Talking Therapy:
Impacts of a Nationwide Mental Health Service

Ekaterina Oparina[1] ⓡ Christian Krekel[2] ⓡ Sorawoot Srisuma[3] ⓡ

[1]CEP, LSE

[2]PBS & CEP, LSE

[3]NUS, University of Surrey

EEA-ESEM 2024

Motivation

- Mental ill health is widespread:
  - Nearly 1 bln. people globally (WHO, 2022)
  - 30% of non-fatal diseases globally (World Bank, 2016)
  - 1 in 6 adults experienced depressive symptoms in the past month in the UK (ONS, 2023)
- Mental ill health is costly:
  - 4% of UK GDP (OECD, 2017)
  - 4% and 8% of GDP across different regions (Arias et al., 2022)
- Evidence of effectiveness of psychological therapies:
  - Extensive for individual treatment, RCTs (Lambert, 2013; Nathan & Gorman, 2015; Roth & Fonagy, 2005)
  - Very scarce for public policies (Serena, 2022) $\leftarrow$ less than 2% of average healthcare budgets goes to mental health (WHO, 2022)

This Paper: What We Do

- Estimate causal effects of a population-wide mental health policy:
  - Since 2008: NHS treatment for anxiety and depression (CBT: *IAPT* programme, now *Talking Therapies*)
  - 7 million people treated since 2008 (13% of the English population)
  - Replicated around the world
- Research questions:
  - Is IAPT effective?
  - Who benefits the most/least and under what conditions?
- Earlier evaluations:
  - Pre-post comparisons (correlation): Clark et al. (2009), Gyani et al. (2013)
  - Small-scale RCTs: Knapstad et al. (2020), Cano-Vindel et al. (2022)
- Focus on the causal **treatment effects** of the *policy*
- Aim: Improving the programme and informing policy choices globally

This Paper: What We Find

- Is the IAPT effective?
  - Yes. 53% recovery: 43pp (causal) + 10pp (natural recovery rate)
  - Treated: 36pp more likely to improve, 8pp less likely to deteriorate
  - Positive short-term ripple effects on work and social life
- Who benefits the most/least and under what conditions?
  - Significant effects even for those who benefit the least (recovery & improvement, but not deterioration)
  - Higher severity, deprivation, long-term illness, service funding (individuals at higher risk of developing mental health problems)
  - Unemployed patients respond to the treatment less favourably
  - Allowing self-referrals, improves access to mental healthcare (shorter time since onset)

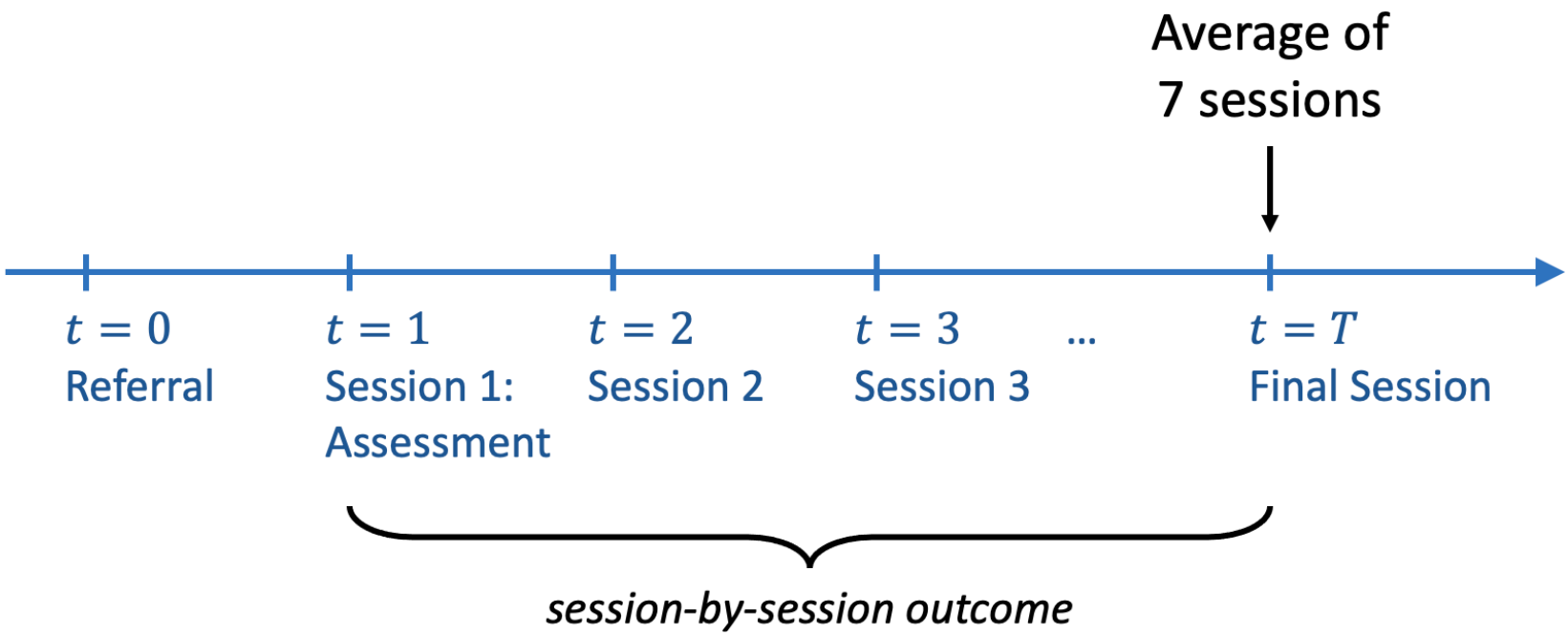
Data and Setup

Data

- Mental health service in England (universal public healthcare system)
- Patient level data on all the IAPT patients from 2016 to 2018: 1,246,792 courses of treatment
- Local area characteristics / service characteristics

- Resulting dataset:
  - Patient individual and treatment characteristics
  - Service characteristics (funding, staff numbers, size)
  - Local area characteristics (incomes, employment, deprivation)

Setting



- Include all patients who had at least 3 sessions

Outcome Data (Session by Session)

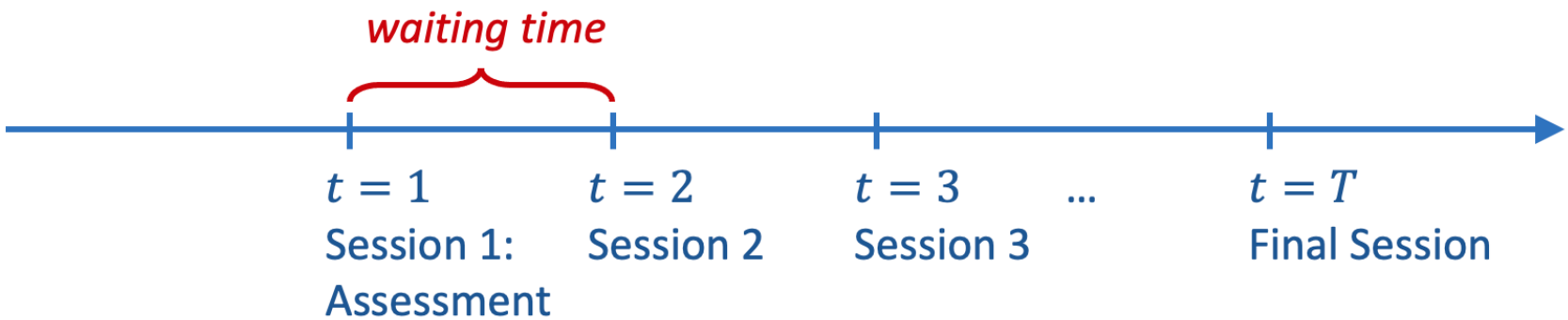
- Depression measure: Patient Health Questionnaire, PHQ-9 [Qs]
- Anxiety measure: Generalised Anxiety Disorder, GAD-7 [Qs]

- Primary outcomes:
  - Reliable Recovery (binary)
  - Reliable Improvement (binary)
  - Reliable Deterioration (binary)
- Secondary outcomes:
  - Work and Social Adjustment Scale (0-40)
  - PHQ-9 (0-27)
  - GAD-7 (0-27)
  - Mental Health Index (zPHQ-9 + zGAD-7)
  - Employment

Empirical Strategy

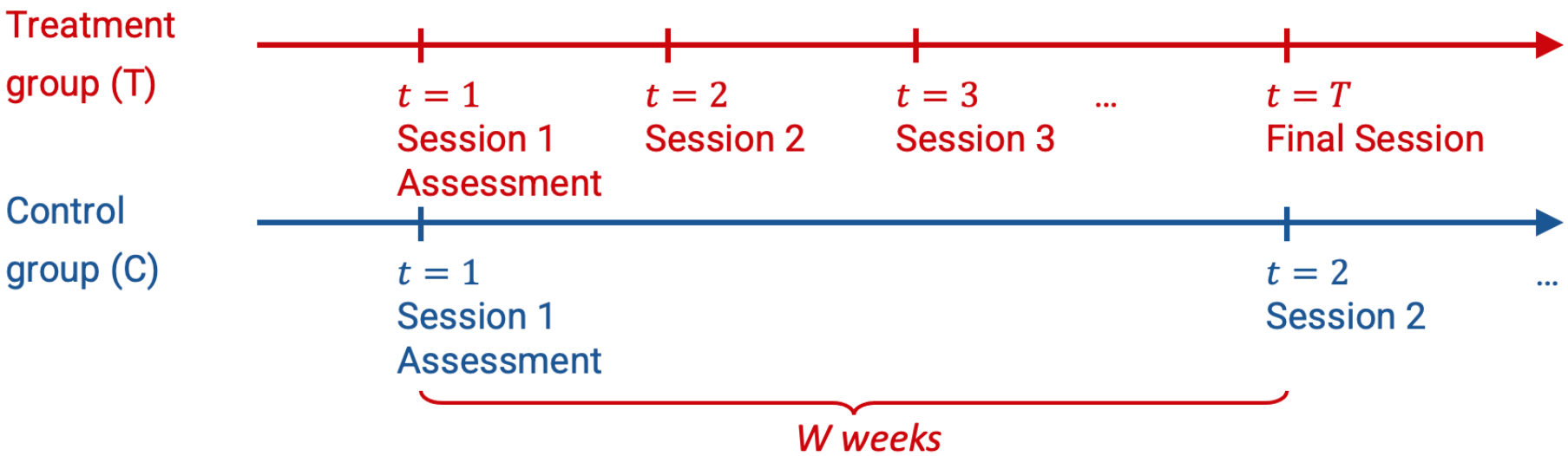
Identification

- No RCT: Construct control group (patients waiting for treatment)
- Session 1: Assessment session (no treatment)
- Session 2: Treatment begins



Identification Cont'd

$$(T_{t=T} - T_{t=1}) - (C_{t=2} - C_{t=1}) = ATT$$



Treatment group	below the median of waiting time
Control group	above the median of waiting time

Identification - Formally

- Potential outcomes framework (Rubin, 1974) two-period model
- Diff-in-Diff assumptions (Roth et al., 2023) assume w fixed:

  Assumption 1: Parallel trends. For all i,

  $$E\left[\Delta Y_{it_i}(0) \,|\, D_{it_i} = 1, X_{it_i}\right] = E\left[\Delta Y_{it_i}(0) \,|\, D_{it_i} = 0, X_{it_i}\right] \text{ almost surely.}$$

  Assumption 2: No anticipatory effects. For all i,

  $$E\left[Y_{it_{i1}}(0) \,|\, D_{it_i} = 1, X_{it_i}\right] = E\left[Y_{it_{i1}}(1) \,|\, D_{it_i} = 1, X_{it_i}\right] \text{ almost surely.}$$

X_{it_i} is a vector of observed characteristics of patient i; define $\Delta Y_{it_i} := Y_{it_{i2}} - Y_{it_{i1}}$ and $\Delta Y_{it_i}(d) := Y_{it_{i2}}(d) - Y_{it_{i1}}(d)$ for $d = 0, 1$.

Validity of Identification Assumptions

- Oversubscription that varies over space and time (25p = 2 w., 75p = 8 w.)

Within-sample selection

- No prioritisation by severity apart from therapy intensity (well-balanced)
- Large dataset: Split by characteristics associated with waiting times
- Detailed dataset: Control for the diagnosis, severity of symptoms, service type, along other characteristics

Out-of-sample selection

- Selective attrition (bound effects) [details]
- Waitlisted patients seeking treatment elsewhere $\rightarrow$ lower bound

Estimation: Average Treatment Effect

- Reduced form (Diff-in-Diff) $\rightarrow$ ATT:

$$\Delta Y_{it_i} = \beta_0 + \beta_1 D_{it_i} + \beta_2 W_{it_i} + \beta_3^\top \tilde{X}_{it_i} + \mu_{ir} + \nu_{it_i} + u_{it_i}. \tag{1}$$

D_{it_i} is the treatment dummy that takes value 1 for the treated; $\Delta Y_{it_i} := Y_{it_{i2}} - Y_{it_{i1}}$ and $\Delta Y_{it_i}(d) := Y_{it_{i2}}(d) - Y_{it_{i1}}(d)$ for $d = 0, 1$ over the period W_{it_i}, which is waiting time for control, treatment duration for treatment; $\tilde{X}_{it_i}$ are individual, service and local-area characteristics; and μ_{ir} and ν_{it_i} are service and time FE.

Estimation: Heterogeneous Treatment Effects

- Pre-selected sources (Nonparametric), matched sample [details]:
  - Pre-selected characteristics (earlier literature) – e.g., high severity, symptoms, deprived area, no disability
  - Higher comparability of the treatment-control match
  - CATT for each group by a combination of covariates

- Data-driven sources (Machine Learning): Generalised Random Forest (Athey, Tibshirani, Wager (2019)) [details]:
  - Data-driven characteristics
  - CATT for each individual

Results

Average Treatment Effect: Baseline

	Reliable Recovery (0-1)		Reliable Improvement (0-1)		Reliable Deterioration (0-1)	
Treatment	0.443***	0.431***	0.388***	0.377***	-0.085***	-0.084***
	(0.004)	(0.004)	(0.004)	(0.003)	(0.002)	(0.001)
Therapy Controls	No	Yes	No	Yes	No	Yes
Individual Controls	No	Yes	No	Yes	No	Yes
Service Controls	No	Yes	No	Yes	No	Yes
Local-Area Controls	No	Yes	No	Yes	No	Yes
Service Fixed Effects	No	Yes	No	Yes	No	Yes
Time Fixed Effects	No	Yes	No	Yes	No	Yes
Mean (last session)	0.536	0.536	0.745	0.745	0.050	0.050
Number of Individuals	1,246,792	1,246,792	1,246,792	1,246,792	1,246,792	1,246,792
Treatment Group	618,574	618,574	618,574	618,574	618,574	618,574
Control Group	628,218	628,218	628,218	628,218	628,218	628,218
R Squared	0.228	0.289	0.152	0.187	0.022	0.064

Note: Linear probability models. Binary dependent variables. Robust standard errors clustered at service level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Individual controls: Age, gender, ethnicity, religion, sexual orientation, disability, employment, and armed forces status

Therapy: Diagnosis, treatment mode, medication, severity at the start, time from referral

Service: Number of referrals, staff and patients, funding

Local area: Unemployment, wages, relative deprivation indices, incomes, employment, education, crime, quality of housing and living environment

Fixed effects: Service, year, month, day of the week

Similar results from non-parametric & ML

Average Treatment Effect: Work and Social Functioning

		Work and Social Adjustment Scale				
	Δ Overall (0-40)	Δ Work (0-8)	Δ Home Management (0-8)	Δ Social Leisure (0-8)	Δ Private Leisure (0-8)	Δ Close Relationships (0-8)
Treatment	-5.709***	-1.091***	-0.998***	-1.390***	-1.084***	-1.145***
	(0.079)	(0.019)	(0.016)	(0.017)	(0.017)	(0.017)
Therapy Controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Service Controls	Yes	Yes	Yes	Yes	Yes	Yes
Local-Area Controls	Yes	Yes	Yes	Yes	Yes	Yes
Service Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Individuals	750,351	750,351	750,351	750,351	750,351	750,351
Treatment Group	369,506	369,506	369,506	369,506	369,506	369,506
Control Group	380,845	380,845	380,845	380,845	380,845	380,845
R Squared	0.138	0.069	0.068	0.104	0.072	0.074

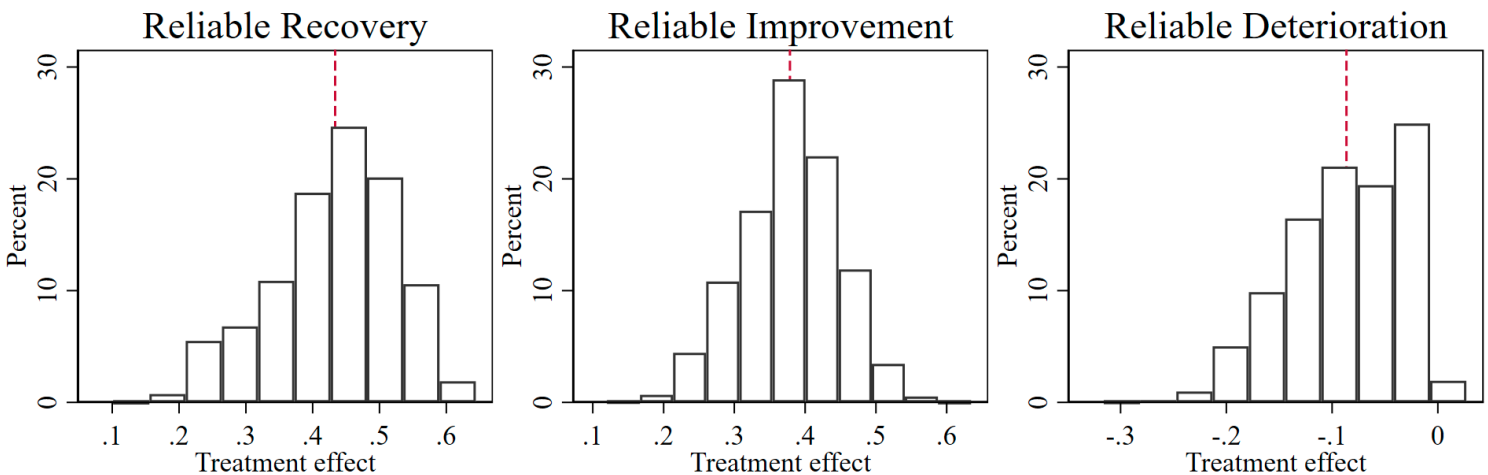
Note: Robust standard errors clustered at service level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Average Treatment Effect: Robustness Checks

- Different definitions of treatment and control group – 25th, 75th, and 90th percentile of waiting time duration [details]
- Different model: Logit instead of a linear probability model + nonparametric & ML results
- Different diagnosis mix [details]:
  - Exclude substance abuse disorders $\rightarrow$ distinct behaviors on a waitlist (Williams and Bretterville-Jensen, 2022)
  - Focus only on depression and anxiety disorders, the main target population of the IAPT programme
- Different outcome definition: ΔPHQ-9, ΔGAD-7, ΔMental Health Index [details]
- Assumption on selective attrition (bound effects) [details]

Heterogeneities: Pre-Selected Sources (Nonparametric Results)

- Keep 76% of the sample
- Significant heterogeneity for all 3 outcomes
- Significant recovery and improvement treatment effects for all groups
- No significant deterioration effect for some groups



Note: The histograms plot the distributions of conditional average treatment effects. The estimates are weighted by the number of treatment-group observations in each sub-population.

Heterogeneities: Pre-Selected Sources (Nonparametric Results) II

	Reliable Recovery	Reliable Improvement	Reliable Deterioration
Severity above median # Treated	-0.088***	-0.071***	0.096***
	(0.002)	(0.002)	(0.001)
Deprivation above median # Treated	-0.026***	0.004**	-0.014***
	(0.002)	(0.002)	(0.001)
Long-term health condition # Treated	-0.026***	0.003	-0.008***
	(0.003)	(0.003)	(0.002)
Service size above median (number of staff) # Treated	-0.004**	-0.006***	0.003**
	(0.002)	(0.002)	(0.001)
Service funding per patient above median # Treated	0.021***	0.026***	-0.010***
	(0.002)	(0.002)	(0.001)
Not religious # Treated	-0.025***	-0.013***	0.007***
	(0.003)	(0.003)	(0.002)
Other religion and missing # Treated	-0.030***	-0.021***	0.006***
	(0.003)	(0.004)	(0.002)
Other ethnicity# Treated	-0.018**	0.000	-0.016***
	(0.007)	(0.008)	(0.005)
Missing ethnicity# Treated	-0.055***	-0.030***	0.002
	(0.003)	(0.003)	(0.002)
R2	0.26	0.16	0.05
Observations	947,945	947,945	947,945

Notes: Omitted categories: Religion: Christian; Ethnicity: White British.

Recovery:

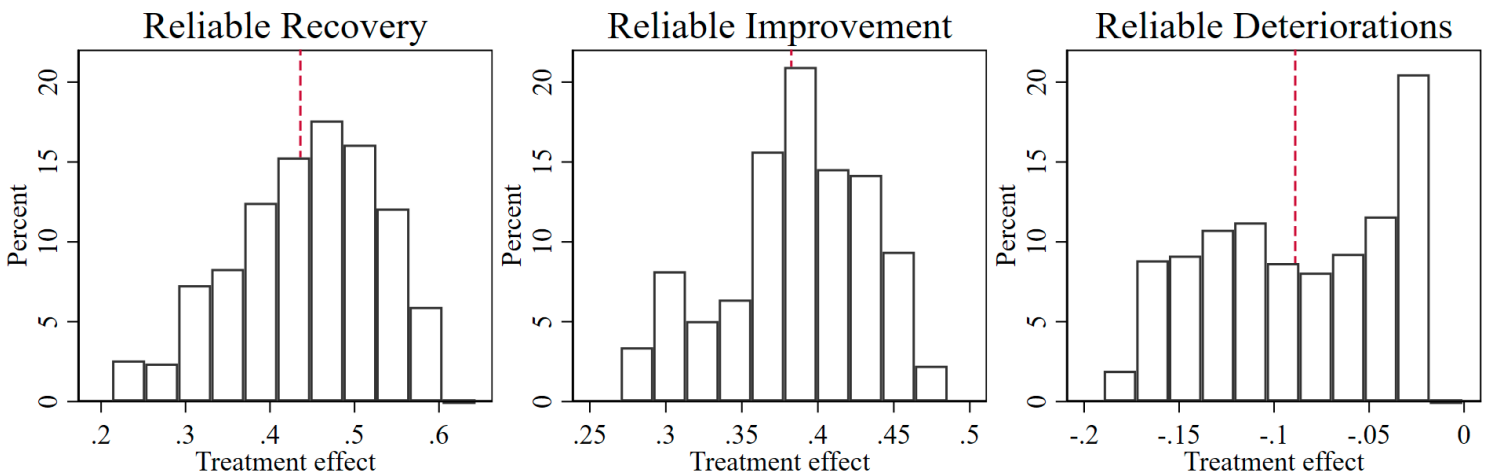
↓ Severity, deprivation, health condition, larger services, non-White-British, poor data quality

↑ Higher funding

Improvement and deterioration largely in-line

Heterogeneities: Data-Driven Sources (ML Results)

- Characteristics by treatment effect quartiles:
  - Recovery: ↓ Unemployed or long-term sick, poor data quality, more severe symptoms at the start, less likely to self-refer, deprived areas, larger services
  - Improvement: ↓ as above + lower funding
  - Deterioration: Patterns less clear ↑ Deprivation, more severe symptoms



Note: The histograms plot the distributions of individual conditional average treatment effects.

Heterogeneities: Data-Driven Sources (ML Results) II

	Reliable Recovery	Reliable Improvement	Reliable Deterioration
Unemployed vs. Employed			
Treated	0.468***	0.387***	-0.085***
	(0.004)	(0.004)	(0.001)
Unemployed	-0.012***	-0.083***	0.029***
	(0.001)	(0.003)	(0.002)
Unemployed # Treated	-0.133***	-0.042***	0.009***
	(0.004)	(0.004)	(0.002)
Number of Individuals	828,356	828,356	828,356
R Squared	0.30	0.19	0.06
Self Referral vs. Non-Self Referral			
Treated	0.404***	0.373***	-0.089***
	(0.005)	(0.004)	(0.002)
Self Referral	0.016***	0.043***	-0.022***
	(0.006)	(0.006)	(0.004)
Self Referral # Treated	0.038***	0.006	0.007***
	(0.005)	(0.004)	(0.002)
Number of Individuals	1,246,792	1,246,792	1,246,792
R Squared	0.29	0.19	0.06

Note: Linear probability models. All controls. Binary dependent variables. Robust standard errors clustered at service level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

- Unemployed *vs.* Employed
  - Unemployed recover less both treated and waiting
  - Important for labour effects of mental health treatment
- Self Referral *vs.* Non-Self Referral
  - Unusual
  - 71.5% of all patients in our sample self-referred
  - Improves access: 364 days since the onset *vs.* 461 days

A (Very Conservative) Cost-Benefit Analysis

- Benefits – mental health only: 5-point decrease in PHQ-9 $\approx$ an increase in the EuroQol-5 Dimensions (EQ-5D) index of about 0.03 points (Furukawa et al., 2021)
- UK Government values 1.0 QALYs at £70,000 (Treasury, 2022)
- Conservative relapse assumptions:
  - Instantaneous relapse of 40% of the patients (≈40% 6y after treatment)
- Monetised benefits over 3 years: **£3,745** per patient
- Costs of **£680** per patient (Clark, 2018)
- Benefit-cost ratio of **5.5**

Summing Up

- Study the effectiveness of a nationwide mental health programme
- It is effective for mental health and social functioning:
  - Increases recovery, improvement, and social functioning
  - On average, reduces deterioration
- There are policy-relevant differences in how well it works (patients at the higher risk, employment, self-referral)
- Benefit-cost ratio of 5.5 on a 3-year period (based on mental health + QALY at £70,000)

Appendix

PHQ-9 questions [back]

Over the past two weeks, how often have you been bothered by:

1. Little interest or pleasure in doing things?
2. Feeling down, depressed, or hopeless?
3. Trouble falling or staying asleep, or sleeping too much?
4. Feeling tired or having little energy?
5. Poor appetite or overeating?
6. Feeling bad about yourself, or that you are a failure or have let yourself or your family down?
7. Trouble concentrating on things, such as reading the newspaper or watching television?
8. Moving or speaking so slowly that other people could have noticed? Or the opposite—being so fidgety or restless that you have been moving around a lot more than usual?
9. Thoughts that you would be better off dead or of hurting yourself in some way?

Not at all, Several days, More than half the days, Nearly every day

GAD-7 questions [back]

Over the past two weeks, how often have you been bothered by:

1. Feeling nervous, anxious, or on edge?
2. Not being able to stop or control worrying?
3. Worrying too much about different things?
4. Trouble relaxing?
5. Being so restless that it is hard to sit still?
6. Becoming easily annoyed or irritable?
7. Feeling afraid, as if something awful might happen?

Not at all, Several days, More than half the days, Nearly every day

Attrition [back]

Table: Average Treatment Effects on Mental Health for Different Recovery Scenarios of Drop-Out Patients

	Main result	Scenario 1	Scenario 2	Scenario 3	Scenario 4
			Reliable Recovery		
	0.431***	0.218***	0.296***	0.404***	0.483***
	(0.004)	(0.009)	(0.007)	(0.004)	(0.004)
R Squared	0.29	0.10	0.16	0.27	0.36
			Reliable Improvement		
Treatment	0.377***	0.195***	0.273***	0.381***	0.460***
	(0.003)	(0.008)	(0.005)	(0.005)	(0.004)
R Squared	0.19	0.07	0.12	0.21	0.28
			Reliable Deterioration		
Treatment	-0.084***	0.016***	-0.063***	-0.171***	-0.249***
	(0.001)	(0.005)	(0.001)	(0.007)	(0.007)
R Squared	0.06	0.06	0.05	0.16	0.21
Number of Individuals	1,246,792	1,507,012	1,507,012	1,507,012	1,507,012
Treatment Group	628,218	684,786	684,786	684,786	684,786
Control Group	618,574	822,226	822,226	822,226	822,226

Note: Linear probability model with all controls. Binary dependent variables. Robust standard errors clustered at service level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Scenario 1: All patients who dropped out of the treatment group deteriorated; hence, none recovered. All patients who dropped out of the control group improved and recovered, none deteriorated. Extreme lower bound.

Scenario 2: All patients who dropped out of the treatment and the control group improved and recovered, none deteriorated.

Scenario 3: All patients who dropped out of the treatment and the control group deteriorated, and none improved or recovered.

Scenario 4: All patients who dropped out of the treatment group improved and recovered, and none deteriorated. All patients who dropped out of the control group deteriorated; hence, none recovered. Extreme upper bound.

Estimation Heterogeneous Treatment Effects [back]

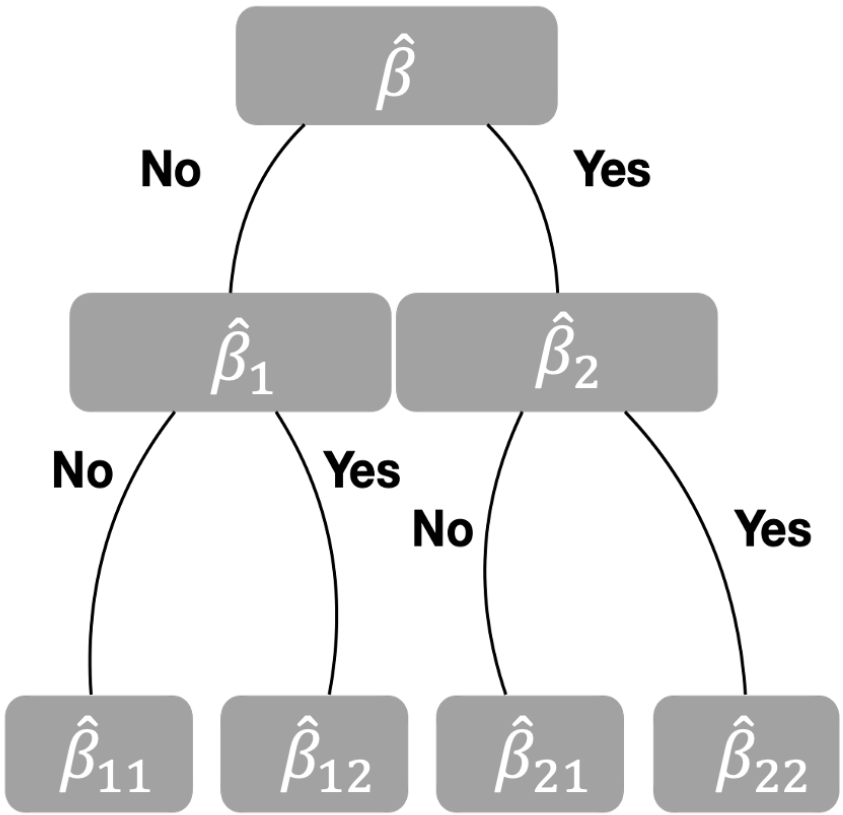
- Nonparametric, matched sample:
  - pre-selected characteristics (earlier literature) – e.g. high severity symptoms, deprived area, no disability
  - higher comparability of the treatment-control match

$$\theta(w,q) := E[\Delta Y^{tr}_{it_i} | W_{it_i} = w, Q_{it_i} = q] - E[\Delta Y^{c}_{it_i} | W_{it_i} = w, Q_{it_i} = q].$$

W_{it_i} and Q_{it_i} are times and combinations of covariates shared by treatment and control group

- CATT for each (w, q) group
- Can be bootstrapped; OLS framework for inference

Generalised Random Forest, CATT [back]



$$\Delta Y_{it_i} = \alpha + \beta D_{it_i} + \epsilon_{it_i}$$

- 'Includes' covariates by orthogonalizing (Robinson, 1988)

$$\Delta \tilde{Y}_{it_i} = \Delta Y_{it_i} - \Delta \hat{Y}_{it_i}(X_{ict})$$
$$\tilde{D}_{it_i} = D_{it_i} - \hat{D}_{it_i}(X_{ict})$$

- Honest approach: different samples for structure and estimation

Reduced-Form Results: Low vs High Intensity

Table: Average Treatment Effects on Mental Health by Treatment Intensity (Low)

Low Intensity						
	Reliable Recovery		Reliable Improvement		Reliable Deterioration	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.440***	0.430***	0.368***	0.360***	-0.078***	-0.078***
	(0.005)	(0.005)	(0.003)	(0.004)	(0.002)	(0.002)
	Δ PHQ-9 (0-27)		Δ GAD-7 (0-21)		Δ Mental Health (Z-Score)	
Treatment	-4.579***	-4.514***	-4.488***	-4.409***	-0.732***	-0.720***
	(0.059)	(0.054)	(0.054)	(0.050)	(0.009)	(0.008)
Number of Individuals	491,942	491,942	491,942	491,942	491,942	491,942
Treatment Group	245,433	245,433	245,433	245,433	245,433	245,433
Control Group	246,509	246,509	246,509	246,509	246,509	246,509
R Squared	0.216	0.284	0.138	0.179	0.020	0.053
Individual Controls	No	Yes	No	Yes	No	Yes
Therapy Controls	No	Yes	No	Yes	No	Yes
Local-Area Controls	No	Yes	No	Yes	No	Yes
Local-Area Fixed Effects	No	Yes	No	Yes	No	Yes
Time Fixed Effects	No	Yes	No	Yes	No	Yes

Table: Average Treatment Effects on Mental Health by Treatment Intensity (High)

High Intensity						
	Reliable Recovery		Reliable Improvement		Reliable Deterioration	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.439***	0.429***	0.404***	0.393***	-0.084***	-0.084***
	(0.008)	(0.008)	(0.007)	(0.006)	(0.003)	(0.002)
	Δ PHQ-9 (0-27)		Δ GAD-7 (0-21)		Δ Mental Health (Z-Score)	
Treatment	-5.458***	-5.486***	-5.047***	-5.035***	-0.846***	-0.847***
	(0.110)	(0.084)	(0.084)	(0.077)	(0.015)	(0.013)
Number of Individuals	275,990	275,990	275,990	275,990	275,990	275,990
Treatment Group	136,379	136,379	136,379	136,379	136,379	136,379
Control Group	139,611	139,611	139,611	139,611	139,611	139,611
R Squared	0.234	0.298	0.164	0.198	275,990	0.069
Individual Controls	No	Yes	No	Yes	No	Yes
Therapy Controls	No	Yes	No	Yes	No	Yes
Local-Area Controls	No	Yes	No	Yes	No	Yes
Local-Area Fixed Effects	No	Yes	No	Yes	No	Yes
Time Fixed Effects	No	Yes	No	Yes	No	Yes

Note: Linear probability model. Binary dependent variables. Robust standard errors clustered at clinical-commissioning-group level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Other definitions of treatment and control group

Table C.III: Average Treatment Effects: Robustness – Other Percentiles of Waiting Time

	Reliable Recovery (0-1)		Reliable Improvement (0-1)		Reliable Deterioration (0-1)	
	(1)	(2)	(3)	(4)	(5)	(6)
25th Percentile of Waiting Time						
Treatment	0.443***	0.458***	0.402***	0.419***	-0.079***	-0.076***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.002)	(0.001)
Number of Individuals	1,246,792	1,246,792	1,246,792	1,246,792	1,246,792	1,246,792
Treatment Group	294,571	294,571	294,571	294,571	294,571	294,571
Control Group	952,221	952,221	952,221	952,221	952,221	952,221
R Squared	0.228	0.280	0.119	0.148	0.011	0.062
75th Percentile of Waiting Time						
Treatment	0.438***	0.464***	0.373***	0.396***	-0.092***	-0.093***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.002)	(0.001)
Number of Individuals	1,246,792	1,246,792	1,246,792	1,246,792	1,246,792	1,246,792
Treatment Group	926,894	926,894	926,894	926,894	926,894	926,894
Control Group	319,898	319,898	319,898	319,898	319,898	319,898
R Squared	0.145	0.222	0.116	0.155	0.023	0.058
90th Percentile of Waiting Time						
Treatment	0.437***	0.456***	0.365***	0.385***	-0.097***	-0.095***
	(0.004)	(0.005)	(0.003)	(0.004)	(0.002)	(0.002)
Number of Individuals	1,246,792	1,246,792	1,246,792	1,246,792	1,246,792	1,246,792
Treatment Group	1,121,181	1,121,181	1,121,181	1,121,181	1,121,181	1,121,181
Control Group	125,611	125,611	125,611	125,611	125,611	125,611
R Squared	0.069	0.153	0.058	0.101	0.015	0.044
Therapy Controls	No	Yes	No	Yes	No	Yes
Individual Controls	No	Yes	No	Yes	No	Yes
Service Controls	No	Yes	No	Yes	No	Yes
Local-Area Controls	No	Yes	No	Yes	No	Yes
Service Fixed Effects	No	Yes	No	Yes	No	Yes
Time Fixed Effects	No	Yes	No	Yes	No	Yes

Other Outcomes and Diagnosis Mix [back]

	Reliable Recovery (0-1)			Other Outcomes		
	Logit Marginal Effect	Without Substance Abuse	Only Depression, Anxiety	Δ PHQ-9 (0-27)	Δ GAD-7 (0-21)	Δ Mental Health Index (Z-Score)
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.381***	0.431***	0.431***	-5.126***	-4.808***	-0.800***
	(0.003)	(0.004)	(0.004)	(0.052)	(0.044)	(0.008)
Therapy Controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Service Controls	Yes	Yes	Yes	Yes	Yes	Yes
Local-Area Controls	Yes	Yes	Yes	Yes	Yes	Yes
Service Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Individuals	1,246,729	1,246,155	996,358	1,246,792	1,246,792	1,246,792
Treatment Group	618,521	618,239	491,358	618,574	618,574	618,574
Control Group	628,208	627,916	504,761	628,218	628,218	628,218
(Pseudo) R Squared	0.263	0.289	0.290	0.286	0.281	0.324

Note: Robust standard errors clustered at service level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

IAPT Programme Overview

- A part of universal healthcare system
- Aim: evidence based psychological therapies available within the NHS
- Focus on depression and anxiety disorders (most common)
- Trained 10,500 therapists: CBT principles
- 7 million people treated since 2008 (13% of English population)
- Step-care model (high, low, step-up treatments)
- IAPT is the largest programme of its kind in the world
- NHS is committed to expansion ($\approx$32% of the community prevalence)

A (Very Conservative) Cost-Benefit Analysis - Algebra

- 5-point decrease in PHQ-9 $\sim$ an increase in the EuroQol-5 Dimensions (EQ-5D) index of about 0.03 points (Furukawa et al., 2021)
- UK Government values 1.0 QALYs at £70,000 (Treasury, 2022)
- Benefits accrue linearly over the course of treatment (2 months)
- Instantaneous relapse of 40% of the patients (conservative, $\sim$40% six years after the end of treatment (cf. Fava et al., 2004))

Monetised benefits over 3 years:

$$\left[\left(\frac{0.00 + 0.03}{2} \times 2 \text{ months} + 0.03 \times 0.6 \times 10 \text{ months}\right) /\right.$$

$12 \text{ months} + 0.03 \times 0.6 \times 2 \text{ years} \times 70,000 = 3,745$ per patient
Costs: £680 per patient (Clark, 2018 - if one divides the total investment into IAPT in 2015–2016 by the total number of courses of treatment)
Net Benefits: £3,745 - £680 = £3,065 per patient three years after the end of treatment, or a benefit-cost ratio of 5.5