

Online appendix for Webber, M. et al. The implementation of Connecting People in community mental health teams in England: A quasi-experimental study.

APPENDIX A. Outcome evaluation additional analysis

Post-estimation checks and sensitivity analyses for the primary outcome analysis

Post-estimation checks and sensitivity analyses were performed to evaluate the robustness of the results. Neither the visual inspection of the three continuous variables in the regression model nor augmented component-plus-residual plots with fitted splines pointed to potential non-linear relationships. Plotting residuals versus fitted values and added variable plots did not show any trends. The dependent variable was identified to be right-skewed and the regression model was bootstrapped to rely less on assumptions of normality and symmetry for the estimation of confidence intervals for the regression coefficients (see last column in table 3). Although this leads to more precise estimation of the group effect (smaller standard error for the regression coefficient), the conclusion from the model remains the same. Visual inspection and leverage plots pointed to a number of potential outlying cases. Outlying cases were mainly observed at pre- and post-test and more likely to be from teams which did not engage with the implementation processes.

The LOCF analysis was also repeated as a difference-in-difference regression with age, gender and team as control variables, effects for group and pre- versus post-assessment and an interaction effect between the latter two variables, which estimates the differential change between the groups over the assessment period. Again, the implementation effect remained non-significant ($p=0.82$ for model with control variables; $p=0.83$ without controls) and the effect size was very small and estimated as partial $r=-0.03$ (bivariate $r=-0.01$).

Results for secondary outcomes

The analyses of the secondary outcomes followed the same approach as presented for the primary outcome in the main paper. These analyses were pre-planned, but the study was only powered to evaluate the primary outcome robustly (i.e. with pre-specified false-positive and false-negative discovery rates). The following analyses are therefore exploratory, and the statistical significance levels are purely descriptive and cannot be used to evaluate the (non-)effectiveness of the implementation. These analyses provide potential pointers to future research and especially negative effects that would have been found could inform future studies and systematic reviews to potentially take such unwanted outcomes or unintended consequences into account. For the secondary outcomes and their references please refer to the section on “Measures” in the main manuscript.

For all regression models the same post-estimation tests and sensitivity analyses were performed as for the primary outcome and again only minor violations (especially of normality and homoscedasticity) were found, which were reasonably addressed using the bootstrap approach as a standard additional analysis.

Table A1 presents the findings on goal attainment scaling, which shows that the intervention did not have a statistically significant effect on the GAS score. Although this result is consistent across the four analyses, there is for this measure a difference between the LOCF and observed data analyses. While for the LOCF analysis the effect is very small (partial $r=0.01$) and in the wrong direction (higher scores in the control group), in the observed data the effect is in the correct direction (lower scores in the control group), and it is larger (partial $r=-0.18$). Nevertheless, both results are non-significant and given the exploratory nature of these analyses likely to be due to variation around a null effect.

Table A1. Ordinary least squares regression for GAS scores

	Group-only Model	LOCF Model^a	Observed Data^a	LOCF Model^a, bootstapped^b
Baseline GAS		1.96** (0.69)	0.60 (0.67)	1.96** (0.67)
Gender		-1.18 (1.91)	-2.35 (1.71)	-1.18 (1.98)
Age		-0.11 (0.07)	0.01 (0.06)	-0.11 (0.07)
Group (control)	0.36 (1.77)	0.13 (4.50)	-2.71 (3.53)	0.13 (3.86)
Constant	50.69*** (1.37)	-15.75 (27.70)	36.87 (26.33)	-15.75 (26.85)
Observations	151	149	100	149
R-squared	0.00	0.15	0.08	0.15

Notes. Standard errors in brackets. The dependent variable for all regressions is the final Goal Attainment Scale score with the last observation carried forward, apart from the observed data model (only observed final scores).

^aRegression controlled for delivery teams (seven dummy variables; teams BI and HI were merged due to their small sample sizes; teams that recruited no participants were excluded). ^b500 bootstrap samples. LOCF, last observation carried forward.

*** p<0.001, ** p<0.01, * p<0.05

Table A2 presents the results for the QPR score. Consistently across all four analyses the control group shows lower values in the outcome variable, but the effect sizes are small, with a partial $r=-0.09$ when controlling for other variables (bivariate $r=-0.07$). Therefore, even if this were not pointing to a null effect, the size of the effect is likely to be too small to meet the threshold for practical relevance (which was set at a partial correlation of $r=0.20$ for the primary outcome, for example).

Table A2. Ordinary least squares regression for QPR scores

	Group-only Model	LOCF Model^a	Observed Data^a	LOCF Model^a, bootstapped^b
Baseline QPR		0.81***	0.77***	0.81***
		(0.06)	(0.07)	(0.05)
Gender		0.07	0.52	0.07
		(1.55)	(1.83)	(1.56)
Age		-0.05	-0.07	-0.05
		(0.06)	(0.06)	(0.05)
Group (control)	-1.75	-2.35	-3.31	-2.35
	(2.11)	(3.66)	(4.10)	(4.78)
Constant	50.20***	16.77**	20.17***	16.77**
	(1.64)	(4.74)	(5.59)	(5.19)
Observations	151	149	125	149
R-squared	0.00	0.60	0.56	0.60

Note. Standard errors in brackets. The dependent variable for all regressions is the final [QPR] score with the last observation carried forward, apart from the observed data model (only observed final scores). ^aRegression controlled for delivery teams (seven dummy variables; teams BI and HI were merged due to their small sample sizes; teams that recruited no participants were excluded). ^b500 bootstrap samples. LOCF, last observation carried forward.

*** $p<0.001$, ** $p<0.01$, * $p<0.05$

Table A3 presents the results for the WEMWBS score. Across all analyses including the specified control variables the control group shows higher values in the WEMWBS. The effect size is close to practical relevance as defined for the primary outcome (partial $r=0.18$ when controlling for other variables). The average expected score gain in the control group is estimated at about 3.5 WEMWBS points, which is about half the size of potentially relevant differences identified by Maheswaran et al (2012). Being non-significant throughout and below the threshold for practical relevance makes it likely that this points to a null effect (rather than an adverse effect of the implementation of Connecting People).

Table A3. Ordinary least squares regression for WEMWBS scores

	Group-only Model	LOCF Model ^a	Observed Data ^a	LOCF Model ^a , bootstapped ^b
Baseline WEMWBS		0.72***	0.65***	0.72***
		(0.07)	(0.08)	(0.07)
Gender		0.57	0.65	0.57
		(1.35)	(1.59)	(1.33)
Age		-0.02	-0.04	-0.02
		(0.05)	(0.06)	(0.05)
Group (control)	-1.09	3.54	3.41	3.54
	(1.60)	(3.17)	(3.54)	(3.00)
Constant	38.42***	9.80*	13.30*	9.80**
	(1.24)	(4.28)	(5.22)	(3.69)
Observations	151	149	125	149
R-squared	0.00	0.49	0.40	0.49

Note. Standard errors in brackets. The dependent variable for all regressions is the final WEMWBS score with the last observation carried forward, apart from the observed data model (only observed final scores).

^aRegression controlled for delivery teams (seven dummy variables; teams BI and HI were merged due to their small sample sizes; teams that recruited no participants were excluded). ^b500 bootstrap samples. LOCF, last observation carried forward.

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table A4 presents the results for the EQ-5D index value. Consistently across all four analyses, the coefficient for the intervention is not significant and associated with a very small effect size, with all (partial) $r < 0.04$ and the expected group difference is also below the threshold for practical relevance. Based on these results it is likely that there was no effect on health-related quality of life.

Since the distribution of the index values was non-normal and is also limited to a small meaningful range, additional analyses were conducted. The predicted values from the LOCF model were in the range of $min = -0.15$ and $max = 0.97$, which indicates that in terms of conditional expectations the OLS approach was performing acceptably as a descriptive summary. We ran in addition a beta regression (as outlined by Hunger et al, (2011) and Smithson & Verkuilen (2006)), predicting transformed (0-1 range) index values, which led to a similar result: the coefficient for the intervention was still non-significant ($b = 0.15$, $SE = 0.28$, $p = 0.60$).

Table A4. Ordinary least squares regression for EQ-5D index values

	Group-only Model	LOCF Model ^a	Observed Data ^a	LOCF Model ^a , bootstrapped ^b
Baseline EQ-5D index value		0.76***	0.72***	0.76***
		(0.05)	(0.06)	(0.06)
Gender		-0.01	-0.00	-0.01
		(0.03)	(0.04)	(0.03)
Age		0.00*	0.00*	0.00*
		(0.00)	(0.00)	(0.00)
Group	0.02	0.01	0.00	0.01
	(0.05)	(0.08)	(0.09)	(0.09)
Constant	0.52***	0.28**	0.33**	0.28**
	(0.04)	(0.09)	(0.11)	(0.10)
Observations	151	148	124	148
R-squared	0.00	0.67	0.62	0.67

Note. Standard errors in brackets. The dependent variable for all regressions is the final EQ-5D index score with the last observation carried forward, apart from the observed data model (only observed final scores).

^aRegression controlled for delivery teams (seven dummy variables; teams BI and HI were merged due to their small sample sizes; teams that recruited no participants were excluded). ^b500 bootstrap samples. LOCF, last observation carried forward.

*** p<0.001, ** p<0.01, * p<0.05

Table A5 presents the results for the EQ-5D-VAS score, which was the only not pre-planned analysis and is reported for comparability with other research using this outcome of the EQ-5D-3L. Consistently across all three analyses including the control variables the control group showed lower values in the outcome variable, but again this difference was statistically not significant. The effect sizes with a partial $r=-0.29$ when controlling for other variables in the LOCF analysis was above the relevance threshold we used for the primary outcome. The consistency of the result and the size of the effect would support further qualitative exploration. If a plausible theoretical link could be established connecting the implementation of Connecting People with self-rated health and this result were replicated in a quantitative, the EQ-5D-VAS could be considered as an outcome in future evaluations of Connecting People and potentially similar interventions.

Table A5. OLS regression for EQ-5D-VAS

	Group-only Model	LOCF Model^a	Observed Data^a	LOCF Model^a, bootstapped^b
Baseline EQ-5D-VAS		0.65*** (0.07)	0.60*** (0.08)	0.65*** (0.08)
Gender		-3.81 (3.41)	-4.22 (4.01)	-3.81 (3.11)
Age		-0.18 (0.12)	-0.24 (0.14)	-0.18 (0.14)
Group (control)	2.22 (3.91)	-13.63 (8.08)	-16.09 (9.06)	-13.63 (7.27)
Constant	51.93*** (3.04)	44.96*** (9.32)	52.99*** (10.88)	44.96*** (10.18)
Observations	151	149	125	149
R-squared	0.00	0.45	0.42	0.45

Note. Standard errors in brackets. The dependent variable for all regressions is the final EQ-5D-VAS score with the last observation carried forward, apart from the observed data model (only observed final scores).

^aRegression controlled for delivery teams (seven dummy variables; teams BI and HI were merged due to their small sample sizes; teams that recruited no participants were excluded). ^b500 bootstrap samples. LOCF, last observation carried forward.

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Appendix B. Economic evaluation additional analysis

Table B1: Unit costs

Item	Cost (£)	Unit	Source
GP	4	per minute	Curtis & Burns (2018), table 10.3b
Psychiatrist	5.77	per minute	Curtis & Burns (2018), supplementary table 1
Other doctor	77.21	per hour	Curtis & Burns (2018), table 14 (Average)
Psychologist	53.74	per hour	Curtis & Burns (2018), table 9- assumed band 7
Counsellor	91	per hour	Curtis & Burns (2018), table 6.15. Assumed equivalent cost to child specialist
Crisis team member	40.56	per hour	Curtis & Burns (2017), table 12.3
Assertive outreach team member	23	per hour	Curtis & Burns (2018), table 11.7
Early intervention team member	40.56	per hour	Curtis & Burns (2017), table 12.5
Social worker	84	per hour	Curtis & Burns (2018), table 11.1
Mental health nurse	65.22	per hour	Curtis & Burns (2018), supplementary table 1 (Specialist nurse)
Occupational therapist	47	per hour	Curtis & Burns (2018), table 11.4
drug/alcohol service / advisor	118	per contact	Curtis & Burns (2018), table 3.1, average
drop-in centre	39.55	per hour	Curtis & Burns (2017), table 12.2. Assumed one-to-one contacts.
self-help/support group	5.80	per session	D'Amico et al. (2015)
class/group at a leisure centre	10.86	per session	D'Amico et al. (2015)
adult education class	5.80	per session	D'Amico et al. (2015)
Generic inpatient stay	628.1	per day	Department of Health (2017), weighted average of non-paediatric elective and non-elective inlier and excess bed-days.

Item	Cost (£)	Unit	Source
Mental health inpatient stay	420.0 3	per day	Department of Health (2017), weighted average of mental health cluster bed days
Medication	0.01 - 6.53	Range of unit costs	NHS Digital (2018)
police officer contact	48	per hour	Beecham and Bonin (2019)
probation officer	223	per hour face to face	Brookes et al. (2013)
court attendance	448.6 7	per hour	The Law Society (2018), and assuming a 6-hour average operating day.
Solicitor	152.5 5	per hour	Phelan (2016)
police cell	229	per bed day	Brookes et al. (2013)
Supported housing (LA, NHS)	148.3 1	per week	Curtis (2010), table 1.8
High intensity supported housing (LA,NHS)	231.5 9	per week	Curtis (2010), table 1.10
Supported housing (voluntary sector/private)	172.2 7	per week	Curtis (2010), table 1.9
High intensity supported housing (voluntary sector/private)	420.9 7	per week	Curtis (2010), table 1.11
Unpaid care/lost employment minimum wage	7.5	per hour	Department for Business, Energy & Industrial Strategy (2017)
Unpaid care replacement cost valuation of a home care worker	22.00	per hour	Curtis & Burns (2018), table 11.5
lost employment-average pay of workers	555	per week	ONS (2018), table 20.1a

Item	Cost (£)	Unit	Source
lost employment-average pay of full-time workers	685.3	per week	ONS (2018), table 20.1a
lost employment-average pay of part time workers	227.5	per week	ONS (2018), table 20.1a
Travel- cost of running a car	0.05	Per mile	Calculated as £388.45 per month (Kwik Fit, 2018) divided by 7900 miles per month (Bancroft, 2017)

Note: Unit costs for earlier years were uprated to 2017/18 prices by uprating to 2016/17 prices using the hospital & community health services (HCHS) index and from 2016/17 to 2017/18 using the Consumer Price Index (health) as described in Unit costs of Health and Social Care (2018) table 15.3

Table B2: Service utilisation for community and hospital contacts over 6 months for the Economic evaluation subsample

SERVICE	Control mean frequency (N=77)		Implementation group mean frequency (N=49)	
	Baseline	Follow-up	Baseline	Follow-up
General Practitioner	5.19	4.16	5.02	3.45
Psychiatrist	3.27	1.52	2.47	1.61
Other doctor	0.91	0.86	1.59	1.35
Psychologist	2.04	2.35	1.90	1.18
Drug & Alcohol advisor	0.92	0.13	0.69	0.12
Counsellor	2.39	3.40	4.36	3.37
Crisis team member	1.94	0.58	1.67	1.02
Assertive outreach team member	0.05	0	0	0.04
Early intervention team member	0	0	0.06	0
Social worker	3.71	4.18	5.24	3.88
Mental health Nurse	6.91	7.14	5.06	3.98
Occupational therapist	1.38	0.16	0.10	0.22
Drug/alcohol service	0.55	0.05	0.53	0.45
Drop-in centre	1.53	0.55	0.78	0
Support group	4.60	7.18	1.76	6.84
Class at a leisure centre	2.99	5.71	4.08	5

SERVICE	Control mean frequency (N=77)		Implementation group mean frequency (N=49)	
Adult education class	1.62	1.48	0.59	1.06
Inpatient bed day	3.84	2.07	4.01	2.31

Figure B1 Bootstrapped mean cost versus mean outcome: QALY (EQ-5D), HSC perspective

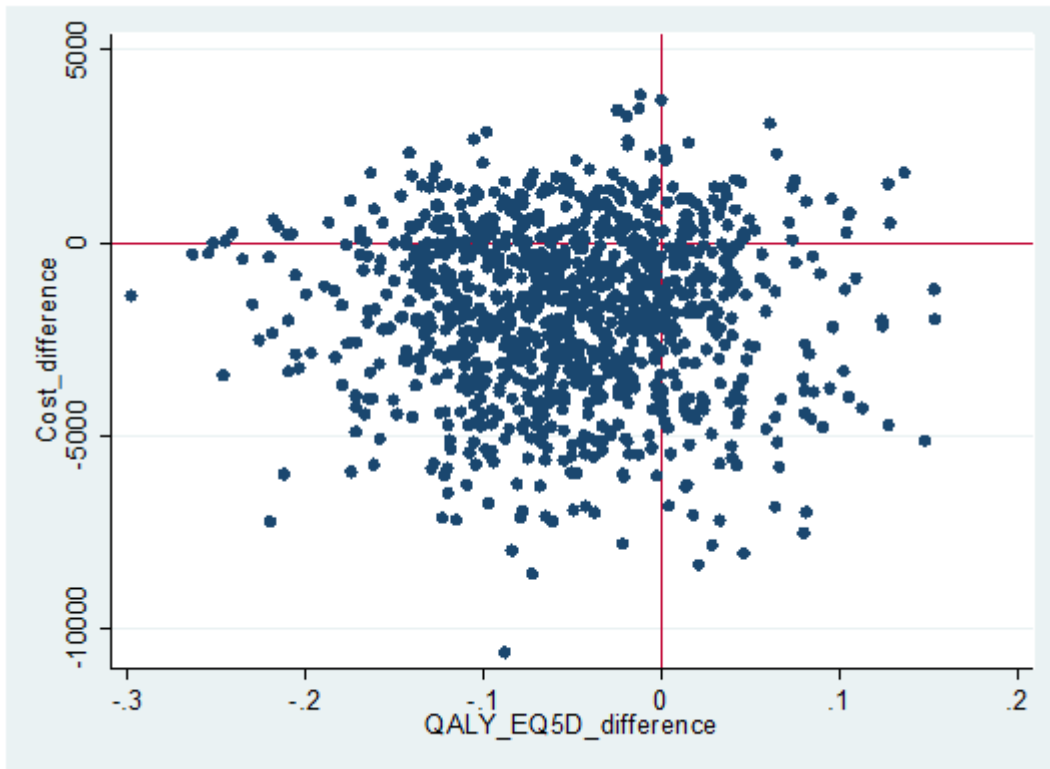


Figure B2 Bootstrapped mean cost versus mean outcome: GAS, HSC perspective

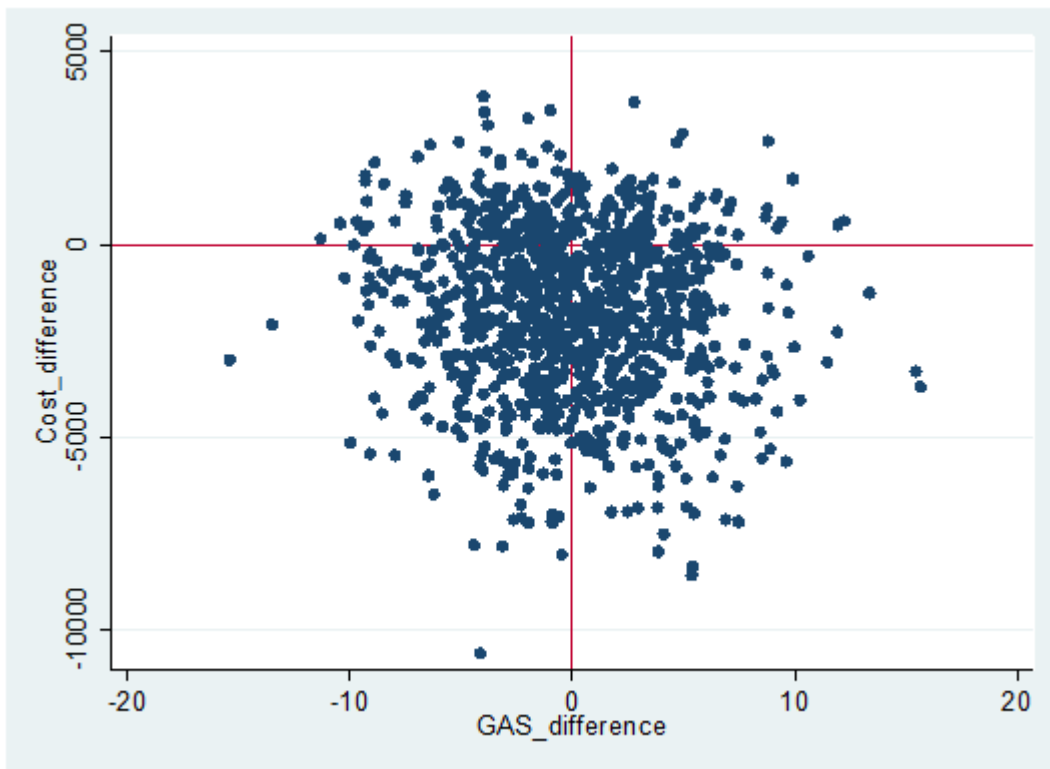


Figure B3 Bootstrapped mean cost verses mean outcome: QPR, HSC perspective

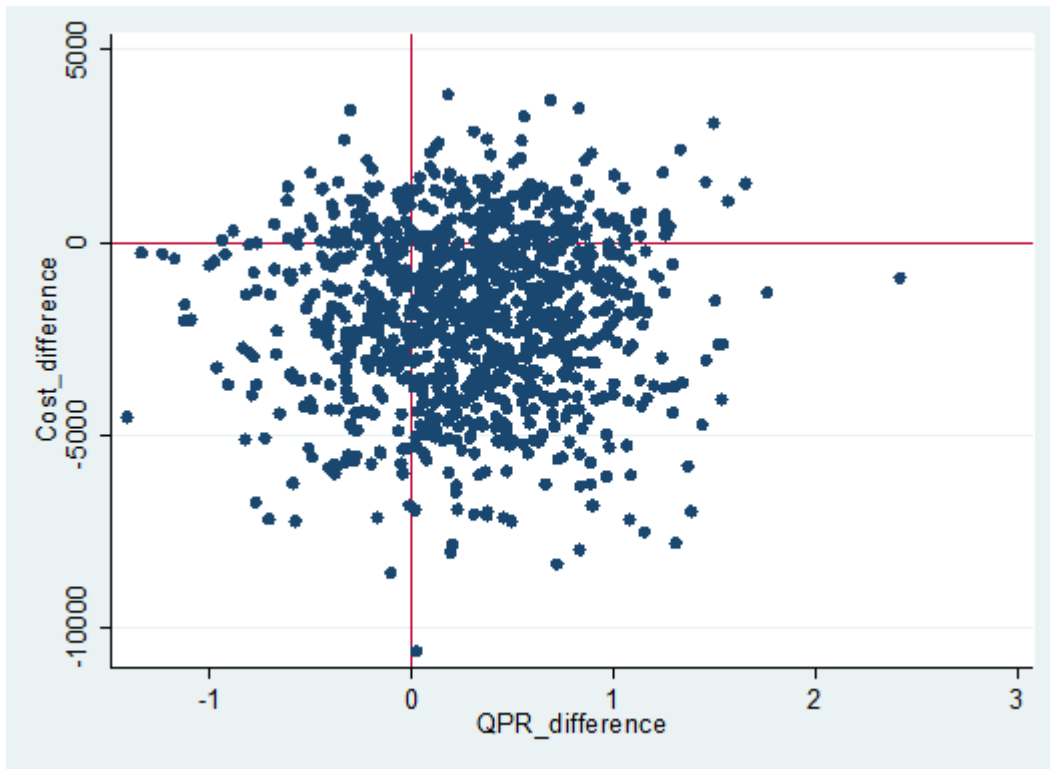


Figure B4 Bootstrapped mean cost verses mean outcome: WEMWBS, HSC perspective

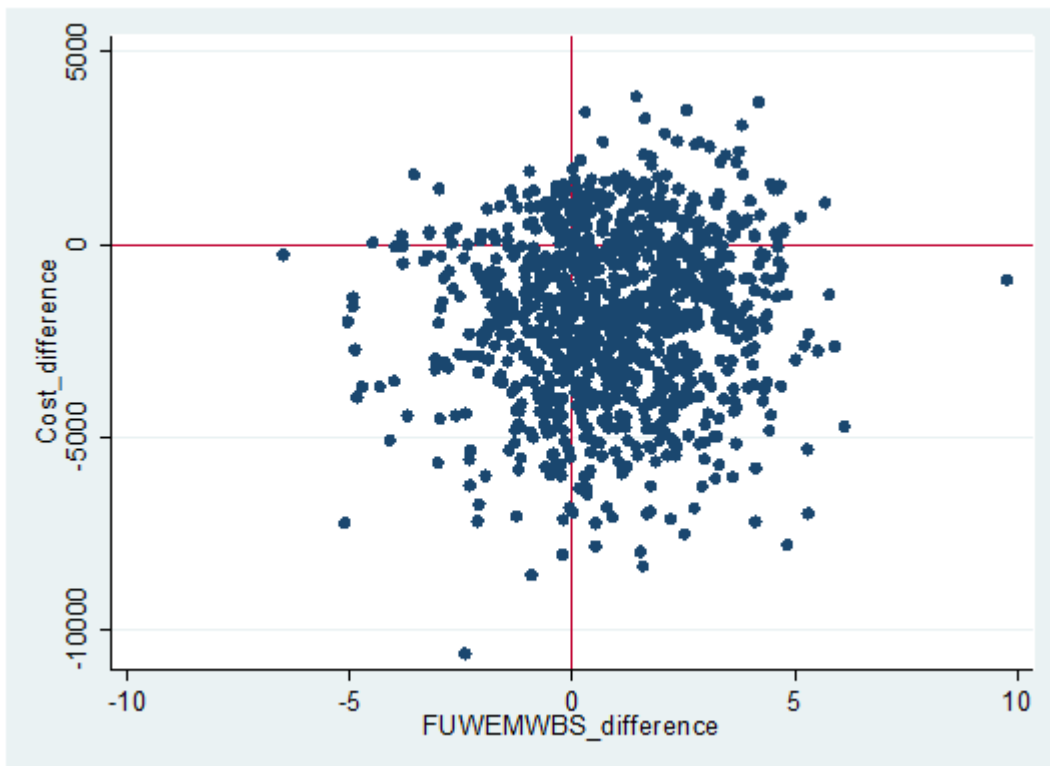


Figure B5 CEAC: RG-UK, HSC perspective

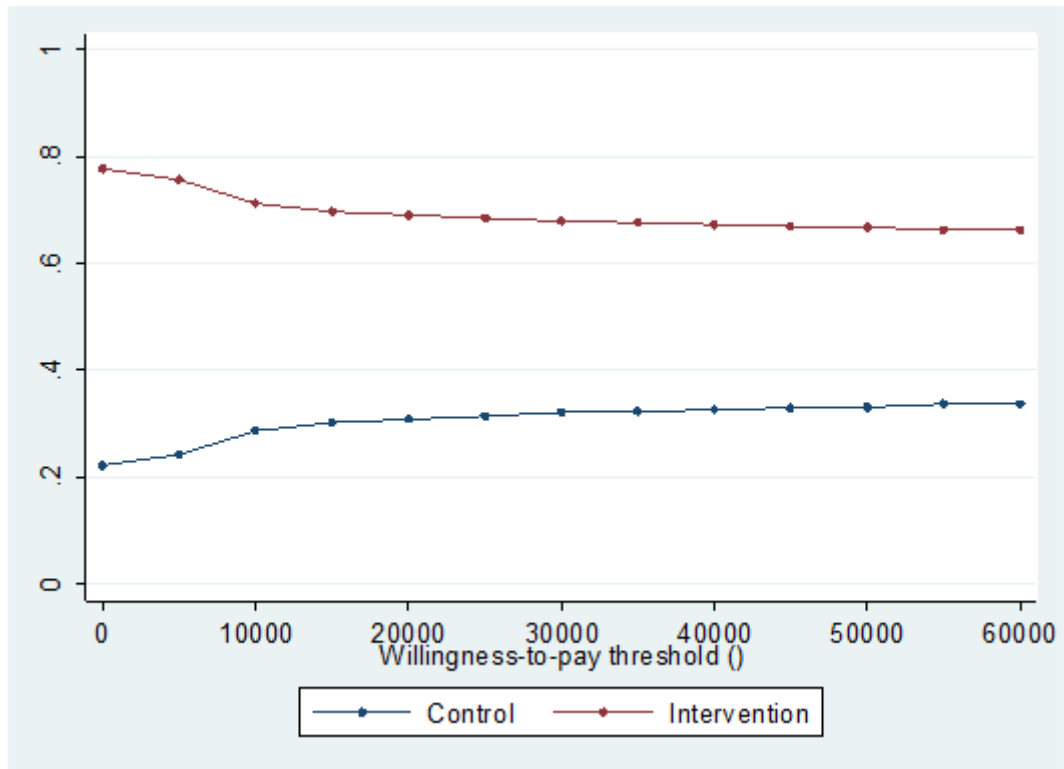


Figure B6 CEAC: QALY (EQ-5D), HSC perspective

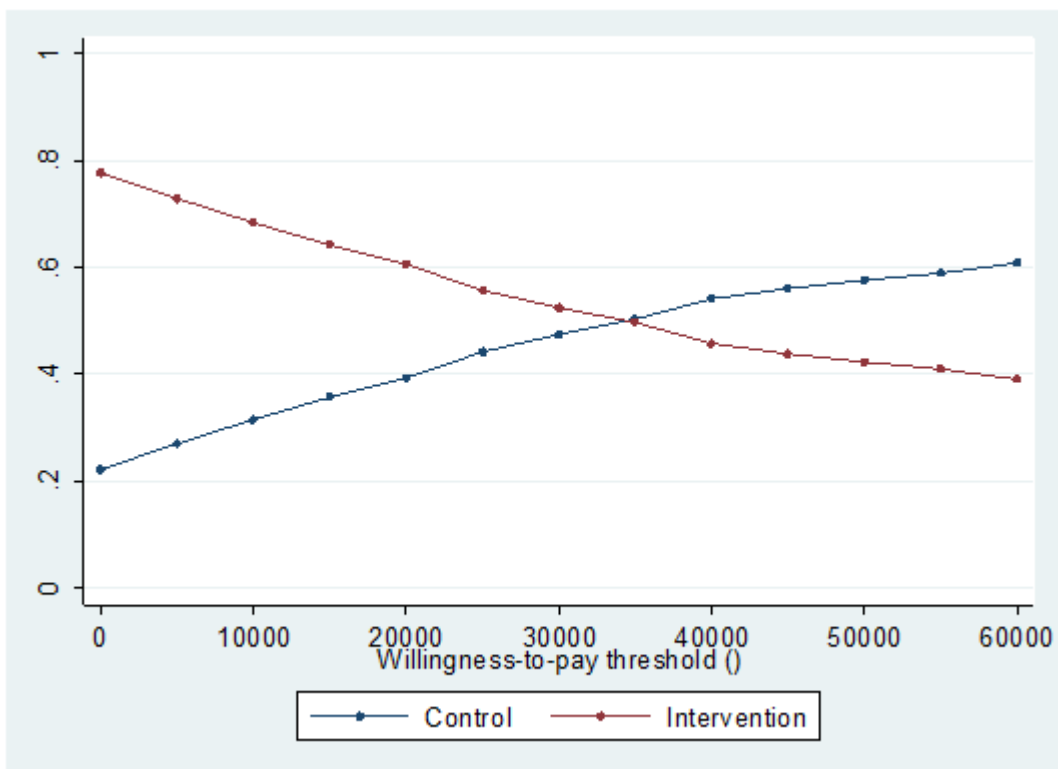


Figure B7 CEAC: GAS, HSC perspective

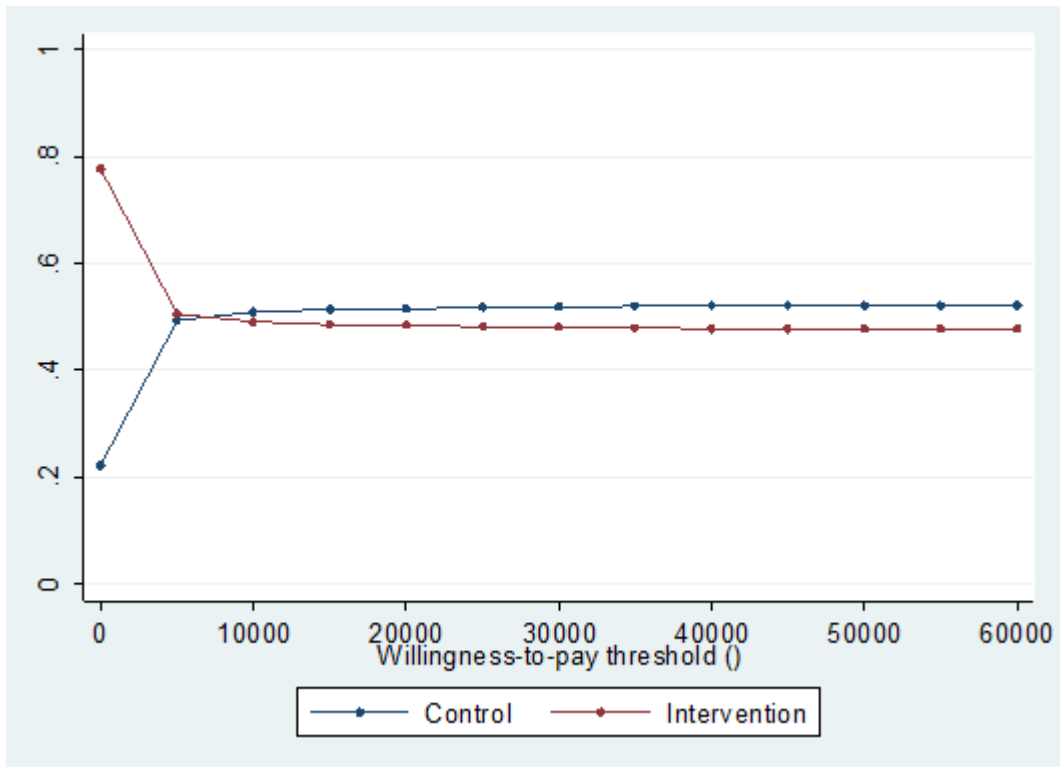


Figure B8 CEAC: QPR, HSC perspective

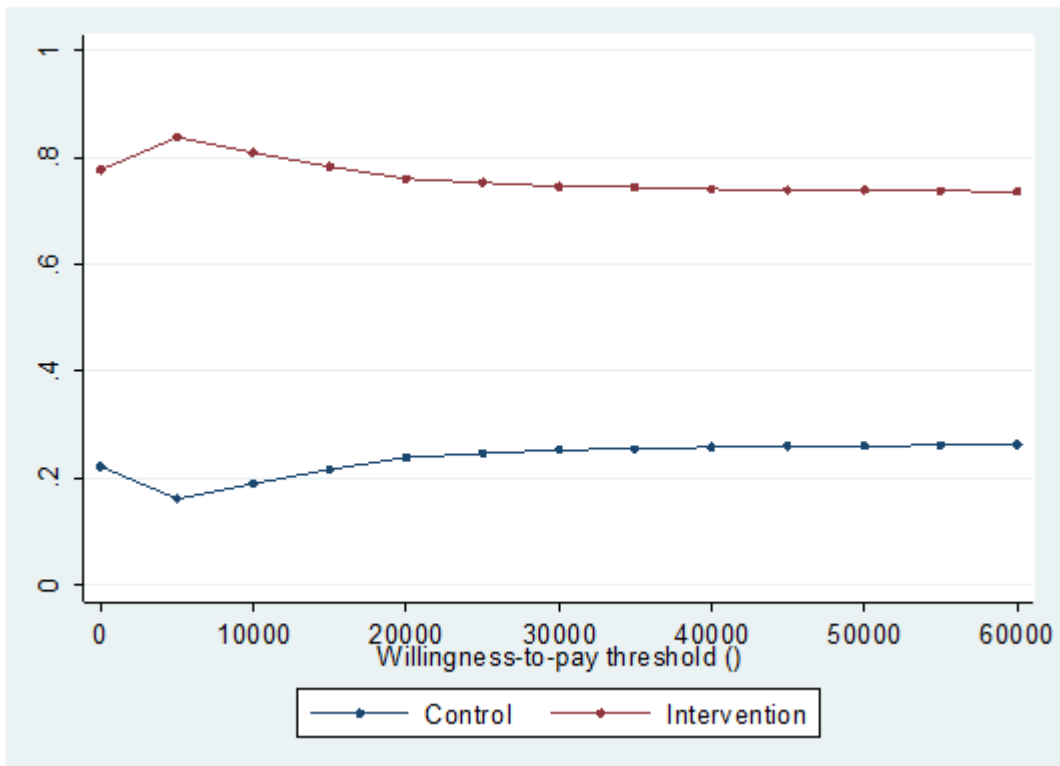
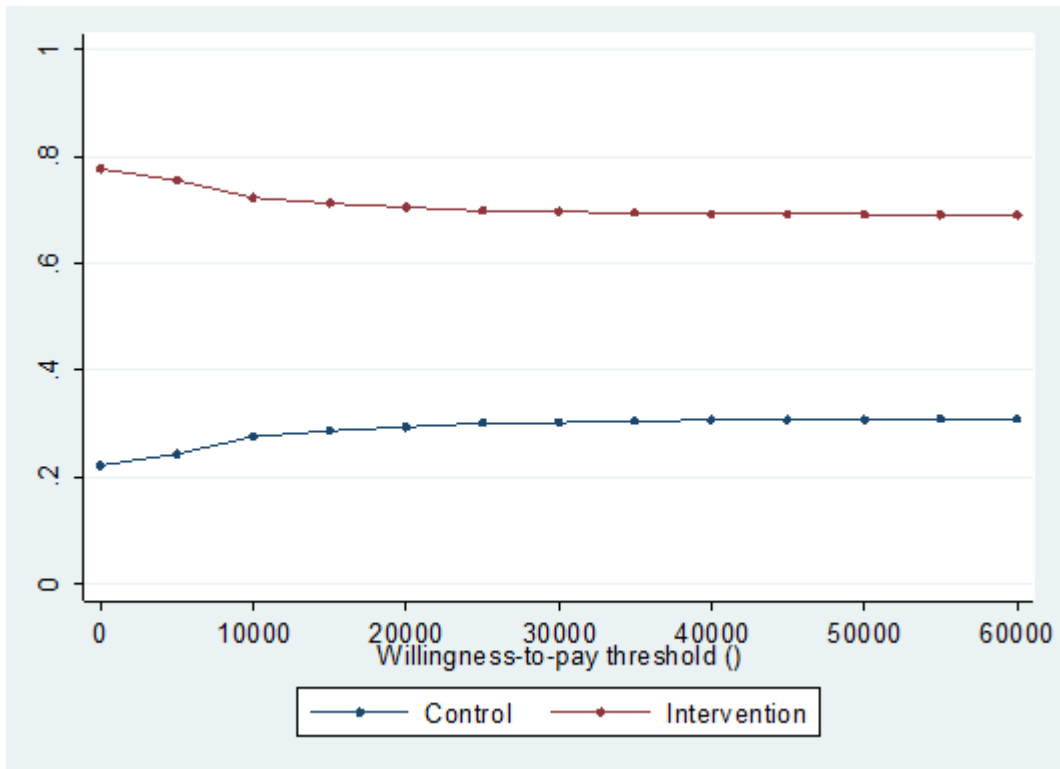


Figure B9 CEAC: WEMWBS, HSC perspective



References

Bancroft, E. (2017) 'What is the average miles driven per year in the UK?'.

Beecham, J.K. and Bonin, E.-M. (2019) 'Economic analysis plan for the effectiveness and cost-effectiveness trial of humanistic counselling in schools (ETHOS)', Canterbury, PSSRU, University of Kent.

Brookes, N., Barrett, B., Netten, A. and Knapp, E. (2013) 'Unit Costs in Criminal Justice (UCCJ)', Canterbury, PSSRU, University of Kent.

Curtis, L. (2010) 'Unit Costs of Health and Social Care 2010', Canterbury, PSSRU, University of Kent.

Curtis, L. and Burns, A. (2017) 'Unit Costs of Health and Social Care 2017', Canterbury, Personal Social Services Research Unit, University of Kent.

Curtis, L. and Burns, A. (2018) 'Unit Costs of Health and Social Care 2018', Canterbury, Personal Social Services Research Unit, University of Kent.

D'Amico, F., Rehill, A., Knapp, M., Aguirre, E., Donovan, H., Hoare, Z., Hoe, J., Russell, I., Spector, A., Streater, A., Whitaker, C., Woods, R.T. and Orrell, M. (2015) 'Maintenance cognitive stimulation therapy: an economic evaluation within a randomized controlled trial', *J Am Med Dir Assoc*, **16**(1), pp. 63-70.

Department for Business Energy & Industrial Strategy (2017) 'National Living Wage and National Minimum Wage. Government response to the Low Pay Commission's Autumn 2017 Report', London, Department for Business, Energy & Industrial Strategy.

Department of Health (2017) 'NHS reference costs 2017/18', London, Department of Health.

Hunger, M., Baumert, J. and Holle, R. (2011) 'Analysis of SF-6D Index Data: Is Beta Regression Appropriate?', *Value in Health*, **14**(5), pp. 759-767.

Kwik Fit (2018) 'Motorists' Costs Running at £162 Per Month-Excluding Cost Of The Car'.

Maheswaran, H., Weich, S., Powell, J. and Stewart-Brown, S. (2012) 'Evaluating the responsiveness of the Warwick Edinburgh Mental Well-Being Scale (WEMWBS): Group and individual level analysis', *Health and Quality of Life Outcomes*, **10**(1), p. 156.

NHS Digital (2018) 'Prescription Cost Analysis - England, 2017'.

Office for National Statistics (2018) 'Employee earnings in the UK: 2018', London, ONS.

Phelan, J. (2016) 'How much does a solicitor charge?'.

Smithson, M. and Verkuilen, J. (2006) 'A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables', *Psychol Methods*, **11**(1), pp. 54-71.

The Law Society (2018) 'Cost of a day in court- New analysis by the Law Society'.