>Age:</i>															
Proportion < 31	-0.074*	-0.367				0.194	4.847				-0.002	0.641			
	(0.043)	(0.445)				(0.412)	(3.574)				(0.029)	(0.432)			
SD age			-0.002**	-0.027				0.001	0.355				0.000	0.047	
			(0.001)	(0.034)				(0.007)	(0.270)				(0.001)	(0.034)	
Mean age					0.001					-0.005					0.001
					(0.001)					(0.005)					(0.000)
<i>Gender:</i>															
Proportion female	0.031*	0.388**	0.031*	0.268	0.034**	-0.122	-0.486	-0.120	1.094	-0.136	-0.002	0.058	-0.002	0.267	0.001
	(0.017)	(0.179)	(0.017)	(0.293)	(0.017)	(0.153)	(1.504)	(0.152)	(2.431)	(0.154)	(0.016)	(0.123)	(0.016)	(0.261)	(0.016)
<i>Trustee controls:</i>															
Proportion non-British	-0.000	-0.014	-0.001	-0.009	0.000	-0.415*	-0.451*	-0.413*	-0.513	-0.419*	0.018	0.008	0.018	-0.000	0.019
	(0.033)	(0.036)	(0.033)	(0.039)	(0.033)	(0.250)	(0.271)	(0.250)	(0.324)	(0.250)	(0.021)	(0.022)	(0.021)	(0.030)	(0.021)
Ln(Board size)	0.096***	0.086***	0.099***	0.136*	0.097***	0.386***	0.340***	0.387***	-0.330	0.382***	-0.012**	-0.023*	-0.013**	-0.111	-0.011*
	(0.010)	(0.015)	(0.010)	(0.075)	(0.010)	(0.073)	(0.114)	(0.074)	(0.590)	(0.074)	(0.006)	(0.012)	(0.006)	(0.074)	(0.006)
<i>Charity controls:</i>															
Ln(Total assets)	0.221***	0.221***	0.221***	0.220***	0.221***	0.177***	0.192***	0.177***	0.208***	0.178***	0.073***	0.075***	0.073***	0.077***	0.072***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.049)	(0.051)	(0.049)	(0.059)	(0.049)	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)
Leverage	-0.005	0.003	-0.005	-0.002	-0.005	-0.520**	-0.569**	-0.518**	-0.508**	-0.522**	-0.029	-0.035	-0.029	-0.027	-0.029
	(0.032)	(0.033)	(0.032)	(0.033)	(0.032)	(0.228)	(0.232)	(0.228)	(0.238)	(0.228)	(0.047)	(0.047)	(0.047)	(0.048)	(0.047)
Observations	66,277	65,442	66,277	65,442	66,277	66,277	65,442	66,277	65,442	66,277	66,277	65,442	66,277	65,442	66,277
Number of charities	8,865	8,030	8,865	8,030	8,865	8,865	8,030	8,865	8,030	8,865	8,865	8,030	8,865	8,030	8,865
R-squared	0.165	0.150	0.165	0.129	0.165	0.003	-0.003	0.003	-0.094	0.003	0.012	0.005	0.012	-0.095	0.012

Notes: Heteroskedasticity-robust standard errors, adjusted for clustering at the charity level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. SD = standard deviation. All models include year dummies and a constant term. All independent variables are lagged by one year. FE models include charity-level fixed effects, IV models additionally use the proportion of male (aged 31 or over) trustees sitting on another board with a female (aged under 31) trustee as an instrument for 'proportion female' ('proportion < 31' and 'standard deviation age'). These are the core results, as reported in *Tables 7-9*, presented in this form to allow straightforward comparison with subsequent robustness checks.

APPENDIX 11 — ROBUSTNESS: WINSORIZING KEY VARIABLES

Dependent variable:	Ln(Total income) [winsor]					Ln(Voluntary income) [winsor]					Spending ratio [winsor]				
Age explanatory variable:	Proportion < 31		SD age		Mean age	Proportion < 31		SD age		Mean age	Proportion < 31		SD age		Mean age
Model:	FE	IV	FE	IV	FE	FE	IV	FE	IV	FE	FE	IV	FE	IV	FE
Independent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
<i>Age:</i>															
Proportion < 31	-0.075*	-0.297				0.196	4.936				0.003	0.234			
	(0.043)	(0.444)				(0.412)	(3.578)				(0.023)	(0.218)			
SD age			-0.002**	-0.022				0.001	0.361				0.000	0.017	
			(0.001)	(0.034)				(0.007)	(0.271)				(0.000)	(0.017)	
Mean age					0.001					-0.005					0.000
					(0.001)					(0.005)					(0.000)
<i>Gender:</i>															
Proportion female	0.027	0.403**	0.027	0.306	0.030*	-0.123	-0.481	-0.121	1.133	-0.137	0.007	0.075	0.007	0.152	0.008
	(0.017)	(0.177)	(0.017)	(0.289)	(0.017)	(0.153)	(1.503)	(0.152)	(2.441)	(0.154)	(0.008)	(0.085)	(0.008)	(0.140)	(0.008)
<i>Trustee controls:</i>															
Proportion non-British	0.000	-0.016	-0.001	-0.012	0.001	-0.415*	-0.452*	-0.413*	-0.515	-0.419*	0.006	-0.000	0.006	-0.003	0.006
	(0.032)	(0.036)	(0.032)	(0.038)	(0.032)	(0.250)	(0.270)	(0.250)	(0.324)	(0.250)	(0.014)	(0.016)	(0.014)	(0.018)	(0.014)
Ln(Board size)	0.096***	0.084***	0.099***	0.125*	0.097***	0.385***	0.337***	0.385***	-0.346	0.381***	-0.009**	-0.015**	-0.010**	-0.048	-0.009**
	(0.009)	(0.015)	(0.010)	(0.074)	(0.009)	(0.073)	(0.114)	(0.074)	(0.592)	(0.074)	(0.004)	(0.007)	(0.004)	(0.037)	(0.004)
<i>Charity controls:</i>															
Ln(Total assets) [winsor]	0.222***	0.222***	0.222***	0.221***	0.222***	0.183***	0.199***	0.182***	0.215***	0.184***	0.061***	0.062***	0.061***	0.063***	0.061***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.049)	(0.052)	(0.050)	(0.060)	(0.049)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Leverage [winsor]	-0.005	0.003	-0.006	0.000	-0.005	-0.553**	-0.611**	-0.551**	-0.559**	-0.555**	-0.067***	-0.068***	-0.067***	-0.066***	-0.067***
	(0.036)	(0.036)	(0.035)	(0.036)	(0.035)	(0.254)	(0.259)	(0.254)	(0.268)	(0.254)	(0.016)	(0.017)	(0.016)	(0.017)	(0.016)
Observations	66,277	65,442	66,277	65,442	66,277	66,277	65,442	66,277	65,442	66,277	66,277	65,442	66,277	65,442	66,277
Number of charities	0.164	0.147	0.164	0.135	0.164	0.003	-0.003	0.003	-0.098	0.003	0.032	0.027	0.032	-0.027	0.032
R-squared	8,865	8,030	8,865	8,030	8,865	8,865	8,030	8,865	8,030	8,865	8,865	8,030	8,865	8,030	8,865

Notes: Heteroskedasticity-robust standard errors, adjusted for clustering at the charity level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. SD = standard deviation.

All models include year dummies and a constant term. All independent variables are lagged by one year. FE models include charity-level fixed effects, IV models additionally use the proportion of male (aged 31 or over) trustees sitting on another board with a female (aged under 31) trustee as an instrument for 'proportion female' ('proportion < 31' and 'standard deviation age'). [Winsor] indicates variables have been winsorized at 1% and 99% of their sample distributions.

APPENDIX 12 — ROBUSTNESS: STRICTER GENDER CLASSIFICATION

Dependent variable: Age explanatory variable: Model: Independent variable	Ln(Total income)					Ln(Voluntary income)					Spending ratio				
	Proportion < 31		SD age		Mean age	Proportion < 31		SD age		Mean age	Proportion < 31		SD age		Mean age
	FE	IV	FE	IV	FE	FE	IV	FE	IV	FE	FE	IV	FE	IV	FE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
<i>Age:</i>															
Proportion < 31	-0.116**	-0.608				0.015	10.021**				0.052	0.493			
	(0.053)	(0.511)				(0.554)	(4.352)				(0.034)	(0.410)			
SD age			-0.002	-0.057				0.003	0.946*				0.001	0.047	
			(0.001)	(0.056)				(0.009)	(0.547)				(0.001)	(0.044)	
Mean age					0.001					-0.005					-0.000
					(0.001)					(0.007)					(0.001)
<i>Gender:</i>															
Proportion female	0.026	0.378*	0.025	0.013	0.029	-0.189	0.193	-0.191	6.220	-0.208	0.016	-0.020	0.016	0.276	0.016
	(0.022)	(0.229)	(0.022)	(0.529)	(0.022)	(0.193)	(1.840)	(0.192)	(5.233)	(0.195)	(0.014)	(0.148)	(0.014)	(0.398)	(0.014)
<i>Trustee controls:</i>															
Proportion non-British	-0.042	-0.059	-0.043	-0.013	-0.042	-0.279	-0.433	-0.281	-1.184	-0.287	-0.000	-0.004	-0.000	-0.041	0.000
	(0.044)	(0.050)	(0.044)	(0.086)	(0.044)	(0.339)	(0.375)	(0.339)	(0.902)	(0.338)	(0.025)	(0.028)	(0.025)	(0.059)	(0.025)
Ln(Board size)	0.098***	0.090***	0.099***	0.215	0.098***	0.297***	0.138	0.292***	-1.908	0.289***	-0.017***	-0.022*	-0.018***	-0.123	-0.017**
	(0.012)	(0.018)	(0.012)	(0.135)	(0.012)	(0.091)	(0.141)	(0.092)	(1.312)	(0.092)	(0.007)	(0.012)	(0.007)	(0.105)	(0.007)
<i>Charity controls:</i>															
Ln(Total assets)	0.213***	0.213***	0.213***	0.208***	0.213***	0.111*	0.140**	0.111*	0.209**	0.112*	0.078***	0.079***	0.078***	0.083***	0.078***
	(0.013)	(0.013)	(0.013)	(0.014)	(0.013)	(0.058)	(0.062)	(0.058)	(0.104)	(0.058)	(0.007)	(0.007)	(0.007)	(0.009)	(0.007)
Leverage	0.018	0.026	0.016	0.010	0.017	-0.617**	-0.777***	-0.617**	-0.526	-0.622**	-0.097***	-0.104***	-0.096***	-0.092***	-0.096***
	(0.037)	(0.039)	(0.037)	(0.042)	(0.037)	(0.269)	(0.285)	(0.269)	(0.376)	(0.269)	(0.023)	(0.026)	(0.023)	(0.027)	(0.023)
Observations	40,881	40,282	40,881	40,282	40,881	40,881	40,282	40,881	40,282	40,881	40,881	40,282	40,881	40,282	40,881
Number of charities	5,692	5,093	5,692	5,093	5,692	5,692	5,093	5,692	5,093	5,692	5,692	5,093	5,692	5,093	5,692
R-squared	0.162	0.146	0.162	0.005	0.162	0.003	-0.022	0.003	-0.782	0.003	0.020	0.016	0.020	-0.155	0.020

Notes: Heteroskedasticity-robust standard errors, adjusted for clustering at the charity level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. SD = standard deviation. All models include year dummies and a constant term. All independent variables are lagged by one year. FE models include charity-level fixed effects, IV models additionally use the proportion of male (aged 31 or over) trustees sitting on another board with a female (aged under 31) trustee as an instrument for 'proportion female' ('proportion < 31' and 'standard deviation age'). Charities are not included if they have one or more trustees whose gender is classified based on their first name and Genderize.io reports this classification has a probability of less than 95% or is based on reference samples of under 50 names.

APPENDIX 13 — ROBUSTNESS: CHARITIES WITH TOTAL INCOME > £500,000 THROUGHOUT SAMPLE

Dependent variable:	Ln(Total income)					Ln(Voluntary income)					Spending ratio				
	Proportion < 31		SD age		Mean age	Proportion < 31		SD age		Mean age	Proportion < 31		SD age		Mean age
Age explanatory variable:	FE	IV	FE	IV	FE	FE	IV	FE	IV	FE	FE	IV	FE	IV	FE
Model:															
Independent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
<i>Age:</i>															
Proportion < 31	-0.076 (0.051)	-0.636 (0.514)				0.033 (0.496)	6.623 (4.037)				-0.020 (0.035)	0.950* (0.524)			
SD age			-0.002** (0.001)	-0.040 (0.035)				-0.006 (0.008)	0.413 (0.265)				0.000 (0.001)	0.059 (0.036)	
Mean age					0.002** (0.001)					-0.005 (0.006)					0.001 (0.001)
<i>Gender:</i>															
Proportion female	0.024 (0.018)	0.423** (0.209)	0.024 (0.018)	0.249 (0.323)	0.029 (0.018)	-0.096 (0.170)	-0.349 (1.777)	-0.093 (0.170)	1.469 (2.675)	-0.113 (0.171)	-0.008 (0.018)	0.087 (0.146)	-0.008 (0.018)	0.348 (0.299)	-0.005 (0.018)
<i>Trustee controls:</i>															
Proportion non-British	0.001 (0.037)	-0.015 (0.041)	0.000 (0.037)	-0.007 (0.045)	0.002 (0.037)	-0.387 (0.285)	-0.447 (0.311)	-0.386 (0.285)	-0.529 (0.384)	-0.393 (0.284)	0.021 (0.023)	0.005 (0.025)	0.021 (0.023)	-0.006 (0.037)	0.022 (0.024)
Ln(Board size)	0.086*** (0.010)	0.077*** (0.017)	0.088*** (0.011)	0.149** (0.075)	0.087*** (0.010)	0.361*** (0.081)	0.292** (0.125)	0.373*** (0.082)	-0.466 (0.574)	0.355*** (0.082)	-0.008 (0.007)	-0.024* (0.013)	-0.009 (0.007)	-0.132* (0.078)	-0.008 (0.007)
<i>Charity controls:</i>															
Ln(Total assets)	0.235*** (0.012)	0.235*** (0.012)	0.236*** (0.012)	0.233*** (0.013)	0.235*** (0.012)	0.214*** (0.056)	0.236*** (0.059)	0.214*** (0.056)	0.251*** (0.068)	0.216*** (0.056)	0.062*** (0.006)	0.066*** (0.006)	0.062*** (0.006)	0.068*** (0.007)	0.062*** (0.006)
Leverage	-0.015 (0.037)	-0.007 (0.037)	-0.015 (0.036)	-0.009 (0.038)	-0.014 (0.036)	-0.437* (0.250)	-0.501** (0.255)	-0.436* (0.250)	-0.483* (0.270)	-0.441* (0.250)	-0.016 (0.056)	-0.025 (0.057)	-0.016 (0.056)	-0.022 (0.059)	-0.016 (0.056)
Observations	54,081	53,943	54,081	53,943	54,081	54,081	53,943	54,081	53,943	54,081	54,081	53,943	54,081	53,943	54,081
Number of charities	5,773	5,635	5,773	5,635	5,773	5,773	5,635	5,773	5,635	5,773	5,773	5,635	5,773	5,635	5,773
R-squared	0.200	0.177	0.200	0.124	0.200	0.003	-0.007	0.003	-0.130	0.003	0.008	-0.006	0.008	-0.154	0.008

Notes: Heteroskedasticity-robust standard errors, adjusted for clustering at the charity level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. SD = standard deviation. All models include year dummies and a constant term. All independent variables are lagged by one year. FE models include charity-level fixed effects, IV models additionally use the proportion of male (aged 31 or over) trustees sitting on another board with a female (aged under 31) trustee as an instrument for 'proportion female' ('proportion < 31' and 'standard deviation age'). Charities are not included if they are active but have total annual income of under £500,000 in any year 2007-18. Because detailed financial information is only reported in years when a charity's income exceeds this threshold, they may enter and exit the main sample in a non-random fashion.

APPENDIX 14 — SUBSAMPLE: HELPS CHILDREN/YOUNG PEOPLE

Dependent variable:	Ln(Total income)					Ln(Voluntary income)					Spending ratio				
	Proportion < 31		SD age		Mean age	Proportion < 31		SD age		Mean age	Proportion < 31		SD age		Mean age
Age explanatory variable:	FE	IV	FE	IV	FE	FE	IV	FE	IV	FE	FE	IV	FE	IV	FE
Model:															
Independent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
<i>Age:</i>															
Proportion < 31	-0.070 (0.058)	0.090 (0.527)				0.518 (0.550)	4.330 (4.332)				-0.021 (0.039)	0.504 (0.429)			
SD age			-0.002** (0.001)	0.007 (0.041)				0.007 (0.010)	0.342 (0.348)				0.000 (0.001)	0.040 (0.038)	
Mean age					0.001* (0.001)					-0.007 (0.008)					-0.000 (0.001)
<i>Gender:</i>															
Proportion female	0.039* (0.021)	0.447** (0.219)	0.039* (0.021)	0.472 (0.319)	0.043** (0.021)	0.039 (0.202)	-0.901 (1.885)	0.043 (0.201)	0.276 (2.774)	0.024 (0.205)	0.009 (0.013)	0.075 (0.125)	0.009 (0.013)	0.212 (0.239)	0.008 (0.013)
<i>Trustee controls:</i>															
Proportion non-British	-0.026 (0.045)	-0.047 (0.050)	-0.026 (0.045)	-0.049 (0.056)	-0.025 (0.045)	-0.185 (0.340)	-0.204 (0.368)	-0.180 (0.340)	-0.321 (0.466)	-0.186 (0.340)	-0.000 (0.025)	-0.012 (0.028)	-0.001 (0.025)	-0.025 (0.042)	-0.001 (0.025)
Ln(Board size)	0.096*** (0.011)	0.077*** (0.017)	0.099*** (0.012)	0.064 (0.088)	0.097*** (0.011)	0.335*** (0.093)	0.323** (0.141)	0.328*** (0.095)	-0.309 (0.737)	0.332*** (0.095)	-0.012** (0.006)	-0.022** (0.011)	-0.013** (0.006)	-0.095 (0.078)	-0.013** (0.006)
<i>Charity controls:</i>															
Ln(Total assets)	0.232*** (0.015)	0.233*** (0.015)	0.232*** (0.015)	0.233*** (0.015)	0.232*** (0.015)	0.177*** (0.067)	0.193*** (0.070)	0.175*** (0.067)	0.193*** (0.074)	0.177*** (0.067)	0.065*** (0.006)	0.068*** (0.006)	0.066*** (0.006)	0.068*** (0.007)	0.066*** (0.006)
Leverage	-0.032 (0.041)	-0.026 (0.042)	-0.033 (0.041)	-0.025 (0.042)	-0.032 (0.041)	-0.698** (0.303)	-0.749** (0.307)	-0.694** (0.303)	-0.712** (0.314)	-0.699** (0.303)	-0.030* (0.018)	-0.033* (0.019)	-0.030* (0.018)	-0.029 (0.022)	-0.030* (0.018)
Observations	38,741	38,324	38,741	38,324	38,741	38,741	38,324	38,741	38,324	38,741	38,741	38,324	38,741	38,324	38,741
Number of charities	5,038	4,621	5,038	4,621	5,038	5,038	4,621	5,038	4,621	5,038	5,038	4,621	5,038	4,621	5,038
R-squared	0.173	0.153	0.173	0.146	0.173	0.003	-0.002	0.003	-0.078	0.003	0.024	0.011	0.024	-0.158	0.024

Notes: Heteroskedasticity-robust standard errors, adjusted for clustering at the charity level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. SD = standard deviation.

All models include year dummies and a constant term. All independent variables are lagged by one year. FE models include charity-level fixed effects, IV models additionally use the proportion of male (aged 31 or over) trustees sitting on another board with a female (aged under 31) trustee as an instrument for 'proportion female' ('proportion < 31' and 'standard deviation age'). Sub-sample only includes charities that help children / young people, as per the Charity Register (Charity Commission, 2020a).

APPENDIX 15 — SUBSAMPLE: ADVANCES HUMAN RIGHTS, RELIGIOUS/RACIAL HARMONY, OR EQUALITY AND DIVERSITY

Dependent variable:	Ln(Total income)					Ln(Voluntary income)					Spending ratio				
	Proportion < 31		SD age		Mean age	Proportion < 31		SD age		Mean age	Proportion < 31		SD age		Mean age
Age explanatory variable:	FE	IV	FE	IV	FE	FE	IV	FE	IV	FE	FE	IV	FE	IV	FE
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Independent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
<i>Age:</i>															
Proportion < 31	-0.049 (0.173)	0.996 (1.153)				0.180 (0.872)	-2.899 (7.087)				-0.050 (0.086)	2.257 (2.397)			
SD age			-0.007 (0.004)	0.113 (0.161)				0.043 (0.027)	-0.330 (0.879)				0.005** (0.003)	0.257 (0.352)	
Mean age					-0.000 (0.003)					0.039* (0.021)					0.004 (0.004)
<i>Gender:</i>															
Proportion female	0.037 (0.091)	0.571 (0.597)	0.041 (0.092)	1.423 (1.682)	0.035 (0.091)	0.194 (0.586)	4.410 (4.944)	0.169 (0.590)	1.930 (9.293)	0.349 (0.595)	0.002 (0.064)	0.591 (0.725)	-0.001 (0.064)	2.522 (3.294)	0.017 (0.074)
<i>Trustee controls:</i>															
Proportion non-British	0.014 (0.139)	-0.012 (0.144)	0.014 (0.140)	-0.055 (0.236)	0.014 (0.139)	0.055 (0.719)	-0.123 (0.831)	0.054 (0.726)	0.003 (0.930)	0.019 (0.732)	0.204 (0.225)	0.173 (0.205)	0.203 (0.224)	0.075 (0.418)	0.200 (0.221)
Ln(Board size)	0.170*** (0.043)	0.126** (0.063)	0.187*** (0.046)	-0.178 (0.467)	0.169*** (0.043)	0.185 (0.295)	0.012 (0.428)	0.085 (0.294)	0.897 (2.554)	0.224 (0.300)	-0.096** (0.048)	-0.163 (0.112)	-0.110** (0.050)	-0.852 (1.024)	-0.093** (0.045)
<i>Charity controls:</i>															
Ln(Total assets)	0.228*** (0.043)	0.236*** (0.044)	0.227*** (0.043)	0.256*** (0.064)	0.229*** (0.044)	0.246 (0.182)	0.233 (0.197)	0.255 (0.182)	0.175 (0.301)	0.233 (0.182)	0.103*** (0.017)	0.119*** (0.030)	0.105*** (0.016)	0.164* (0.099)	0.102*** (0.016)
Leverage	-0.050 (0.140)	-0.131 (0.226)	-0.070 (0.138)	0.209 (0.468)	-0.055 (0.142)	0.144 (1.010)	0.462 (1.088)	0.251 (1.038)	-0.525 (2.398)	0.287 (0.966)	0.055 (0.113)	-0.131 (0.262)	0.063 (0.114)	0.638 (0.940)	0.063 (0.113)
Observations	2,830	2,781	2,830	2,781	2,830	2,830	2,781	2,830	2,781	2,830	2,830	2,781	2,830	2,781	2,830
Number of charities	426	377	426	377	426	426	377	426	377	426	426	377	426	377	426
R-squared	0.156	0.106	0.158	-0.709	0.156	0.009	-0.030	0.011	-0.141	0.012	0.030	-0.126	0.032	-3.682	0.031

Notes: Heteroskedasticity-robust standard errors, adjusted for clustering at the charity level, in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. SD = standard deviation. All models include year dummies and a constant term. All independent variables are lagged by one year. FE models include charity-level fixed effects, IV models additionally use the proportion of male (aged 31 or over) trustees sitting on another board with a female (aged under 31) trustee as an instrument for 'proportion female' ('proportion < 31' and 'standard deviation age'). Sub-sample includes only charities which, as per the Charity Register (Charity Commission, 2020a), have a charitable purpose of "the advancement of human rights, conflict resolution or reconciliation or the promotion of religious or racial harmony or equality and diversity".

APPENDIX 16 — ROBUSTNESS: CONNECTIONS IV, CONTROLLING FOR TOTAL CONNECTIONS

Dependent variable: Age explanatory variable: Model: Independent variable	Ln(Total income)				Ln(Voluntary income)				Spending ratio			
	Proportion < 31		SD age		Proportion < 31		SD age		Proportion < 31		SD age	
	IV	IV TC	IV	IV TC	IV	IV TC	IV	IV TC	IV	IV TC	IV	IV TC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
<i>Age:</i>												
Proportion < 31	-0.367 (0.445)	-0.629 (0.405)			4.847 (3.574)	5.420* (3.292)			0.641 (0.432)	0.612 (0.397)		
SD age			-0.027 (0.034)	-0.038 (0.026)			0.355 (0.270)	0.328 (0.202)			0.047 (0.034)	0.037 (0.025)
<i>Mean age</i>												
<i>Gender:</i>												
Proportion female	0.388** (0.179)	0.183 (0.134)	0.268 (0.293)	0.100 (0.172)	-0.486 (1.504)	-0.037 (1.158)	1.094 (2.431)	0.686 (1.437)	0.058 (0.123)	0.035 (0.101)	0.267 (0.261)	0.117 (0.143)
<i>Trustee controls:</i>												
Proportion non-British	-0.014 (0.036)	0.001 (0.035)	-0.009 (0.039)	0.002 (0.037)	-0.451* (0.271)	-0.485* (0.262)	-0.513 (0.324)	-0.486* (0.292)	0.008 (0.022)	0.010 (0.021)	-0.000 (0.030)	0.010 (0.025)
Ln(Board size)	0.086*** (0.015)	0.090*** (0.014)	0.136* (0.075)	0.160*** (0.059)	0.340*** (0.114)	0.331*** (0.108)	-0.330 (0.590)	-0.273 (0.443)	-0.023* (0.012)	-0.022** (0.011)	-0.111 (0.074)	-0.091 (0.055)
Total connections		0.000*** (0.000)		0.000 (0.000)		-0.000 (0.000)		0.000 (0.001)		0.000 (0.000)		0.000 (0.000)
<i>Charity controls:</i>												
Ln(Total assets)	0.221*** (0.011)	0.219*** (0.011)	0.220*** (0.011)	0.218*** (0.011)	0.192*** (0.051)	0.195*** (0.050)	0.208*** (0.059)	0.205*** (0.055)	0.075*** (0.005)	0.075*** (0.005)	0.077*** (0.006)	0.076*** (0.006)
Leverage	0.003 (0.033)	0.002 (0.033)	-0.002 (0.033)	-0.004 (0.034)	-0.569** (0.232)	-0.566** (0.232)	-0.508** (0.238)	-0.515** (0.236)	-0.035 (0.047)	-0.035 (0.047)	-0.027 (0.048)	-0.029 (0.047)
Observations	65,442	65,442	65,442	65,442	65,442	65,442	65,442	65,442	65,442	65,442	65,442	65,442
Number of charities	8,030	8,030	8,030	8,030	8,030	8,030	8,030	8,030	8,030	8,030	8,030	8,030
R-squared	0.150	0.158	0.129	0.101	-0.003	-0.004	-0.094	-0.078	0.005	0.005	-0.095	-0.051

Notes: Heteroskedasticity-robust standard errors, adjusted for clustering at the charity level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. SD = standard deviation. All models include year dummies and a constant term. All independent variables are lagged by one year. All models use the proportion of male (aged 31 or over) trustees sitting on another board with a female (aged under 31) trustee as an instrument for 'proportion female' ('proportion < 31' and 'standard deviation age'). IV TC columns add an additional control of the total number of external directorships trustees on a board hold.

APPENDIX 17 — ROBUSTNESS: INTERACTION TERMS

Dependent variable: Age explanatory variable: Model: Independent variable	Ln(Total income)					Ln(Voluntary income)					Spending ratio				
	Proportion < 31		SD age		Mean age	Proportion < 31		SD age		Mean age	Proportion < 31		SD age		Mean age
	FE	IV	FE	IV	FE	FE	IV	FE	IV	FE	FE	IV	FE	IV	FE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
<i>Age:</i>															
Proportion < 31	-0.079 (0.073)	0.363 (0.820)				-0.275 (0.791)	4.236 (6.516)				-0.104* (0.061)	0.109 (0.643)			
SD age			-0.002 (0.001)	0.008 (0.081)				-0.013 (0.012)	0.585 (0.686)				-0.001 (0.001)	0.043 (0.073)	
Mean age					0.002* (0.001)					0.001 (0.008)					0.001 (0.001)
<i>Gender:</i>															
Proportion female	0.031* (0.017)	0.448** (0.193)	0.035 (0.033)	0.996 (1.350)	0.170 (0.107)	-0.146 (0.156)	-0.537 (1.597)	-0.467 (0.302)	5.850 (11.703)	0.813 (0.958)	-0.007 (0.017)	0.014 (0.130)	-0.034 (0.035)	0.179 (1.164)	0.031 (0.084)
<i>Interaction terms:</i>															
(Prop < 31)×(Prop fem)	0.012 (0.117)	-1.867 (1.626)				1.079 (1.513)	1.564 (13.761)				0.236** (0.102)	1.361 (1.241)			
(SD age)×(Prop fem)			-0.000 (0.003)	-0.076 (0.124)				0.037 (0.027)	-0.498 (1.094)				0.003 (0.003)	0.009 (0.106)	
(Mean age)×(Prop fem)					-0.002 (0.002)					-0.017 (0.017)					-0.001 (0.002)
<i>Trustee controls:</i>															
Proportion non-British	-0.000 (0.033)	-0.018 (0.037)	-0.001 (0.033)	-0.013 (0.041)	0.000 (0.033)	-0.413* (0.250)	-0.448* (0.272)	-0.412* (0.250)	-0.535 (0.350)	-0.420* (0.250)	0.019 (0.021)	0.011 (0.023)	0.018 (0.021)	0.000 (0.032)	0.019 (0.021)
Ln(Board size)	0.096*** (0.010)	0.086*** (0.015)	0.099*** (0.010)	0.115 (0.098)	0.097*** (0.010)	0.386*** (0.074)	0.340*** (0.114)	0.390*** (0.074)	-0.467 (0.803)	0.382*** (0.074)	-0.012** (0.006)	-0.023* (0.012)	-0.012** (0.006)	-0.109 (0.093)	-0.011* (0.006)
<i>Charity controls:</i>															
Ln(Total assets)	0.221*** (0.011)	0.221*** (0.011)	0.221*** (0.011)	0.220*** (0.011)	0.221*** (0.011)	0.177*** (0.049)	0.192*** (0.051)	0.177*** (0.049)	0.209*** (0.062)	0.178*** (0.049)	0.072*** (0.005)	0.075*** (0.005)	0.073*** (0.005)	0.077*** (0.006)	0.072*** (0.005)
Leverage	-0.005 (0.032)	0.006 (0.034)	-0.005 (0.032)	0.010 (0.039)	-0.005 (0.032)	-0.521** (0.228)	-0.571** (0.233)	-0.524** (0.228)	-0.431 (0.304)	-0.522** (0.228)	-0.030 (0.047)	-0.036 (0.047)	-0.030 (0.047)	-0.028 (0.051)	-0.029 (0.047)
Observations	66,277	65,442	66,277	65,442	66,277	66,277	65,442	66,277	65,442	66,277	66,277	65,442	66,277	65,442	66,277
Number of charities	8,865	8,030	8,865	8,030	8,865	8,865	8,030	8,865	8,030	8,865	8,865	8,030	8,865	8,030	8,865
R-squared	0.165	0.142	0.165	0.124	0.165	0.003	-0.003	0.003	-0.140	0.003	0.012	0.002	0.012	-0.091	0.012

Notes: Heteroskedasticity-robust standard errors, adjusted for clustering at the charity level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. SD = standard deviation. All models include year dummies and a constant term. All independent variables are lagged by one year. FE models include charity-level fixed effects, IV models additionally use the proportion of male (aged 31 or over) trustees sitting on another board with a female (aged under 31) trustee as an instrument for 'proportion female' ('proportion < 31' and 'standard deviation age').

APPENDIX 18 — ROBUSTNESS: BLAU INDEX OF GENDER DIVERSITY

Dependent variable: Age explanatory variable: Model: Independent variable	Ln(Total income)			Ln(Voluntary income)			Spending ratio		
	Proportion < 31	SD age	Mean age	Proportion < 31	SD age	Mean age	Proportion < 31	SD age	Mean age
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
FE									
<i>Age:</i>									
Proportion < 31	-0.073* (0.043)			0.182 (0.411)			-0.000 (0.029)		
SD age		-0.002** (0.001)			0.001 (0.007)			0.000 (0.001)	
Mean age			0.001 (0.001)			-0.004 (0.005)			0.001 (0.000)
<i>Gender:</i>									
Blau index	0.033 (0.022)	0.034 (0.022)	0.035 (0.022)	0.010 (0.178)	0.012 (0.178)	0.001 (0.178)	-0.026 (0.030)	-0.026 (0.030)	-0.024 (0.030)
<i>Trustee controls:</i>									
Proportion non-British	0.000 (0.033)	-0.000 (0.033)	0.001 (0.033)	-0.421* (0.250)	-0.419* (0.250)	-0.425* (0.250)	0.019 (0.021)	0.019 (0.021)	0.020 (0.021)
Ln(Board size)	0.095*** (0.010)	0.097*** (0.010)	0.095*** (0.010)	0.380*** (0.075)	0.381*** (0.076)	0.377*** (0.076)	-0.010 (0.007)	-0.011 (0.007)	-0.009 (0.007)
<i>Charity controls:</i>									
Ln(Total assets)	0.221*** (0.011)	0.221*** (0.011)	0.221*** (0.011)	0.177*** (0.049)	0.177*** (0.049)	0.178*** (0.049)	0.072*** (0.005)	0.072*** (0.005)	0.072*** (0.005)
Leverage	-0.005 (0.032)	-0.006 (0.032)	-0.005 (0.032)	-0.518** (0.228)	-0.516** (0.228)	-0.519** (0.228)	-0.029 (0.047)	-0.029 (0.047)	-0.029 (0.047)
Observations	66,277	66,277	66,277	66,277	66,277	66,277	66,277	66,277	66,277
Number of charities	8,865	8,865	8,865	8,865	8,865	8,865	8,865	8,865	8,865
R-squared	0.165	0.165	0.165	0.003	0.003	0.003	0.012	0.012	0.012

Notes: Heteroskedasticity-robust standard errors, adjusted for clustering at the charity level, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. SD = standard deviation. All models include year dummies, a constant term and charity fixed effects. The Blau index, also known as the Gini-Simpson index, is a diversity index equal to $1 - \sum_{i=1}^K P_i^2$ where K is the number of gender categories present in a charity and P_i the proportion of trustees in each (Shehata et al., 2017).