Ministry of Justice

Assessing the effectiveness of Radio Frequency Electronic Monitoring for Community and Suspended Sentence Orders

Technical report

Professor Ian Brunton-Smith

University of Surrey

Ministry of Justice Analytical Series

2025

Data and Analysis exists to improve policy making, decision taking and practice by the Ministry of Justice. It does this by providing robust, timely and relevant data and advice drawn from research and analysis undertaken by the department’s analysts and by the wider research community.

Disclaimer

The views expressed are those of the authors and are not necessarily shared by the Ministry of Justice (nor do they represent Government policy).

First published 2025

OGL

© Crown copyright 2025

This publication is licensed under the terms of the Open Government Licence v3.0 except where otherwise stated. To view this licence, visit nationalarchives.gov.uk/doc/open-government-licence/version/3

Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.

Any enquiries regarding this publication should be sent to us at Evidence_partnerships@justice.gov.uk

This publication is available for download at http://www.justice.gov.uk/publications/research-and-analysis/moj

ISBN 978 1 911691 69 3

Acknowledgements

This work was funded by the Evaluation Accelerator Fund (EAF) in collaboration with the Economic and Social Research Council (ESRC). I am grateful for the help and insights provided by the David McAlonan, Charlotte Hartwright, Saleyah Miah and Ann-Marie Jordan in the Ministry of Justice throughout the completion of the project and for providing detailed comments on earlier drafts. In addition, the work was supported by Amy Summerfield, the Evidence and Partnerships Hub, and the Data First Project in the MoJ. I would also like to express my gratitude to the external academic reviewers whose detailed and critical comments were crucial to the final direction of the project.

Contents

List of tables

List of figures

1. Summary

This technical report supplements the quasi-experimental evaluations reported in Brunton-Smith (2025a) and Brunton-Smith (2025b).[1] The evaluations assessed whether Radio Frequency (RF) Electronic Monitoring (EM) was effective for reducing proven reoffending (Brunton-Smith, 2025a), court reconvictions during and after disposals (Brunton-Smith, 2025b) and compliance with other sentence requirements (Brunton-Smith, 2025b).

Focusing solely on RF EM technology and taking a retrospective approach, the evaluations considered: (i) the effectiveness of curfew requirements with RF EM when used in conjunction with a community order; and (ii) the effectiveness of curfew requirements with RF EM in support of a suspended sentence order. The cohort of offenders in Brunton-Smith (2025b) started their sentence between January 2014 and December 2018. The cohort of offenders in Brunton-Smith (2025a) was a subset of this group that started their sentence between April 2016 and March 2017.

Chapter 2 provides an in-depth theoretical rationale for the quasi-experimental approaches (Propensity Score Matching in Brunton-Smith, 2025a; Propensity Score Matching, Coarsened Exact Matching and Causal Machine Learning in Brunton-Smith, 2025b). This includes a discussion of the differences between the causal estimation approaches and their respective strengths and weaknesses.

Details of the data sources and variables used to create the offender cohorts for each evaluation are discussed in chapter 3.

Offender records and details of sentence requirements were extracted from probation information held in the nDelius management information system. Information on previous court convictions and time in prison since January 2011 was extracted from the magistrates' courts and Crown Court datasets (Libra and Xhibit). Data on which offenders

[1] These evaluations were developed as part of a recent collaboration between the Ministry of Justice, the Economic and Social Research Council (ESRC) and the Cabinet Office Evaluation Accelerator Fund (EAF) to explore the feasibility of using administrative data to evaluate policy and practice interventions in the justice system. The statistics reported in Brunton-Smith (2024a) and Brunton-Smith (2024b) should therefore be considered as experimental.

were subject to RF EM as part of their community sentences were included from EM provider administrative records. Finally, Brunton-Smith (2025b) also includes data on offender needs extracted from the offender risk assessment system (OASys).

Chapter 4 includes descriptive statistics for all variables and provides details on the quality of statistical matches (using PSM and CEM) for offenders completing a curfew requirement with RF EM as part of their sentence. Descriptive statistics for included variables are reported for the full cohort of offenders and the subset of offenders sentenced between April 2016 and March 2017.

Finally, chapter 5 considers the primary limitations of the quasi-experimental evaluation approach and the use of administrative data sources.

2. Methodology

2.1 Potential outcomes framework

To measure the impact of RF EM the evaluations reported in Brunton-Smith (2025a; b) adopted the potential outcomes framework to causal estimation (Rubin, 1974).

The potential outcomes framework is grounded in the counterfactual theory, which posits that the causal effect of an intervention is the contrast between the outcome that actually occurred and the outcome that would have occurred had the intervention not been implemented. At the core of this approach is the concept of individual-level counterfactuals (e.g. the difference in outcome for an individual when in receipt of a given treatment compared to not being given a treatment).

Formally, for a given treatment (T=0, T=1), we wish to understand *for each individual* the difference between the outcome under treatment, denoted $Y_i(1)$ and the outcome if they had not received treatment, $Y_i(0)$.

$$\sigma_i = Y_i(1) - Y_i(0)$$

But we can never observe both outcomes for the same individual. Instead, for those subjected to treatment (T=1) we observe $Y_i(1)$ and for those not subjected to treatment (T=0) we observe $Y_i(0)$. However, whilst the individual counterfactual is, by definition, unobservable, it is still possible to estimate the *average treatment effect (ATE)* by considering the difference in the average outcomes for the group that received the treatment, $E[Y(1)]$ and the group that did not receive treatment, $E[Y(0)]$.

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

Rubin (1974) shows that this can yield an unbiased estimate of the causal effect of the treatment if a number of core assumptions are satisfied.

First, is the Ignorability or Unconfoundedness Assumption, which posits that (conditional on a set of observed covariates) the assignment to the treatment is independent of the potential outcomes. This ensures that the treatment and control groups are comparable.

Second, is the Stable Unit Treatment Value Assumption (SUTVA), which asserts that the potential outcome for any unit is independent of the particular assignment of treatments to other units. In other words, the impact of RF EM for a given individual is not dependent on how many (or which) other people have are also being monitored. This assumption is critical for ensuring that the effects observed are attributable solely to the treatment and not to external or inter-unit influences.

Third, the Positivity Assumption requires that every unit has a non-zero probability of receiving treatment, ensuring that the causal effect is identifiable for all units within the population.

Finally, we must also assume Exchangeability, which implies that the potential outcomes are independent of treatment assignment.

These assumptions underlie the rationale for using randomisation in experimental designs as it helps to balance both observed and unobserved covariates across treatment and control groups, thereby mitigating selection bias. However, in most cases randomisation is not possible, either because of difficulties associated with correct assignment of treatment across the population, or in observational studies because treatment assignment has already occurred. When randomisation cannot be assumed, researchers can try to approximate balance between treatment and control groups statistically.

Three statistical approaches to causal estimation were used (Propensity Score Matching, Coarsened Exact Matching and Causal Machine Learning[2]) so that weaknesses inherent in each approach could be mitigated by the strengths of the alternative strategies.

[2] PSM and CEM models were estimated in R, with matching completed using the 'matchit' package and effect estimation completed using the 'marginaleffects' package. CML models were estimated in Python using the double machine learning module from 'econml', and ensemble models and randomised grid-searching module from 'sklearn'.

2.2 Propensity Score Matching (PSM)

Perhaps the most widely used approach for improving the estimation of causal effects in observational studies is the propensity score:

$$e(X) = Pr\,(T = 1|X = x)$$

The propensity score, $e(X)$ measures the *conditional probability* of receiving the treatment (T=1) given an explicit set of covariates, X. Assuming the list of covariates effectively captures the range of potential determinants of treatment assignment that are also correlated with the outcome of interest, this approach can approximate the conditions of a randomised experiment. By conditioning on $e(X)$, it allows for the comparison of expected outcomes between treated and untreated groups, effectively addressing the issue of unobservable counterfactuals.

Estimation is undertaken in two steps. In the first step, the propensity score model is estimated and the propensity scores for each individual are extracted. Observations that did not receive treatment (T=0) are selected to match each treated observation (T=1) based on the proximity of their propensity scores. Typically, for a match to be identified, untreated individuals with propensity scores within around 0.1 standard deviations of a treated individual are needed. Untreated observations are usually selected without replacement (i.e., they can only be matched to a single treated unit) and ties are selected at random.[3] In the second step, the matched group of observations are then compared as if assignment to treatment is random, or a model is estimated which includes the treatment effect alongside the same set of conditioning variables to further correct for imbalance across the treatment and control groups (Greifer, 2023).

Importantly, when using propensity score matching, treated and control units are also usually only selected from within the 'area of common support' (where propensity scores overlap). In practice this means that the sample of control units is often restricted to a subset of all potential control observations that share a similar profile to treated units, and

3 The precise distance between matches, how this distance is calculated, whether control units can be reselected and how many untreated units to match with each treated unit are all at the discretion of the researcher, as is the degree of balance needed between the final treated and control sample. The impact of many of these choices have been explored in this study using sensitivity analysis.

as a result, PSM is generally used to produce consistent estimates of the Average Treatment Effect on the Treated (ATT).

$$ATT = E[Y(1) - Y(0)|T = 1]$$

It is common to find matches for *all* treated offenders using PSM because exact matches are not required. Instead, observations only need to be 'similar', as defined by the closeness of the estimated propensity score between individuals in the treatment and control groups. But the quality of matches is dependent on careful tuning of the similarity measure, decisions around replacement criteria for the control group and the congruence of the propensity model with the data generating process.

To correctly approximate a randomised experiment, the propensity model should include all relevant covariates that are associated with both the treatment assignment and the outcome. This inclusion criterion helps to reduce selection bias by balancing these covariates between the treatment and control groups.[4]

However, while it is essential to include relevant covariates, overfitting the model with too many irrelevant variables (those unrelated to treatment assignment or the outcome) can reduce its effectiveness by capturing too much noise rather than the underlying relationship. It is therefore important to be selective in the identification of variables for inclusion in the matching model. The current analysis identified confounder variables with a combination of prior theory, data availability and data quality (see section 3.3).

The propensity score model also assumes a linear relationship between covariates and the log odds of receiving treatment. Any interactions between covariates must be correctly specified in the propensity model, along with any nonlinear terms. As a result, whilst propensity scores can provide a robust causal estimate, they can be sensitive to modelling decisions and matching criteria.

[4] Balance can be assessed by examining standardised mean differences (with 0.05 difference or less considered optimal), variance ratios (where values should be, at a minimum, between 0.5 and 2, but ideally between 0.8 and 1.25) and overlap statistics (values should ideally be lower than 0.1 across groups) (Greifer, 2023; Rubin, 2001).

In the current evaluations, propensity models with a logit link[5] were used to match the treatment and control groups. To identify the optimal balance-variance trade off a range of different model specifications were estimated before the final matched groups were selected. Balance and match rates for each model configuration were compared against a baseline specification including a 0.1 caliper with 1:1 matching without replacement.[6]

2.3 Coarsened Exact Matching (CEM)

A less model dependent matching approach is exact matching. This does not require the correct specification of the functional form of the relationship of covariates with treatment assignment. Instead, each treated individual is paired with an observation from the control pool only if it shares identical values on all included covariates.

If multiple observations can be matched, they are all retained and equally weighted so that the effective sample size for the treatment and control pools is similar.

Unlike propensity score models, exact matching therefore ensures perfect balance between the treatment and control groups, at least for those where an exact match can be found. Having matched the groups, estimation of the treatment effect (ATT) then proceeds as before.

However, exact matching is a data intensive matching approach that works best when the number of covariates to match on is comparatively small and the size of the pool of potential control units is large relative to the treated units. Exact matching does not perform well when a larger number of continuous covariates are included because it becomes harder to identify suitable matches.

Coarsening the matching variables can help in this instance (Iacus, King and Porro, 2012). Qualitative variables can be simplified by grouping together different categories (for example, grouping together similar offences into broader offence categories). Numeric

5 This assumes model errors are distributed according to the logistic distribution (the natural log of the odds that an event occurs). Treatment and control group assignment were similar when a probit link (based on the cumulative distribution function of the standard normal distribution) was used.

6 Alternate model specifications included: (i) replacing the 14 broad offence groupings with the full offence classification; (ii) replacing the 0.1 caliper with a 0.2 caliper; and (iii) including the interaction between gender and age as well as a quadratic term for age. There was no improvement to balance across these alternative specifications.

variables can be converted into ranged variables (for example, grouped age categories or grouping based on the average distance between each individual score and the overall mean). But in most cases it still becomes necessary to remove some individuals from the treatment group where matches cannot be identified. As a result, the target of inference must change from all those who received treatment to the subset of those who received treatment and who have digital "twins" in the control pool. Assuming the proportion of unmatched units is small, this may have minimal impact on overall conclusions.

However, it remains prudent to supplement conclusions from an exact matching approach with another approach that does not require treated units without exact matches to be dropped prior to analysis. The current impact evaluation therefore uses (coarsened) exact matching in addition to, rather than instead of, a propensity score based match.

In Brunton-Smith (2025a; b) all qualitative confounders were matched exactly between the treatment and control groups. Numeric variables were first grouped using Sturges (1926) rule for identifying appropriate class intervals when summarising numeric data.[7] Here, the intention was to optimise the balance between the treatment and control group without requiring exact numeric matches. Models were estimated separately using the full offence classification and the 14 broad offence groupings. Matching models with the smallest ethnic groups (i.e., Asian Chinese and Asian Other) combined were used because these resulted in a modest increase in match rates with no material changes to balance.

2.4 Causal Machine Learning (CML)

Matching approaches have a degree of intuitive appeal and can yield unbiased estimates of the ATT when the core model assumptions are satisfied, but they can underperform when not all treated units are matched or balance between treated and control groups is poor. They are also not well suited to correct estimation of ATE because matching is focused on identifying digital twins for treated units (meaning the control pool is a subset of eligible cases). This means that conclusions are usually limited to the effects of the intervention on those individuals who actually received the intervention, rather than a more

7 Each group has the same class interval, with the number of groups, k, determined by the formula $k=1+\log2(n)$

general assessment of the effects of the intervention on any offender (whether or not they were actually in receipt of the intervention).

One solution is to estimate causal effects using a model, with many studies relying on a generalised linear framework that includes covariates alongside the treatment indicator in a regression, as well as (optionally) modelling treatment assignment. Like the propensity score model, however, this relies on correct specification of the form of the relationship between predictors and the outcome, and between predictors and the treatment, something which is not often possible to specify in advance.

Recent developments in causal models and machine learning present a way forward here (e.g. Athey, Tibshirani and Wager, 2019; Chernozhukov et al., 2017), allowing researchers to leverage the cutting edge in automated model selection and testing, as well as utilising a broad range of non-parametric model forms (e.g., forest models).

Essentially, this involves using machine learning models to: (i) predict the outcome from the set of included (antecedent) variables; (ii) correctly classify membership of the treatment group from the same pool of confounder variables; and (iii) combine the two prediction models in a final step to obtain an estimate of the treatment effect (Chernozhukov et al., 2017; Wager and Athey, 2018).

As a machine learning approach, model specification is optimised to minimise prediction error at each stage, enabling a more robust functional form of the expected relationships with the treatment and outcome to be identified. This can leverage machine learning tools for model selection with cross-validation, effectively introducing a degree of automation into the process.

However, the emphasis of CML on optimising prediction means it can be susceptible to overfitting without careful tuning of model hyperparameters. It is also necessary to assume a linear probability model when estimating the outcome model, with no equivalent nonlinear modelling approach. This is unlikely to be problematic in the context of court reconvictions for two reasons. First, interest is in the *difference* in the percentage of court reconvictions for the group in receipt of RF EM compared to the group not in receipt of RF EM (which can be considered linear). Second, the probability of reconviction in each group is not expected to be close to 0 or 1 (where the linear probability model is known to underperform).

Importantly, by retaining all observations from the control pool in the estimation stage, these causal machine learning models can also produce consistent estimates of ATE. This also means that they are well suited to estimating heterogeneous treatment effects, also referred to as the conditional average treatment effect (CATE).

$$CATE(x) = E[Y(1) - Y(0)|X = x]$$

This has a similar form to the ATT, but instead of only estimating the average effectiveness of treatment on the treated (T=1) it is possible to estimate the average effectiveness of treatment for groups defined by the covariates, X. In our case, we might anticipate that impact of RF EM might depend on offence type, with a different average treatment effect for different offence categories.

A double machine learning approach to causal estimation using causal forest models (Athey, Tibshirani and Wager, 2019; Chernozhukov et al., 2017; Wager and Athey, 2018) was used to determine if there was evidence of differences in the effectiveness of RF EM for offenders sentenced for different categories of offence. A gradient boosting binary classifier was used to estimate treatment conditional on the full set of matching covariates. A gradient boosting regressor was used to estimate *y* conditional on the same predictors and a causal forest model was used to estimate the final (heterogeneous) treatment effect sizes, assuming a linear probability model. Selection of appropriate hyperparameters for the treatment and outcome models was achieved using a randomised grid-based search strategy where candidate hyperparameters were randomly selected and compared.[8]

A total of 10 random sets of hyperparameters were examined for each outcome, with each test using three-fold cross-validation on a training sample of 80 per cent of the data.

---

[8] As a random forest modelling approach, hyperparameter tuning relates to the number and complexity of the decision trees that are estimated. The number of decision trees was varied from 100-300 with learning rates (the contribution of each tree to the final model) of 0.01, 0.05 and 0.1. The maximum depth of each decision tree (it's overall complexity) was varied from 3-5 with minimum sample split (the number of samples to draw from X to train each base estimator) of 2, 4 and 6 and minimum sample leaf (the minimum number of observations required in an individual leaf node) of 1-3. Final selection was based on predictive accuracy and all models were re-estimated with alternative hyperparameters to determine overall model sensitivity.

Having selected the best performing hyperparameters and estimated the final causal forest model using the training data, estimated effect sizes were produced using the remaining 20 per cent test sample.[9] The CML algorithm enables direct estimation of the effect of treatment for all units included in the test sample, with ATE simply the average across all units and ATT the average effectiveness for treated units.

[9] Results for the overall effect of EM in each case were substantively in line with the results using PSM and CEM with consistent estimates of statistical significance.

3. Data and measures

3.1 Data

The impact evaluations of RF EM within community sentences, as reported in Brunton-Smith (2025a; b), used data extracted from the probation service management information system, nDelius. The analysis reported in Brunton-Smith (2025b) used the cohort of offenders aged 18–90 whose probationary requirements commenced between January 2014 and December 2018.[10] A total of 371,977 records were for community orders, and 191,384 records were for suspended sentence orders.

The impact evaluation of 12 month proven reoffending using linked PNC data reported in Brunton-Smith (2025a) was based on a subset of this offender cohort whose probationary requirements commenced between April 2016 and March 2017.[11] Offenders could have been completing single requirement or multiple requirement orders within their community sentences. Offenders may have had multiple probation records if they were managed by the probation service on more than one occasion.

The data were combined with data from the magistrates' courts and Crown Court, enabling a more complete picture of criminal justice system journeys for those individuals subject to any form of probationary supervision. Offender records were linked to EM service provider data to correctly identify the cohort of offenders subject to a curfew requirement with RF EM. Finally, Brunton-Smith (2025b) also included details of offender needs extracted from OASys risk reports.

[10] Quality issues with EM service data before 2014 mean it was not possible to examine effectiveness prior to this date. The evaluation window is limited 2018 to allow a sufficient follow up period for offenders to reappear in reoffending data whilst ensuring results are not complicated by the timing of the Covid-19 pandemic of 2020 (which saw lower reoffending overall).

[11] April 2016 to March 2017 covers the first full financial year of data following the change in methodology used to calculate PNC reoffending (for a full explanation on the changes see http://www.gov.uk/government/statistics/proven-reoffending-statistics-october-2015-to-december-2015).

Linkage between all databases was probabilistic using the Ministry of Justice's Splink package (Ministry of Justice, 2021), with links generated based on similarity of names (forename, surnames, other names), date of birth and location.

3.2 Measuring Electronic Monitoring (EM)

Probation data do not contain a comprehensive measure of which offenders were in receipt of RF EM, with curfew orders only flagged as involving RF EM if they were actively monitored by probation or an additional EM qualifier was attached to a particular requirement.

Probation data contained a total of 10,787 curfew requirements directly flagged as being subject to RF EM between January 2014 and December 2018, with the majority of curfew orders (92%) not identified as involving EM. This figure falls some way below the official reported levels of EM usage.[12]

As a result, the cohort of EM treated offenders was identified using direct source data from the main EM service provider. The EM service provider data included a total of 121,370 records where EM was used as part of a community order or suspended sentence order (started between 1 January 2014 and 31 December 2018) and the offender had a valid record in the probation database.[13] It is important to note that all of these records involved the use of RF EM as opposed to other forms of EM (e.g., GPS enabled or alcohol monitors).

The EM service data were first linked to individual offenders. For each offender with linked RF EM data, an EM record was then linked to a specific probation record if (i) the start date recorded by the EM service provider was the same (or up to seven days later) as the curfew start date recorded in the probation database, or (ii) if the start date for any requirement was the same as the EM start date.

The seven day window between curfew start date and RF EM start date allows for possible delays in EM installation, but leaves open the possibility that some curfew requirements

12 https://data.justice.gov.uk/contracts/electronic-monitoring

13 An additional 15,778 records were available where the offender could not be identified in the probation data.

may erroneously be flagged as involving EM – for example, if an offender was sentenced to a community order with no EM within seven days of them being granted bail with EM for another offence. The impact of this is likely to be small, with the majority of linked records (95%) sharing an identical order date across both databases. Approximately 75 per cent of EM records were directly matched to probation records.[14]

All probation records for offenders not subject to RF EM were assumed eligible for inclusion in the matched control groups.[15] The current official sentencing guidelines[16] do identify additional eligibility criteria determining whether to impose EM as part of a curfew order: (i) there is a person (other than the offender) without whose co-operation it would not be practicable to secure the monitoring and that person does not consent; and/or (ii) electronic monitoring is unavailable and/or impractical; and/or (iii) the particular circumstances of the case, lead the sentencer to consider it inappropriate to do so.

However, no information was available in the extracted data to accurately identify these cases and remove them from the control pool. In practice, the number of cases in which EM would not be used for these reasons is likely to be low (Hucklesby and Holdsworth, 2016).

3.3 Potential confounders

The absence of a randomised design for allocation of RF EM and inability to exclude ineligible records from the control pool meant it was necessary to account for as many potential confounders as possible in the matching/modelling process.

Specifically, the aim was to account for all offender features that plausibly determined whether an offender's sentence included RF EM, as well as being causally related to future reoffending, with the assumption that the mechanisms determining receipt of RF EM are conditionally independent of reoffending. However, it was also important to be

[14] Match rates were slightly lower for community orders (73%) than suspended sentence orders (80%). If reoffending rates (and related outcomes) for these unlinked records are systematically different, it is possible that the estimates of the effectiveness of RF EM would be biased.

[15] Offenders that had multiple probation records including some involving EM and some not involving EM were excluded from the control pool.

[16] https://www.sentencingcouncil.org.uk/overarching-guides/magistrates-court/item/imposition-of-community-and-custodial-sentences/

selective in identifying confounders to avoid inducing too much additional noise into estimates. A total of 51 potential confounding characteristics were selected covering demographic information, details of the sentenced offence, number of previous court convictions (separately by offence category),[17] disposal length and number of probationary requirements. A further eight OASys risk scores were used in the analyses reported in Brunton-Smith (2025b).

No existing evidence was available on the factors influencing judicial decisions to use RF EM as part of a condition of a community sentence. By contrast, much is now known about the key determinants of offending behaviour that is used to inform the selection of key confounder variables.

Existing research has routinely demonstrated the importance of gender and age for offending (Hirshi and Gottfredson, 1983; Nagin and Land, 1993). Age at the time of committing the sentenced offence was measured in integer years.

It is also well known that offender journeys through the criminal justice system are a function of the ethnicity of the offender (Lymperopoulou, 2022). Offender ethnicity was recorded, distinguishing between 18 ethnic groupings. This was the most granular ethnic classification available in administrative records. Information about ethnic origin was not available for approximately nine per cent of offenders and these individuals were excluded from the main analysis. All matching and outcome models were re-estimated excluding the ethnicity variable to assess the sensitivity of results to the presence of this missing data. Results were generally robust to the exclusion of ethnicity.

Disposals are, in part, structured by the type of offence committed. Details of the sentenced offence were recorded using the 14 Home Office offence groupings.[18]

In those instances where more than one offence was considered during sentencing, the most serious offence was used for matching purposes. Details of the number of

[17] This does not account for out of court disposals or overturned convictions.

[18] Sensitivity analyses repeating the matching and outcome evaluations using more granular offence codes (covering more than 140 different offence types) produced almost identical results.

additional offences that were covered by each probation event were also recorded (ranging from 0–69).

Prior offending is a key determinant of subsequent offending (Moffit, 1993). Linked magistrates' courts (Libra) and Crown Court (Xhibit) datasets were used to measure prior court convictions, with separate counts of the number of convictions for offences relating to drugs, possession or use of a weapon, public order, robbery, theft, violence, summary offences and other offences.[19] The number of times each offender had spent time in prison prior to the currently sentenced offence was also recorded.

To determine whether being subject to RF EM within a community sentence was associated with reduced offending during the EM period and during the total disposal period (a duration that is likely to be different for offenders completing multiple requirement orders), it was necessary to ensure comparable disposal durations between the treatment and control groups.[20] As a result, the total length of time between disposal start and disposal termination was included in the matching models.[21]

It is possible that this will induce a post-treatment bias if the disposal length is, in part, determined by the decision to impose a curfew requirement with EM – something that is most likely when offenders are completing a single requirement order. However, the current sentencing guidelines typically require sentencers to identify an initial disposal length at the first step of the sentencing process, prior to the decision to utilise RF EM, so the impact may be minimal. The year when the current disposal started was also recorded.

Other variables relating to the additional requirements imposed on offenders were not included because of the potential risk of post-treatment bias – namely, the use of RF EM may plausibly influence the types of additional requirement that judges/sentencers may

[19] No data was available about offences that were committed prior to 2011 and as a result the measures of offending history are left-censored. It is likely that we are underestimating the true extent of offending histories and that this will be systematically biased for older offenders. However, given that the focus is on identifying (near) identical matches for each treated individual, this bias will be equally felt in both treatment and control groups, and is therefore unlikely to adversely impact results.

[20] It is not possible to also ensure comparable monitoring durations between the treatment and control groups because offenders in the control group have no monitoring duration data.

[21] Offenders that were not completing any additional requirements in addition to their curfew order with EM have the same date recorded for disposal start and termination in nDelius. For these records (n=42,607), the start and end dates record by the EM service provider were used to calculate the total disposal length.

choose to impose as part of a specific sentence and, as such, they are causally dependent on RF EM. Inclusion of these types of post-treatment variables can induce post-treatment bias, resulting in biased estimated treatment effects.

The *number* of additional probationary requirements was selected for matching to enable separate analysis of the impact of RF EM when used as part of a single requirement order and when used as part of a multi-requirement order within a community sentence. However, whilst this may induce additional post-treatment bias, the absolute number of different probationary requirements is less likely to be causally related to RF EM than the specific *types* of requirement.

Specific offender needs related to, for example, accommodation, employment or substance misuse are also likely to be predictive of future reoffending risk. Information on offender needs was extracted from OASys records and linked to each offender. Offender risk assessments must have been made prior to the sentenced offence to minimise the risk of inducing additional post-treatment bias. No additional restrictions on the duration between the assessment being made and the current sentenced offence occurring were imposed, so some assessments may have been completed more than 12 months prior to the offence. Where offenders had more than one valid OASys record, the most recent risk assessment was selected.

OASys data were extracted to capture offender needs relating to: accommodation; employment; relationships; lifestyle and associates; drug misuse; alcohol misuse; thinking and behaviour; and attitudes. In all cases, offenders were flagged as “in need of support” if the responses to a set of questions exceeded an accepted threshold. For example, accommodation need was flagged if an offender had problems with at least two of: currently having no fixed abode or being in transient accommodation, living in unsuitable accommodation, living in temporary accommodation or living in an unsuitable location.

Importantly, valid OASys records were only available for around half (49%) of offenders. These were often individuals who had served a custodial sentence since 2011 (42% of those without a history of custody had a linked OASys record compared to 81% of those with a prior custody record), which means they are highly susceptible to selection bias.

Results including OASys records were only reported in Brunton-Smith (2025b), where analysis was repeated with and without the inclusion of OASys variables.[22]

[22] A similar approach was used in Eaton and Mews (2019) where it was clear that the subsample of OASys matched offenders had, on average, higher reoffending rates than the general offender population. In their research, the relative differences between the treatment (community) and control (custody) groups remained similar, albeit at different absolute levels.

4. Descriptive statistics and post-matching balance

4.1 Community orders

Table 4.1 records how the 371,977 community orders were distributed across offender and index offence types, as well as the differences between community orders with an RF EM condition (n=52,115) and community orders not including RF EM (n=319,862).

Table 4.1. Descriptive statistics for community orders, 2014–18[23]

	non-RF EM (N=319,862)	RF EM (N=52,115)	Total (N=371,977)
Age (at offence)	32.93 (10.76) 31: 18.00–87.00	32.82 (10.72) 31: 18.00–85.00	32.92 (10.75) 31: 18.00–87.00

Gender

	non-RF EM (N=319,862)	RF EM (N=52,115)	Total (N=371,977)
Female (reference)	53,140 (16.6%)	9,171 (17.6%)	62,311 (16.8%)
Male	266,722 (83.4%)	42,944 (82.4%)	309,666 (83.2%)

Ethnicity

	non-RF EM (N=319,862)	RF EM (N=52,115)	Total (N=371,977)
Asian Indian	4,840 (1.5%)	481 (0.9%)	5,321 (1.4%)
Asian Pakistani	6,451 (2.0%)	982 (1.9%)	7,433 (2.0%)
Asian Bangladeshi	2,515 (0.8%)	250 (0.5%)	2,765 (0.7%)
Asian Chinese	287 (0.1%)	16 (0.0%)	303 (0.1%)
Asian Other	3,358 (1.0%)	306 (0.6%)	3,664 (1.0%)
Black Caribbean	9,604 (3.0%)	1,304 (2.5%)	10,908 (2.9%)
Black African	8,701 (2.7%)	933 (1.8%)	9,634 (2.6%)
Black Other	2,776 (0.9%)	304 (0.6%)	3,080 (0.8%)

23 Numeric variables report: mean, (sd), median, min-max. Qualitative variables report N (%)

	non-RF EM	RF EM	Total
	(N=319,862)	(N=52,115)	(N=371,977)
White and Black Caribbean	6,074 (1.9%)	962 (1.8%)	7,036 (1.9%)
White and Black African	1,454 (0.5%)	194 (0.4%)	1,648 (0.4%)
White and Asian	1,253 (0.4%)	203 (0.4%)	1,456 (0.4%)
Mixed Other	2,039 (0.6%)	255 (0.5%)	2,294 (0.6%)
Arab	712 (0.2%)	70 (0.1%)	782 (0.2%)
Other ethnicity	3,954 (1.2%)	351 (0.7%)	4,305 (1.2%)
White British	240,248 (75.1%)	43,178 (82.9%)	283,426 (76.2%)
White Irish	2,569 (0.8%)	436 (0.8%)	3,005 (0.8%)
White Roma	1,682 (0.5%)	303 (0.6%)	1,985 (0.5%)
White Other	21,345 (6.7%)	1,587 (3.0%)	22,932 (6.2%)

Index offence

	non-RF EM	RF EM	Total
	(N=319,862)	(N=52,115)	(N=371,977)
Criminal damage	1,628 (0.5%)	296 (0.6%)	1,924 (0.5%)
Drugs	15,988 (5.0%)	3,041 (5.8%)	19,029 (5.1%)
Fraud	13,197 (4.1%)	1,735 (3.3%)	14,932 (4.0%)
Miscellaneous	13,060 (4.1%)	2,038 (3.9%)	15,098 (4.1%)
Weapons	629 (0.2%)	102 (0.2%)	731 (0.2%)
Public order	9,520 (3.0%)	1,625 (3.1%)	11,145 (3.0%)
Robbery	185 (0.1%)	27 (0.1%)	212 (0.1%)
Sex offences	2,816 (0.9%)	315 (0.6%)	3,131 (0.8%)
Summary	50,437 (15.8%)	8,796 (16.9%)	59,233 (15.9%)
Summary (motoring)	44,333 (13.9%)	5,935 (11.4%)	50,268 (13.5%)
Theft	57,568 (18.0%)	12,725 (24.4%)	70,293 (18.9%)
Violence	109,783 (34.3%)	15,244 (29.3%)	125,027 (33.6%)
Other (Breach)	348 (0.1%)	185 (0.4%)	533 (0.1%)
Other (Child offence)	370 (0.1%)	51 (0.1%)	421 (0.1%)

	non-RF EM	RF EM	Total
	(N=319,862)	(N=52,115)	(N=371,977)
History of drug offences (N)	0.11 (0.54) 0: 0.00–21.00	0.16 (0.67) 0: 0.00–18.00	0.11 (0.56) 0: 0.00–21.00
History of weapons offences (N)	(0.10) 0: 0.00–12.00	(0.11) 0: 0.00–6.00	(0.10) 0: 0.00–12.00
History of public order offences (N)	(0.18) 0: 0.00–19.00	(0.22) 0: 0.00–16.00	(0.19) 0: 0.00–19.00
History of robbery (N)	(0.11) 0: 0.00–9.00	(0.12) 0: 0.00–6.00	(0.11) 0: 0.00–9.00
History of theft (N)	0.73 (2.34) 0: 0.00–75.00	1.27 (3.22) 0: 0.00–74.00	0.81 (2.49) 0: 0.00–75.00
History of Violence (N)	(0.26) 0: 0.00–18.00	(0.35) 0: 0.00–12.00	(0.27) 0: 0.00–18.00
History of summary offences (N)	1.13 (2.37) 0: 0.00–147.00	1.50 (2.85) 0: 0.00–76.00	1.18 (2.44) 0: 0.00–147.00
Other history (N)	0.09 (0.42) 0: 0.00–20.00	0.14 (0.53) 0: 0.00–13.00	0.10 (0.44) 0: 0.00–20.00
Prior prison sentences (N)	0.20 (1.11) 0: 0.00–43.00	0.34 (1.51) 0: 0.00–43.00	0.22 (1.18) 0: 0.00–43.00

	non-RF EM	RF EM	Total
	(N=319,862)	(N=52,115)	(N=371,977)
Number of offences in probation disposal	0.52 (1.00) 0: 0.00–22.00	0.51 (1.01) 0: 0.00–23.00	0.52 (1.00) 0: 0.00–23.00
Number of requirements	1.70 (0.87) 1: 1.00–13.00	1.76 (0.98) 1: 1.00–11.00	1.71 (0.89) 1: 1.00–13.00
Disposal length	322.97 (215.74) 364: 0.00–3,300.00	188.01 (179.15) 99: 0.00–2,280.00	304.06 (216.13) 364: 0.00–3,300.00
Electronic monitoring period		103.08 (94.32) 83: 0.00–2,848.00	

Year (requirement started)

	non-RF EM	RF EM	Total
	(N=319,862)	(N=52,115)	(N=371,977)
2014	76,296 (23.9%)	8,799 (16.9%)	85,095 (22.9%)
2015	71,325 (22.3%)	9,344 (17.9%)	80,669 (21.7%)
2016	58,540 (18.3%)	12,146 (23.3%)	70,686 (19.0%)

	non-RF EM	RF EM	Total
	(N=319,862)	**(N=52,115)**	**(N=371,977)**
2017	57,125 (17.9%)	11,025 (21.2%)	68,150 (18.3%)
2018	56,576 (17.7%)	10,801 (20.7%)	67,377 (18.1%)

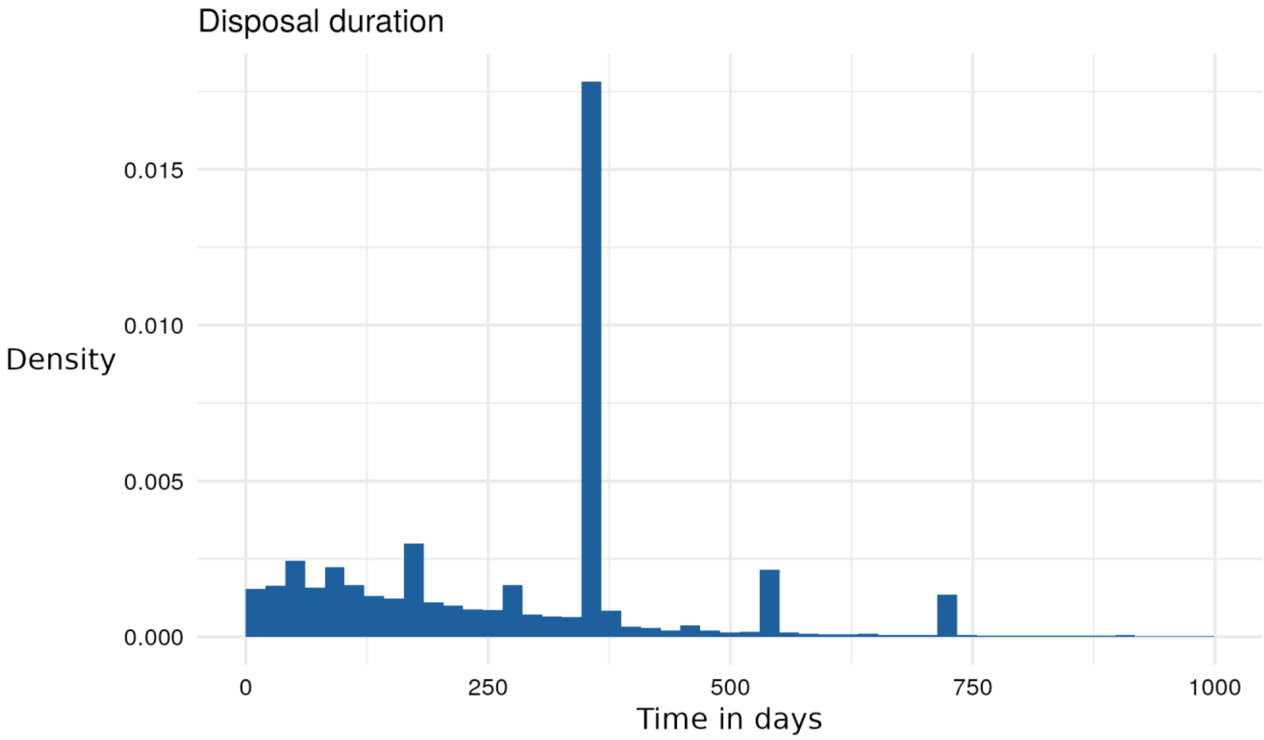
Offenders were mostly male (83%), from a white background (84%) and had a median age of 31. Sentences were predominantly for offences classified as violence against the person (34%), followed by theft (19%) and summary offences (16%).

Under 0.1% of records were for a breach (n=534), offences relating to children (421) or robbery (212).

The mean disposal length was 304 days and there was a concentration of disposals of one year in duration, as can be seen in figure 4.1.[24]

[24] A small number of probation records had disposal lengths recorded as greater than 1,000 days (n = 279 with RF EM, n = 7,213 in the control pool). Whilst these records were retained during the matching process, they were identified as outliers and omitted from the outcome analysis reported in Brunton-Smith (2025a; b).

Figure 4.1. Disposal length duration for probation records identified as a community order, 2014–18



A higher proportion of sentences that included a curfew requirement with RF EM were for theft (24% compared to 18%) and there were fewer violent offences (29% compared to 34%). In general, the distribution was similar for other offence types.

Offender age was also similar and there tended to be a similar number of probationary requirements (mean of 1.8 compared to 1.7) and additional offences being considered (both groups had a mean of 0.5).

By contrast, there was evidence of an over-representation of white British offenders subject to RF EM (83% compared to 75%) and fewer black (Caribbean and African) offenders in receipt of RF EM.

Sentences including a curfew requirement with RF EM were also for offenders characterised by longer and more serious offending histories, with higher average numbers of prior convictions for drug offences, theft, violence and summary offences, plus a higher number of prior custodial sentences (0.3 compared to 0.2).

Sentences including a curfew requirement with RF EM had substantially shorter overall disposal lengths than records without RF EM (a median of 99 days compared to 364). Offenders were monitored for approximately a median of 83 days. The proportion of offenders subject to a single requirement order was similar in the cohort of offenders whose sentence included RF EM and the cohort not subject to RF EM (49% compared to 50%).

The picture was very similar when the focus was restricted to offenders with a valid OASys record, albeit with a greater number of prior offences and higher mean number of prior custodial sentences (Appendix table A.1). The cohort of offenders with disposals starting between April 2016 and March 2017 (linked to PNC) was very similar to the overall cohort of offenders (see Appendix table A.2).

The PSM and CEM achieved a high level of balance with average (absolute) differences of 0.010 and 0.003, as shown in table 4.2. After matching, there was very little evidence of imbalance across individual variables. All probation records subject to RF EM could be matched using PSM. Match rates were lower for CEM at 70 per cent and analysis of the retained records indicates that the matched cohort had shorter offending histories and fewer offences being considered as part of the sentence.

Balance was similarly good when the subset of records with valid OASys risk profiles were included (see Appendix table A.3). The match rate using CEM was less than 10 per cent so matched comparisons were not explored for offenders with OASys records using the CEM approach.

Table 4.2. Balance tables for community orders, 2014–18

CEM

	Control (mean)	Control (SD)	RF EM (mean)	RF EM (sd)	Difference	Ratio	Overlap
Age (at offence)	32.189	10.329	32.199	10.336	0.001	1.001	0.004
Gender: Male	0.833	0.373	0.833	0.373	0.000	*NA*	0.000
Ethnicity: Asian Indian	0.005	0.068	0.005	0.068	0.000	*NA*	0.000
Ethnicity: Asian Pakistani	0.011	0.103	0.011	0.103	0.000	*NA*	0.000
Ethnicity: Asian Bangladeshi	0.002	0.040	0.002	0.040	0.000	*NA*	0.000
Ethnicity: Asian Chinese	0.000	0.013	0.000	0.014	0.002	*NA*	0.000
Ethnicity: Asian Other	0.002	0.048	0.002	0.047	0.000	*NA*	0.000
Ethnicity: Black Caribbean	0.012	0.111	0.012	0.111	0.000	*NA*	0.000
Ethnicity: Black African	0.009	0.096	0.009	0.096	0.000	*NA*	0.000
Ethnicity: Black Other	0.002	0.041	0.002	0.041	0.000	*NA*	0.000
White and Black Caribbean	0.010	0.097	0.010	0.097	0.000	*NA*	0.000
Ethnicity: White and Black African	0.001	0.033	0.001	0.033	0.000	*NA*	0.000
Ethnicity: White and Asian	0.001	0.029	0.001	0.029	0.000	*NA*	0.000
Ethnicity: Mixed Other	0.002	0.040	0.002	0.040	0.000	*NA*	0.000
Ethnicity: Arab	0.001	0.024	0.001	0.026	0.002	*NA*	0.000
Ethnicity: Other ethnicity	0.004	0.060	0.003	0.059	-0.001	*NA*	0.000
Ethnicity: White British	0.910	0.286	0.910	0.286	0.000	*NA*	0.000
Ethnicity: White Irish	0.002	0.047	0.002	0.047	0.000	*NA*	0.000
Ethnicity: White Roma	0.002	0.042	0.002	0.042	0.000	*NA*	0.000
Ethnicity: White Other	0.026	0.159	0.026	0.159	0.000	*NA*	0.000

	Control (mean)	Control (SD)	RF EM (mean)	RF EM (sd)	Difference	Ratio	Overlap
Index offence: Criminal damage	0.003	0.058	0.003	0.058	0.000	*NA*	0.000
Index offence: Drugs	0.046	0.210	0.046	0.210	0.000	*NA*	0.000
Index offence: Fraud	0.028	0.166	0.028	0.166	0.000	*NA*	0.000
Index offence: Miscellaneous	0.029	0.167	0.029	0.167	0.000	*NA*	0.000
Index offence: Weapons	0.001	0.027	0.001	0.027	0.000	*NA*	0.000
Index offence: Public order	0.026	0.160	0.026	0.160	0.000	*NA*	0.000
Index offence: Robbery	0.000	0.007	0.000	0.007	0.000	*NA*	0.000
Index offence: Sex offences	0.003	0.055	0.003	0.055	0.000	*NA*	0.000
Index offence: Summary	0.180	0.385	0.180	0.385	0.000	*NA*	0.000
Index offence: Summary (motoring)	0.119	0.324	0.119	0.324	0.000	*NA*	0.000
Index offence: Theft	0.234	0.424	0.234	0.424	0.000	*NA*	0.000
Index offence: Violence	0.328	0.470	0.328	0.470	0.000	*NA*	0.000
Index offence: Other (Breach)	0.001	0.028	0.001	0.028	0.000	*NA*	0.000
Index offence: Other (Child offence)	0.000	0.017	0.000	0.017	0.000	*NA*	0.000
History of drug offences (N)	0.055	0.287	0.061	0.296	0.009	1.063	0.010
History of weapons offences (N)	0.001	0.026	0.001	0.026	0.000	1.000	0.089
History of public order offences (N)	0.000	0.022	0.000	0.022	0.000	1.000	0.046
History of robbery (N)	0.001	0.024	0.001	0.024	0.000	1.000	0.071
History of theft (N)	0.653	1.733	0.681	1.749	0.009	1.018	0.039
History of Violence (N)	0.015	0.138	0.015	0.138	0.000	1.000	0.045
History of summary offences (N)	0.969	1.588	1.110	1.666	0.050	1.101	0.150
Other history (N)	0.034	0.206	0.034	0.206	0.000	1.000	0.011

	Control (mean)	Control (SD)	RF EM (mean)	RF EM (sd)	Difference	Ratio	Overlap
Prior prison sentences (N)	0.084	0.460	0.084	0.457	0.000	0.985	0.000
Number of offences in probation disposal	0.367	0.703	0.356	0.701	-0.011	0.994	0.017
Number of requirements	1.667	0.862	1.667	0.862	0.000	1.000	0.003
Disposal length	195.752	159.206	182.239	163.033	-0.075	1.049	0.150
Year (requirement started): 2014	0.193	0.395	0.193	0.395	0.000	*NA*	0.000
Year (requirement started): 2015	0.195	0.396	0.195	0.396	0.000	*NA*	0.000
Year (requirement started): 2016	0.226	0.418	0.226	0.418	0.000	*NA*	0.000
Year (requirement started): 2017	0.200	0.400	0.200	0.400	0.000	*NA*	0.000
Year (requirement started): 2018	0.186	0.389	0.186	0.389	0.000	*NA*	0.000
PSM (distance)	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
Average					0.003		

	Control (mean)	Control (SD)
All (Effective Sample Size)	319,696	52,114
All (Unweighted)	319,696	52,114
Matched (Effective Sample Size)	48350	36,705
Matched (Unweighted)	125,731	36,705
Unmatched	193,965	15,409
Match rate		70%

PSM

	Control (mean)	Control (SD)	RF EM (mean)	RF EM (sd)	Difference	Ratio	Overlap
Age (at offence)	32.647	10.606	32.824	10.719	0.017	1.022	0.010
Gender: Male	0.827	0.379	0.824	0.381	-0.007	*NA*	0.003
Ethnicity: Asian Indian	0.009	0.093	0.009	0.096	0.006	*NA*	0.001
Ethnicity: Asian Pakistani	0.018	0.134	0.019	0.136	0.005	*NA*	0.001
Ethnicity: Asian Bangladeshi	0.004	0.064	0.005	0.069	0.010	*NA*	0.001
Ethnicity: Asian Chinese	0.000	0.012	0.000	0.018	0.009	*NA*	0.000
Ethnicity: Asian Other	0.005	0.072	0.006	0.076	0.009	*NA*	0.001
Ethnicity: Black Caribbean	0.023	0.151	0.025	0.156	0.011	*NA*	0.002
Ethnicity: Black African	0.017	0.129	0.018	0.133	0.008	*NA*	0.001
Ethnicity: Black Other	0.006	0.075	0.006	0.076	0.002	*NA*	0.000
Ethnicity: White and Black Caribbean	0.019	0.135	0.018	0.135	-0.002	*NA*	0.000
Ethnicity: White and Black African	0.004	0.060	0.004	0.061	0.003	*NA*	0.000
Ethnicity: White and Asian	0.004	0.064	0.004	0.062	-0.004	*NA*	0.000
Ethnicity: Mixed Other	0.005	0.069	0.005	0.070	0.002	*NA*	0.000
Ethnicity: Arab	0.001	0.036	0.001	0.037	0.001	*NA*	0.000
Ethnicity: Other ethnicity	0.006	0.080	0.007	0.082	0.004	*NA*	0.000
Ethnicity: White British	0.836	0.370	0.829	0.377	-0.021	*NA*	0.008
Ethnicity: White Irish	0.008	0.090	0.008	0.091	0.003	*NA*	0.000
Ethnicity: White Roma	0.006	0.076	0.006	0.076	-0.001	*NA*	0.000
Ethnicity: White Other	0.029	0.167	0.030	0.172	0.011	*NA*	0.002
Index offence: Criminal damage	0.006	0.076	0.006	0.075	-0.002	*NA*	0.000

	Control (mean)	Control (SD)	RF EM (mean)	RF EM (sd)	Difference	Ratio	Overlap
Index offence: Drugs	0.057	0.232	0.058	0.234	0.006	*NA*	0.001
Index offence: Fraud	0.032	0.177	0.033	0.179	0.005	*NA*	0.001
Index offence: Miscellaneous	0.039	0.194	0.039	0.194	0.000	*NA*	0.000
Index offence: Weapons	0.002	0.046	0.002	0.044	-0.004	*NA*	0.000
Index offence: Public order	0.031	0.172	0.031	0.174	0.003	*NA*	0.001
Index offence: Robbery	0.001	0.023	0.001	0.023	0.000	*NA*	0.000
Index offence: Sex offences	0.007	0.084	0.006	0.078	-0.013	*NA*	0.001
Index offence: Summary	0.166	0.372	0.169	0.375	0.007	*NA*	0.003
Index offence: Summary (motoring)	0.108	0.310	0.114	0.318	0.020	*NA*	0.006
Index offence: Theft	0.248	0.432	0.244	0.430	-0.010	*NA*	0.004
Index offence: Violence	0.300	0.458	0.293	0.455	-0.016	*NA*	0.007
Index offence: Other (Breach)	0.003	0.052	0.004	0.059	0.015	*NA*	0.001
Index offence: Other (Child offence)	0.001	0.031	0.001	0.031	0.000	*NA*	0.000
History of drug offences (N)	0.154	0.671	0.158	0.675	0.006	1.011	0.010
History of weapons offences (N)	0.008	0.144	0.008	0.112	-0.001	0.609	0.090
History of public order offences (N)	0.016	0.253	0.015	0.219	-0.002	0.743	0.047
History of robbery (N)	0.008	0.120	0.008	0.117	-0.001	0.943	0.071
History of theft (N)	1.272	3.159	1.269	3.218	-0.001	1.038	0.037
History of Violence (N)	0.050	0.338	0.054	0.348	0.012	1.059	0.046
History of summary offences (N)	1.444	3.112	1.499	2.851	0.019	0.839	0.147
Other history (N)	0.132	0.532	0.136	0.533	0.007	1.004	0.012
Prior prison sentences (N)	0.356	1.464	0.339	1.508	-0.012	1.062	0.007

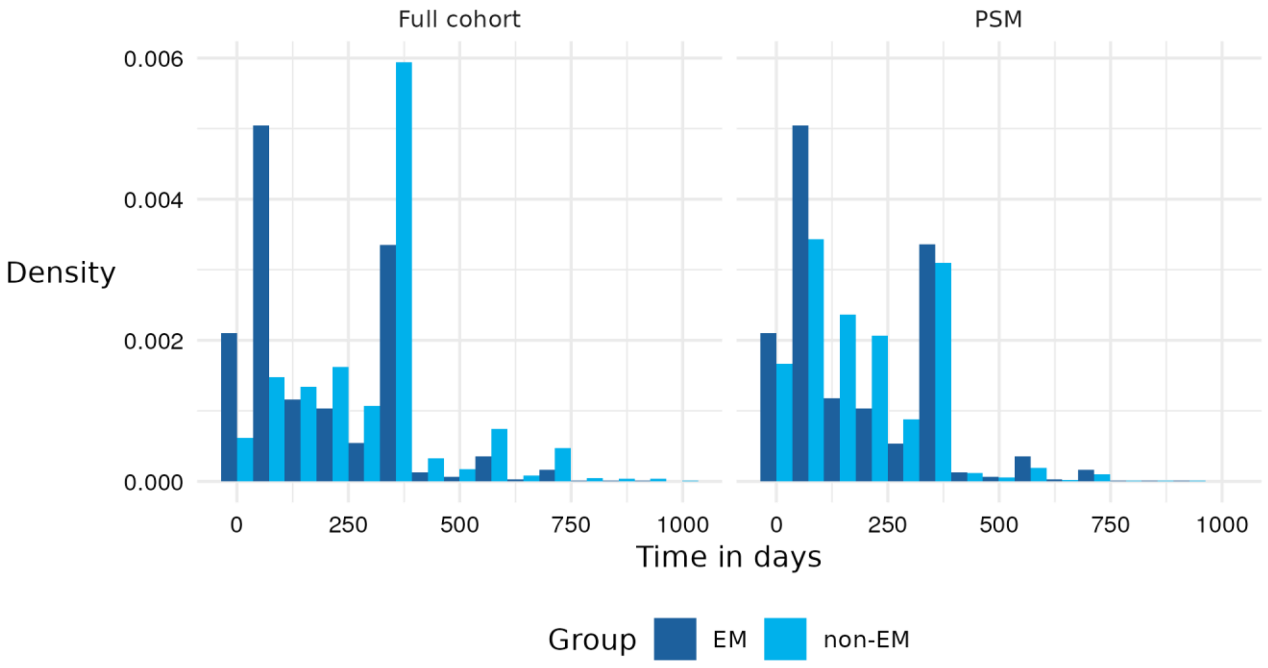
	Control (mean)	Control (SD)	RF EM (mean)	RF EM (sd)	Difference	Ratio	Overlap
Number of offences in probation disposal	0.539	1.037	0.514	1.010	-0.025	0.949	0.020
Number of requirements	1.859	0.990	1.760	0.983	-0.100	0.986	0.083
Disposal length	193.973	156.264	188.011	179.149	-0.033	1.314	0.183
Year (requirement started): 2014	0.176	0.381	0.169	0.375	-0.018	*NA*	0.007
Year (requirement started): 2015	0.188	0.390	0.179	0.384	-0.022	*NA*	0.008
Year (requirement started): 2016	0.216	0.412	0.233	0.423	0.039	*NA*	0.017
Year (requirement started): 2017	0.213	0.409	0.212	0.408	-0.003	*NA*	0.001
Year (requirement started): 2018	0.208	0.406	0.207	0.405	-0.001	*NA*	0.000
PSM (distance)	0.237	0.130	0.237	0.130	0.000	1.000	0.000
Average					0.010		

	Control (mean)	Control (SD)
All (Effective Sample Size)	319,696	52,114
All (Unweighted)	319,696	52,114
Matched (Effective Sample Size)	52,114	52,114
Matched (Unweighted)	52,114	52,114
Unmatched	267,582	0
Match rate		100%

Importantly, to assess the impact of RF EM during the monitoring period and across the whole disposal period, the disposal lengths for treated and control units must be comparable.[25]

Figure 4.2 confirms the over-representation of comparatively short duration sentences with RF EM curfew conditions in the full cohort ("unmatched"). This is likely to reflect the larger proportion of single requirement orders involving RF EM, where the offender was only in receipt of a single requirement. Following the application of PSM, there is a closer correspondence between the treated and control groups in the "PSM" sample. However, there is still evidence of a modest overrepresentation of RF EM offenders with shorter disposals.

Figure 4.2. Full Cohort and PSM comparisons of disposal length duration for probation records identified as a community order, 2014–18

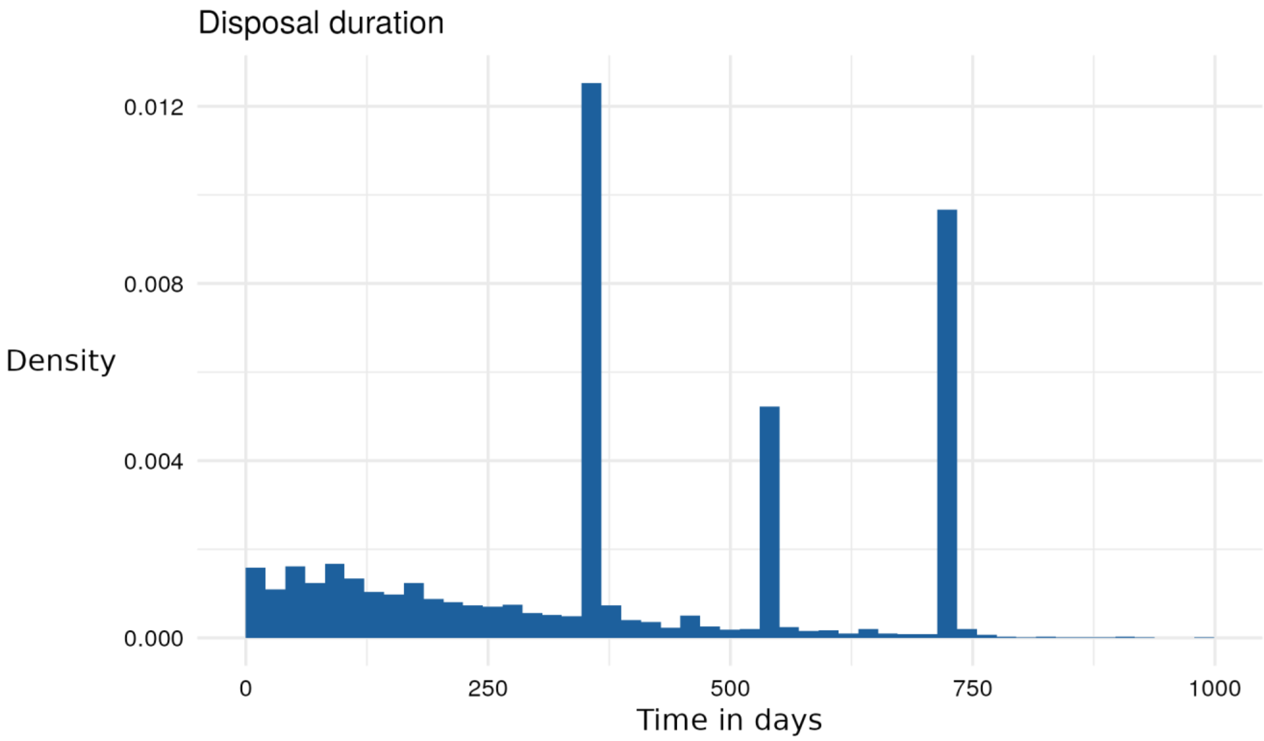


[25] It is not possible to also ensure comparable monitoring durations between the treatment and control groups because offenders in the control group have no monitoring duration data.

4.2 Suspended sentence orders

From a total of 191,384 probation records covering suspended sentence orders, the most common index offence types were violence against the person (33%) and theft (17%), while more than 85 per cent were male offenders. Disposal lengths were moderately longer, on average, than community orders (mean of 390 days), and there was evidence of concentrations of disposal lengths around one and two years, as shown in figure 4.3.[26]

Figure 4.3. Disposal length duration for probation records identified as a suspended sentence order, 2014–18



The offence profile was similar for the subset of sentences including a curfew requirement with RF EM (n=25,858) and the records where RF EM was not included (n=165,526). Differences in prior offending and disposal length were also modest in size. Offenders were monitored for approximately four months (median 140 days).

[26] Very few probation records had disposal lengths recorded as greater than 1,000 days (12 with RF EM, 129 from the control pool). Whilst these records were retained during the matching process they were identified as outliers and omitted from the outcome analysis reported in Brunton-Smith (2025a; b).

Table 4.3. Descriptive statistics for suspended sentence orders, 2014–18[27]

	non-RF EM	RF EM	Total
	(N=165,526)	(N=25,858)	(N=191,384)
Age (at offence)	32.74 (10.76) 31: 18.00–88.00	31.56 (11.05) 29: 18.00–86.00	32.58 (10.81) 30: 18.00–88.00

Gender

	non-RF EM	RF EM	Total
	(N=165,526)	(N=25,858)	(N=191,384)
Female (reference)	23,464 (14.2%)	3,542 (13.7%)	27,006 (14.1%)
Male	142,062 (85.8%)	22,316 (86.3%)	164,378 (85.9%)

Ethnicity

	non-RF EM	RF EM	Total
	(N=165,526)	(N=25,858)	(N=191,384)
Asian Indian	2,530 (1.5%)	313 (1.2%)	2,843 (1.5%)
Asian Pakistani	4,214 (2.5%)	617 (2.4%)	4,831 (2.5%)
Asian Bangladeshi	1,536 (0.9%)	209 (0.8%)	1,745 (0.9%)
Asian Chinese	216 (0.1%)	16 (0.1%)	232 (0.1%)
Asian Other	1,794 (1.1%)	203 (0.8%)	1,997 (1.0%)
Black Caribbean	5,753 (3.5%)	794 (3.1%)	6,547 (3.4%)
Black African	4,871 (2.9%)	618 (2.4%)	5,489 (2.9%)
Black Other	1,481 (0.9%)	221 (0.9%)	1,702 (0.9%)
White and Black Caribbean	3,518 (2.1%)	582 (2.3%)	4,100 (2.1%)
White and Black African	752 (0.5%)	115 (0.4%)	867 (0.5%)
White and Asian	636 (0.4%)	117 (0.5%)	753 (0.4%)
Mixed Other	1,133 (0.7%)	157 (0.6%)	1,290 (0.7%)
Arab	398 (0.2%)	50 (0.2%)	448 (0.2%)
Other ethnicity	2,051 (1.2%)	203 (0.8%)	2,254 (1.2%)
White British	123,200 (74.4%)	20,404 (78.9%)	143,604 (75.0%)
White Irish	1,348 (0.8%)	212 (0.8%)	1,560 (0.8%)
White Roma	887 (0.5%)	174 (0.7%)	1,061 (0.6%)

[27] Numeric variables report: mean, (sd), median, min-max. Qualitative variables report N (%)

	non-RF EM	RF EM	Total
	(N=165,526)	(N=25,858)	(N=191,384)
White Other	9,208 (5.6%)	853 (3.3%)	10,061 (5.3%)

Index offence

	non-RF EM	RF EM	Total
	(N=165,526)	(N=25,858)	(N=191,384)
Criminal damage	1,002 (0.6%)	143 (0.6%)	1,145 (0.6%)
Drugs	15,555 (9.4%)	3,114 (12.0%)	18,669 (9.8%)
Fraud	8,740 (5.3%)	1,249 (4.8%)	9,989 (5.2%)
Miscellaneous	12,057 (7.3%)	1,383 (5.3%)	13,440 (7.0%)
Weapons	616 (0.4%)	112 (0.4%)	728 (0.4%)
Public order	7,139 (4.3%)	1,351 (5.2%)	8,490 (4.4%)
Robbery	1,035 (0.6%)	255 (1.0%)	1,290 (0.7%)
Sex offences	2,437 (1.5%)	277 (1.1%)	2,714 (1.4%)
Summary	14,404 (8.7%)	2,044 (7.9%)	16,448 (8.6%)
Summary (motoring)	18,227 (11.0%)	3,002 (11.6%)	21,229 (11.1%)
Theft	28,686 (17.3%)	4,781 (18.5%)	33,467 (17.5%)
Violence	54,841 (33.1%)	8,040 (31.1%)	62,881 (32.9%)
Other (Breach)	215 (0.1%)	45 (0.2%)	260 (0.1%)
Other (Child offence)	572 (0.3%)	62 (0.2%)	634 (0.3%)

	non-RF EM	RF EM	Total
	(N=165,526)	(N=25,858)	(N=191,384)
History of drug offences (N)	0.12 (0.56) 0: 0.00–30.00	0.16 (0.64) 0: 0.00–11.00	0.13 (0.57) 0: 0.00–30.00
History of weapons offences (N)	0.01 (0.12) 0: 0.00–12.00	0.01 (0.13) 0: 0.00–6.00	0.01 (0.12) 0: 0.00–12.00
History of public order offences (N)	0.01 (0.19) 0: 0.00–35.00	0.01 (0.22) 0: 0.00–22.00	0.01 (0.20) 0: 0.00–35.00
History of robbery (N)	0.01 (0.12) 0: 0.00–7.00	0.01 (0.16) 0: 0.00–9.00	0.01 (0.13) 0: 0.00–9.00
History of theft (N)	0.92 (2.82) 0: 0.00–74.00	0.96 (2.87) 0: 0.00–42.00	0.92 (2.83) 0: 0.00–74.00

	non-RF EM	RF EM	Total
	(N=165,526)	**(N=25,858)**	**(N=191,384)**
History of Violence (N)	0.06 (0.34) 0: 0.00–21.00	0.06 (0.37) 0: 0.00–15.00	0.06 (0.35) 0: 0.00–21.00
History of summary offences (N)	1.12 (2.45) 0: 0.00–163.00	1.23 (2.58) 0: 0.00–79.00	1.14 (2.47) 0: 0.00–163.00
Other history (N)	0.12 (0.49) 0: 0.00–17.00	0.14 (0.55) 0: 0.00–18.00	0.12 (0.50) 0: 0.00–18.00
Prior prison sentences (N)	0.31 (1.43) 0: 0.00–45.00	0.34 (1.58) 0: 0.00–42.00	0.31 (1.45) 0: 0.00–45.00

	non-RF EM	RF EM	Total
	(N=165,526)	**(N=25,858)**	**(N=191,384)**
Number of offences in probation disposal	0.73 (1.30) 0: 0.00–32.00	0.73 (1.27) 0: 0.00–19.00	0.73 (1.30) 0: 0.00–32.00
Number of requirements	1.81 (0.89) 2: 1.00–14.00	2.25 (1.09) 2: 1.00–11.00	1.87 (0.93) 2: 1.00–14.00
Disposal length	395.17 (227.78) 364: 0.00–2,909.00	358.99 (256.53) 364: 0.00–1,729.00	390.28 (232.20) 364: 0.00–2,909.00
Electronic monitoring period		179.97 (151.49) 140: 0.00–1,830.00	

Year (requirement started)

	non-RF EM	RF EM	Total
	(N=165,526)	**(N=25,858)**	**(N=191,384)**
2014	38,052 (23.0%)	4,170 (16.1%)	42,222 (22.1%)
2015	37,650 (22.7%)	4,776 (18.5%)	42,426 (22.2%)
2016	32,945 (19.9%)	6,528 (25.2%)	39,473 (20.6%)
2017	32,261 (19.5%)	5,955 (23.0%)	38,216 (20.0%)
2018	24,618 (14.9%)	4,429 (17.1%)	29,047 (15.2%)

Offenders receiving a sentence including a curfew order with RF EM tended to have a slightly younger age profile (median age 29 compared to 31).

Similar to the data for community orders, there was an over-representation of white British offenders (79% compared to 74%).

Offenders subject to a suspended sentence order may also have additional requirements attached to their sentence. Records with linked RF EM data tended to have a greater number of additional sentence requirements (2.3 compared to 1.8). Approximately 74 per cent of records including RF EM were single requirement orders, compared to 57 per cent for records that did not involve RF EM. The median length of time an offender was monitored was longer than for community orders.

The subset of records that could be linked to an OASys record were again characterised by longer offending histories (e.g., twice as many thefts on average and double the number of custodial spells), but this does not differ by RF EM status (Appendix table A.4). The cohort with disposals starting between April 2016 and March 2017 look very similar to the full cohort on all variables (see Appendix table A.5).

The PSM and CEM again achieved a high level of balance with average (absolute) differences of 0.007 and 0.001, as can be seen in table 4.4. There was very little evidence of imbalance across individual variables. Nearly all probation records with an RF EM condition could be matched using PSM (10 offenders subject to RF EM were dropped).[28] The match rate was lower using CEM at 61 per cent. Examination of the retained records indicates that the matched cohort had shorter offending histories and fewer offences being considered as part of the sentence.

---

[28] Balance was also good when the subset of records with valid OASys risk profiles were included (see appendix table A.6).

Table 4.4. Balance tables for suspended sentence orders, 2014–18

CEM

	Control (mean)	Control (SD)	RF EM (mean)	RF EM (sd)	Difference	Ratio	Overlap
Age (at offence)	30.726	10.335	30.697	10.372	-0.003	1.007	0.005
Gender: Male	0.889	0.314	0.889	0.314	0.000	*NA*	0.000
Ethnicity: Asian Indian	0.003	0.050	0.003	0.050	0.000	*NA*	0.000
Ethnicity: Asian Pakistani	0.011	0.105	0.011	0.105	0.000	*NA*	0.000
Ethnicity: Asian Bangladeshi	0.002	0.041	0.002	0.041	0.000	*NA*	0.000
Ethnicity: Asian Chinese	0.000	0.020	0.000	0.020	-0.001	*NA*	0.000
Ethnicity: Asian Other	0.003	0.051	0.003	0.052	0.000	*NA*	0.000
Ethnicity: Black Caribbean	0.014	0.116	0.014	0.116	0.000	*NA*	0.000
Ethnicity: Black African	0.011	0.105	0.011	0.105	0.000	*NA*	0.000
Ethnicity: Black Other	0.002	0.040	0.002	0.040	0.000	*NA*	0.000
Ethnicity: White and Black Caribbean	0.008	0.090	0.008	0.090	0.000	*NA*	0.000
Ethnicity: White and Black African	0.000	0.016	0.000	0.016	0.000	*NA*	0.000
Ethnicity: White and Asian	0.001	0.025	0.001	0.025	0.000	*NA*	0.000
Ethnicity: Mixed Other	0.001	0.034	0.001	0.034	0.000	*NA*	0.000
Ethnicity: Arab	0.001	0.030	0.001	0.028	-0.003	*NA*	0.000
Ethnicity: Other ethnicity	0.003	0.056	0.003	0.057	0.001	*NA*	0.000
Ethnicity: White British	0.916	0.278	0.916	0.278	0.000	*NA*	0.000
Ethnicity: White Irish	0.001	0.030	0.001	0.030	0.000	*NA*	0.000
Ethnicity: White Roma	0.001	0.033	0.001	0.033	0.000	*NA*	0.000
Ethnicity: White Other	0.023	0.150	0.023	0.150	0.000	*NA*	0.000

	Control (mean)	Control (SD)	RF EM (mean)	RF EM (sd)	Difference	Ratio	Overlap
Index offence: Criminal damage	0.002	0.046	0.002	0.046	0.000	*NA*	0.000
Index offence: Drugs	0.122	0.328	0.122	0.328	0.000	*NA*	0.000
Index offence: Fraud	0.035	0.184	0.035	0.184	0.000	*NA*	0.000
Index offence: Miscellaneous	0.043	0.204	0.043	0.204	0.000	*NA*	0.000
Index offence: Public order	0.049	0.216	0.049	0.216	0.000	*NA*	0.000
Index offence: Robbery	0.005	0.067	0.005	0.067	0.000	*NA*	0.000
Index offence: Sex offences	0.008	0.089	0.008	0.089	0.000	*NA*	0.000
Index offence: Summary	0.076	0.265	0.076	0.265	0.000	*NA*	0.000
Index offence: Summary (motoring)	0.110	0.313	0.110	0.313	0.000	*NA*	0.000
Index offence: Theft	0.169	0.375	0.169	0.375	0.000	*NA*	0.000
Index offence: Violence	0.378	0.485	0.378	0.485	0.000	*NA*	0.000
Index offence: Weapons	0.001	0.033	0.001	0.033	0.000	*NA*	0.000
Index offence: Other (Breach)	0.000	0.018	0.000	0.018	0.000	*NA*	0.000
Index offence: Other (Child offence)	0.001	0.030	0.001	0.030	0.000	*NA*	0.000
History of drug offences (N)	0.063	0.290	0.064	0.292	0.001	1.017	0.043
History of weapons offences (N)	0.001	0.038	0.001	0.038	0.000	1.000	0.104
History of public order offences (N)	0.004	0.065	0.005	0.072	0.005	1.235	0.124
History of robbery (N)	0.001	0.025	0.001	0.025	0.000	1.000	0.017
History of theft (N)	0.476	1.462	0.448	1.451	-0.010	0.985	0.017
History of Violence (N)	0.037	0.199	0.035	0.195	-0.005	0.954	0.038
History of summary offences (N)	0.949	1.586	0.908	1.586	-0.016	1.000	0.173
Other history (N)	0.025	0.172	0.025	0.172	0.000	1.000	0.006

	Control (mean)	Control (SD)	RF EM (mean)	RF EM (sd)	Difference	Ratio	Overlap
Prior prison sentences (N)	0.094	0.518	0.085	0.511	-0.005	0.975	0.010
Number of offences in probation disposal	0.526	0.888	0.514	0.885	-0.010	0.993	0.030
Number of requirements	2.119	0.926	2.119	0.926	0.000	1.000	0.001
Disposal length	367.248	255.266	367.272	253.585	0.000	0.987	0.034
Year (requirement started): 2014	0.190	0.393	0.190	0.393	0.000	*NA*	0.000
Year (requirement started): 2015	0.205	0.403	0.205	0.403	0.000	*NA*	0.000
Year (requirement started): 2016	0.246	0.431	0.246	0.431	0.000	*NA*	0.000
Year (requirement started): 2017	0.214	0.410	0.214	0.410	0.000	*NA*	0.000
Year (requirement started): 2018	0.145	0.352	0.145	0.352	0.000	*NA*	0.000
PSM (distance)	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
Average					0.001		

	Control (mean)	Control (SD)
All (Effective Sample Size)	165,439	25,854
All (Unweighted)	165,439	25,854
Matched (Effective Sample Size)	23765	15,773
Matched (Unweighted)	51,280	15,773
Unmatched	114,159	10,081
Match rate		61%

PSM

	Control (mean)	Control (SD)	RF EM (mean)	RF EM (sd)	Difference	Ratio	Overlap
Age (at offence)	31.622	10.260	31.564	11.055	-0.005	1.161	0.051
Gender: Male	0.858	0.349	0.863	0.344	0.013	*NA*	0.005
Ethnicity: Asian Indian	0.012	0.108	0.012	0.109	0.004	*NA*	0.000
Ethnicity: Asian Pakistani	0.022	0.148	0.024	0.153	0.010	*NA*	0.002
Ethnicity: Asian Bangladeshi	0.008	0.090	0.008	0.090	-0.001	*NA*	0.000
Ethnicity: Asian Chinese	0.001	0.022	0.001	0.025	0.005	*NA*	0.000
Ethnicity: Asian Other	0.007	0.085	0.008	0.088	0.006	*NA*	0.001
Ethnicity: Black Caribbean	0.030	0.171	0.031	0.173	0.003	*NA*	0.001
Ethnicity: Black African	0.023	0.150	0.024	0.153	0.005	*NA*	0.001
Ethnicity: Black Other	0.008	0.087	0.009	0.092	0.009	*NA*	0.001
Ethnicity: White and Black Caribbean	0.022	0.146	0.023	0.148	0.004	*NA*	0.001
Ethnicity: White and Black African	0.004	0.065	0.004	0.066	0.003	*NA*	0.000
Ethnicity: White and Asian	0.005	0.067	0.005	0.067	0.000	*NA*	0.000
Ethnicity: Mixed Other	0.006	0.076	0.006	0.078	0.004	*NA*	0.000
Ethnicity: Arab	0.002	0.046	0.002	0.044	-0.004	*NA*	0.000
Ethnicity: Other ethnicity	0.007	0.083	0.008	0.088	0.010	*NA*	0.001
Ethnicity: White British	0.801	0.400	0.789	0.408	-0.028	*NA*	0.012
Ethnicity: White Irish	0.008	0.089	0.008	0.090	0.003	*NA*	0.000
Ethnicity: White Roma	0.007	0.083	0.007	0.082	-0.002	*NA*	0.000
Ethnicity: White Other	0.028	0.165	0.033	0.179	0.028	*NA*	0.005
Index offence: Criminal damage	0.005	0.070	0.006	0.074	0.009	*NA*	0.001

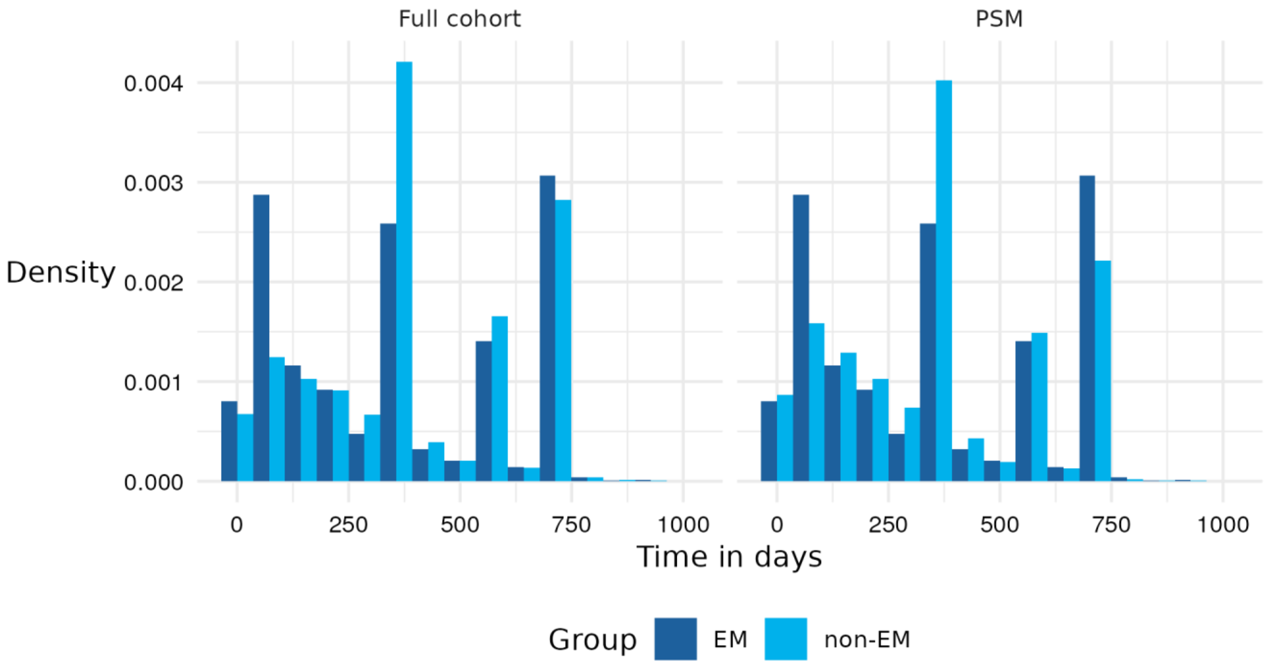
	Control (mean)	Control (SD)	RF EM (mean)	RF EM (sd)	Difference	Ratio	Overlap
Index offence: Drugs	0.119	0.324	0.120	0.325	0.003	*NA*	0.001
Index offence: Fraud	0.049	0.216	0.048	0.214	-0.003	*NA*	0.001
Index offence: Miscellaneous	0.051	0.220	0.053	0.225	0.011	*NA*	0.003
Index offence: Public order	0.050	0.219	0.052	0.223	0.009	*NA*	0.002
Index offence: Robbery	0.009	0.092	0.010	0.099	0.013	*NA*	0.001
Index offence: Sex offences	0.010	0.098	0.011	0.103	0.010	*NA*	0.001
Index offence: Summary	0.080	0.271	0.079	0.270	-0.003	*NA*	0.001
Index offence: Summary (motoring)	0.114	0.318	0.116	0.320	0.006	*NA*	0.002
Index offence: Theft	0.190	0.393	0.185	0.388	-0.014	*NA*	0.006
Index offence: Violence	0.314	0.464	0.311	0.463	-0.006	*NA*	0.003
Index offence: Weapons	0.005	0.069	0.004	0.066	-0.006	*NA*	0.000
Index offence: Other (Breach)	0.001	0.038	0.002	0.042	0.006	*NA*	0.000
Index offence: Other (Child offence)	0.002	0.047	0.002	0.049	0.004	*NA*	0.000
History of drug offences (N)	0.157	0.668	0.155	0.637	-0.002	0.908	0.044
History of weapons offences (N)	0.013	0.178	0.013	0.133	0.000	0.560	0.105
History of public order offences (N)	0.011	0.156	0.013	0.220	0.009	1.997	0.124
History of robbery (N)	0.013	0.151	0.013	0.165	0.000	1.194	0.018
History of theft (N)	0.991	2.734	0.958	2.873	-0.012	1.104	0.019
History of Violence (N)	0.063	0.368	0.062	0.372	-0.003	1.019	0.038
History of summary offences (N)	1.235	2.395	1.232	2.580	-0.001	1.160	0.170
Other history (N)	0.144	0.551	0.137	0.548	-0.013	0.989	0.009
Prior prison sentences (N)	0.349	1.469	0.340	1.578	-0.006	1.153	0.016

	Control (mean)	Control (SD)	RF EM (mean)	RF EM (sd)	Difference	Ratio	Overlap
Number of offences in probation disposal	0.729	1.238	0.730	1.268	0.001	1.049	0.028
Number of requirements	2.234	1.058	2.249	1.079	0.014	1.041	0.050
Disposal length	357.655	224.750	358.978	256.530	0.005	1.303	0.152
Year (requirement started): 2014	0.156	0.363	0.161	0.368	0.013	*NA*	0.005
Year (requirement started): 2015	0.186	0.389	0.185	0.388	-0.004	*NA*	0.001
Year (requirement started): 2016	0.250	0.433	0.252	0.434	0.006	*NA*	0.003
Year (requirement started): 2017	0.230	0.421	0.230	0.421	0.000	*NA*	0.000
Year (requirement started): 2018	0.177	0.382	0.171	0.377	-0.016	*NA*	0.006
PSM (distance)	0.182	0.098	0.182	0.098	0.000	1.001	0.000
Average					0.007		

	Control (mean)	Control (SD)
All (Effective Sample Size)	165,439	25,854
All (Unweighted)	165,439	25,854
Matched (Effective Sample Size)	25,844	25,844
Matched (Unweighted)	25,844	25,844
Unmatched	139,595	10
Match rate		100%

Figure 4.4 shows a good correspondence between the distribution of disposal lengths for the treated and control groups for the “PSM” sample.

Figure 4.4. Full Cohort and PSM comparisons of disposal length duration for probation records identified as a suspended sentence order, 2014–18



5. Limitations

There are important limitations with a quasi-experimental approach to causal estimation.

First, the research is reliant on the assumption that there are no unobserved confounders that are associated with treatment and associated with reoffending. This assumption is untestable and it remains possible that important confounders are missing. Whilst the list of explicit variables covers offending history, offence type and basic demographics, less can be said about offender need. This may be particularly important in this instance when we think about housing need, with fixed accommodation a requirement for a curfew requirement with EM, and those with unstable accommodation more likely to offend.

Importantly, the results reported in Brunton-Smith (2025b) were robust to the inclusion of offender assessed needs from OASys reports covering accommodation, employment, relationships, lifestyle and associates, drug misuse, alcohol misuse, thinking and behaviour and attitudes. This is consistent with earlier work from Eaton and Mews (2019) which showed little difference in the effectiveness of community sentences when compared to short custodial sentences with and without taking account of OASys risks.

However, OASys information was only available for a subset of offenders that tended to have more serious offending histories (82% of those with a valid OASys record spent time in custody since 2011, compared to 42% of those without OASys). OASys assessments are also less likely to be available for first time offenders before sentencing. As a result, whilst the consistency across approaches can be taken as a good indication that the main confounders have been effectively captured, all evaluation results should be judged with caution.

Second, results reported in Brunton-Smith (2025a; b) may also be susceptible to selection bias. In particular, the current official sentencing guidelines identify a number of factors that may lead an individual to be considered ineligible for EM: a lack of consent from someone (other than the offender) required for EM to be installed; a lack of availability of EM; and particular case circumstances. No relevant information was available about these factors to enable records to be excluded from the control group prior to estimation.

However, it remains possible that some offenders deemed ineligible for EM are being included in the comparison and as result, measures of EM effectiveness may be under (or over) estimated.

Third, results may also have been affected by post-treatment bias. In order effectively to examine the impact of RF EM during community sentence disposals, it was necessary to include disposal length in the list of confounder variables. The current sentencing guidelines require sentencers to make an initial judgement on the disposal length in Step One of the sentencing process (when considering harm and culpability), prior to determining additional conditions like the use of RF EM. However, this can be adjusted at later guideline steps and so it remains possible that the total disposal length is, in part, determined by whether EM is used. Similarly, the number of community sentence requirements imposed may also be partially determined by whether to use EM. Consequently, reported estimates of impact reported in Brunton-Smith (2025a; b) may be biased.

Fourth, the analyses reported in Brunton-Smith (2025a; b) are susceptible to researcher dependency effects. Throughout the analysis a range of important decisions have been required – in relation to selection of confounders, choice of matching approach and closeness of matches, whether matching with or without replacement, assessment of balance plus selection of CML model and tuning of hyperparameters. Different choices could plausibly have resulted in different conclusions.

To mitigate this risk, Brunton-Smith (2025b) reports results from across a range of different matching algorithms and estimation approaches, with results showing a good degree of consistency across different approaches. However, it remains possible that a different researcher could have taken a different approach and thus produced different conclusions.

There were also limitations with the use of data collected for routine administrative purposes that cannot be ignored.

All analysis must assume that records were an accurate reflection of the true underlying processes being measured. In some instances, data inaccuracies can be identified and corrected prior to analysis, and a careful process of data screening was undertaken to minimise the impact of measurement error. For example, disposal start and termination dates were user input by probation practitioners or their admin staff and consequently are subject to input error that may, in turn, affect the accuracy of measures of disposal length and returns to the criminal justice system during and after disposal completion.

Records involving a curfew requirement with EM and no additional requirements were initially identified as having a zero-day disposal length because probation staff input the same date for disposal start and termination, necessitating the augmentation of probation recorded disposal start and end dates with EM service provider installation and removal date records. However, it remains possible that additional data errors remain that have not been corrected.

The potential impact of any remaining errors on the conclusions cannot be known *a priori*, and as a result the current findings must be interpreted with a degree of caution. The results are therefore only valid under the assumption that any remaining errors are uncorrelated with the treatment and included outcomes.

The data were also subject to missingness across included covariates. In particular, comparatively high levels of missingness were observed in the measurement of offender ethnicity with nine per cent of records not including a valid ethnic code. These observations were omitted from the analysis.

Results in Brunton-Smith (2025b) were re-estimated using models where the ethnicity variable was omitted, ensuring those records with missing values for ethnicity were retained in the analysis, with no appreciable impact on conclusions. It therefore does not appear that the omission of these records was biasing the results.

Finally, not all EM service data could be accurately mapped on to specific probation events, with approximately 25 per cent of EM records omitted from the analysis. If reoffending rates (and related outcomes) for these unlinked records were systematically different, it is possible that the estimates of the effectiveness of RF EM would be biased.

References

Athey, S., Tibshirani, J., and Wager, S. (2019) *Generalized Random Forests*. Annals of Statistics.

Brunton-Smith (2025a) *Assessing the effectiveness of Radio Frequency Electronic Monitoring for Community and Suspended Sentence Orders: PNC-based proven reoffending analysis.*

Brunton-Smith (2025b) *Assessing the effectiveness of Radio Frequency Electronic Monitoring for Community and Suspended Sentence Orders: Court reconvictions during and after a community sentence, breaches and warnings.*

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., and Newey, a. W. (2016) *Double Machine Learning for Treatment and Causal Parameters*. ArXiv e-prints.

Eaton, G., and Mews, A. (2019) *The impact of short custodial sentences, community orders and suspended sentence orders on reoffending.* Ministry of Justice Analytical Series.

Greifer, N. (2023) *Assessing Balance*. Available at: https://cran.r-project.org/web/packages/MatchIt/vignettes/assessing-balance.html

Hirschi, T; and Gottfredson, M. (1983) *Age and the Explanation of Crime.* The American Journal of Sociology, 89 (3): 552–584.

Hucklesby, A. and Holdsworth, E. (2016) *Electronic Monitoring in England and Wales.* Centre for Criminal Justice Studies, University of Leeds. http://www.antoniocasella.eu/nume/Hucklesby_Holdsworth_2016.pdf

Iacus, S. M., King, G., and Porro, G. (2012) *Causal inference without balance checking: Coarsened Exact Matching.* Political Analysis, 20 (1): 1–24.

Lymperopoulou, K. (2022) *Ethnic inequalities in sentencing in the crown court – evidence from the MoJ Data First Criminal Justice datasets.* ADR UK Data Insights.

Ministry of Justice (2021) *Splink: MoJ's Open Source Library for Probabilistic Record Linkage at Scale*. Available at: https://www.gov.uk/government/publications/joined-up-data-in-government-the-future-of-data-linking-methods/splink-mojs-open-source-library-for-probabilistic-record-linkage-at-scale

Nagin, D., and Land, K. (1993) *Age, Criminal Careers, and Population Heterogeneity: Specification and Estimation of a Nonparametric, Mixed Poisson Model.* Criminology, 31(3): 327–362.

Rubin, D. (1974) *Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies.* Journal of Educational Psychology, 66(5): 688–701. doi:10.1037/h0037350.

Rubin, D. B. (2001) *Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation.* Health Services and Outcomes Research Methodology 2(3–4): 169–88.

Sturges, H. A. (1926) *The choice of class interval.* Journal of the American Statistical Association, 21(153): 65–66.

Wager, S., and Athey. S. (2018) *Estimation and inference of heterogeneous treatment effects using random forests. Journal of the American Statistical Association*, 113(523): 1228–1242.

Appendix A
Tables

Table A.1. Descriptive statistics for community orders (with valid OASys record only), 2014–18[29]

	non-RF EM (N=108,846)	RF EM (N=21,411)	Total (N=130,257)
Age (at offence)	31.60 (9.80) 30: 18.00–84.00	31.80 (9.85) 30: 18.00–81.00	31.63 (9.81) 30: 18.00–84.00

Gender

	non-RF EM (N=108,846)	RF EM (N=21,411)	Total (N=130,257)
Female (reference)	15,276 (14.0%)	2,984 (13.9%)	18,260 (14.0%)
Male	93,570 (86.0%)	18,427 (86.1%)	111,997 (86.0%)

Ethnicity

	non-RF EM (N=108,846)	RF EM (N=21,411)	Total (N=130,257)
Asian Indian	1,115 (1.0%)	152 (0.7%)	1,267 (1.0%)
Asian Pakistani	1,734 (1.6%)	397 (1.9%)	2,131 (1.6%)
Asian Bangladeshi	704 (0.6%)	122 (0.6%)	826 (0.6%)
Asian Chinese	41 (0.0%)	1 (0.0%)	42 (0.0%)
Asian Other	707 (0.6%)	108 (0.5%)	815 (0.6%)
Black Caribbean	3,244 (3.0%)	491 (2.3%)	3,735 (2.9%)
Black African	2,246 (2.1%)	343 (1.6%)	2,589 (2.0%)
Black Other	925 (0.8%)	130 (0.6%)	1,055 (0.8%)
White and Black Caribbean	2,503 (2.3%)	423 (2.0%)	2,926 (2.2%)
White and Black African	517 (0.5%)	95 (0.4%)	612 (0.5%)
White and Asian	436 (0.4%)	104 (0.5%)	540 (0.4%)
Mixed Other	660 (0.6%)	99 (0.5%)	759 (0.6%)
Arab	178 (0.2%)	22 (0.1%)	200 (0.2%)

[29] Numeric variables report: mean, (sd), median, min-max. Qualitative variables report N (%)

	non-RF EM	RF EM	Total
	(N=108,846)	(N=21,411)	(N=130,257)
Other ethnicity	735 (0.7%)	95 (0.4%)	830 (0.6%)
White British	87,761 (80.6%)	18,145 (84.7%)	105,906 (81.3%)
White Irish	887 (0.8%)	175 (0.8%)	1,062 (0.8%)
White Roma	741 (0.7%)	139 (0.6%)	880 (0.7%)
White Other	3,712 (3.4%)	370 (1.7%)	4,082 (3.1%)

Index offence

	non-RF EM	RF EM	Total
	(N=108,846)	(N=21,411)	(N=130,257)
Criminal damage	731 (0.7%)	129 (0.6%)	860 (0.7%)
Drugs	5,195 (4.8%)	1,154 (5.4%)	6,349 (4.9%)
Fraud	1,866 (1.7%)	361 (1.7%)	2,227 (1.7%)
Miscellaneous	3,841 (3.5%)	861 (4.0%)	4,702 (3.6%)
Weapons	181 (0.2%)	39 (0.2%)	220 (0.2%)
Public order	3,988 (3.7%)	755 (3.5%)	4,743 (3.6%)
Robbery	77 (0.1%)	13 (0.1%)	90 (0.1%)
Sex offences	840 (0.8%)	91 (0.4%)	931 (0.7%)
Summary	18,950 (17.4%)	3,610 (16.9%)	22,560 (17.3%)
Summary (motoring)	8,002 (7.4%)	1,754 (8.2%)	9,756 (7.5%)
Theft	27,519 (25.3%)	6,547 (30.6%)	34,066 (26.2%)
Violence	37,382 (34.3%)	6,001 (28.0%)	43,383 (33.3%)
Other (Breach)	180 (0.2%)	79 (0.4%)	259 (0.2%)
Other (Child offence)	94 (0.1%)	17 (0.1%)	111 (0.1%)

	non-RF EM	RF EM	Total
	(N=108,846)	(N=21,411)	(N=130,257)
History of drug offences (N)	0.16 (0.66) 0: 0.00–21.00	0.21 (0.77) 0: 0.00–18.00	0.17 (0.68) 0: 0.00–21.00
History of weapons offences (N)	0.01 (0.15) 0: 0.00–12.00	0.01 (0.14) 0: 0.00–5.00	0.01 (0.15) 0: 0.00–12.00
History of public order offences (N)	0.02 (0.26) 0: 0.00–19.00	0.02 (0.28) 0: 0.00–16.00	0.02 (0.26) 0: 0.00–19.00

	non-RF EM	RF EM	Total
	(N=108,846)	(N=21,411)	(N=130,257)
History of robbery (N)	0.01 (0.14) 0: 0.00–9.00	0.01 (0.15) 0: 0.00–6.00	0.01 (0.15) 0: 0.00–9.00
History of theft (N)	1.42 (3.28) 0: 0.00–75.00	2.02 (4.12) 0: 0.00–74.00	1.52 (3.44) 0: 0.00–75.00
History of Violence (N)	0.07 (0.37) 0: 0.00–18.00	0.09 (0.45) 0: 0.00–12.00	0.07 (0.38) 0: 0.00–18.00
History of summary offences (N)	1.61 (3.06) 1: 0.00–147.00	1.91 (3.41) 1: 0.00–76.00	1.66 (3.13) 1: 0.00–147.00
Other history (N)	0.16 (0.57) 0: 0.00–13.00	0.21 (0.65) 0: 0.00–13.00	0.17 (0.58) 0: 0.00–13.00
Prior prison sentences (N)	0.41 (1.59) 0: 0.00–39.00	0.58 (2.00) 0: 0.00–43.00	0.44 (1.67) 0: 0.00–43.00

	non-RF EM	RF EM	Total
	(N=108,846)	(N=21,411)	(N=130,257)
Number of offences in probation disposal	0.63 (1.12) 0: 0.00–19.00	0.59 (1.09) 0: 0.00–18.00	0.63 (1.12) 0: 0.00–19.00
Number of requirements	1.86 (0.94) 2: 1.00–13.00	1.82 (1.05) 1: 1.00–11.00	1.85 (0.95) 2: 1.00–13.00
Disposal length	311.97 (207.47) 364: 0.00–2,836.00	184.17 (179.18) 92: 0.00–2,280.00	290.96 (208.54) 323: 0.00–2,836.00
Electronic monitoring period		102.72 (97.08) 81: 0.00–2,280.00	

Year (requirement started)

	non-RF EM	RF EM	Total
	(N=108,846)	(N=21,411)	(N=130,257)
2014	25,795 (23.7%)	3,491 (16.3%)	29,286 (22.5%)
2015	24,024 (22.1%)	3,633 (17.0%)	27,657 (21.2%)
2016	19,910 (18.3%)	4,997 (23.3%)	24,907 (19.1%)
2017	19,207 (17.6%)	4,583 (21.4%)	23,790 (18.3%)
2018	19,910 (18.3%)	4,707 (22.0%)	24,617 (18.9%)

OASys flags

	non-RF EM (N=108,846)	RF EM (N=21,411)	Total (N=130,257)
Accommodation	0.45 (0.50): 0.00–1.00	0.43 (0.49): 0.00–1.00	0.45 (0.50): 0.00–1.00
Employment	0.54 (0.50): 0.00–1.00	0.60 (0.49): 0.00–1.00	0.55 (0.50): 0.00–1.00
Relationships	0.71 (0.45): 0.00–1.00	0.70 (0.46): 0.00–1.00	0.71 (0.45): 0.00–1.00
Lifestyle & Associates	0.68 (0.47): 0.00–1.00	0.75 (0.43): 0.00–1.00	0.69 (0.46): 0.00–1.00
Drug Misuse	0.48 (0.50): 0.00–1.00	0.55 (0.50): 0.00–1.00	0.49 (0.50): 0.00–1.00
Alcohol Misuse	0.35 (0.48): 0.00–1.00	0.36 (0.48): 0.00–1.00	0.35 (0.48): 0.00–1.00
Thinking & Behaviour	0.66 (0.47): 0.00–1.00	0.68 (0.47): 0.00–1.00	0.66 (0.47): 0.00–1.00
Attitudes	0.68 (0.47): 0.00–1.00	0.73 (0.44): 0.00–1.00	0.69 (0.46): 0.00–1.00

Table A.2. Descriptive statistics for community orders (PNC cohort), April 2016 – March 2017[30]

	non-RF EM	RF EM	Total
	(N=67,680)	(N=12,137)	(N=79,817)
Age (at offence)	32.99 (10.67) 31: 18.00–82.00	33.18 (10.81) 31 18.00–85.00	33.02 (10.69) 31: 18.00–85.00

Gender

	non-RF EM	RF EM	Total
	(N=67,680)	(N=12,137)	(N=79,817)
Female (reference)	11,046 (16.3%)	2,116 (17.4%)	13,162 (16.5%)
Male	56,634 (83.7%)	10,021 (82.6%)	66,655 (83.5%)

Ethnicity

	non-RF EM	RF EM	Total
	(N=67,680)	(N=12,137)	(N=79,817)
Asian Indian	979 (1.4%)	119 (1.0%)	1,098 (1.4%)
Asian Pakistani	1,396 (2.1%)	244 (2.0%)	1,640 (2.1%)
Asian Bangladeshi	523 (0.8%)	74 (0.6%)	597 (0.7%)
Asian Chinese	54 (0.1%)	3 (0.0%)	57 (0.1%)
Asian Other	675 (1.0%)	71 (0.6%)	746 (0.9%)
Black Caribbean	1,971 (2.9%)	370 (3.0%)	2,341 (2.9%)
Black African	1,757 (2.6%)	257 (2.1%)	2,014 (2.5%)
Black Other	574 (0.8%)	88 (0.7%)	662 (0.8%)
White and Black Caribbean	1,256 (1.9%)	256 (2.1%)	1,512 (1.9%)
White and Black African	289 (0.4%)	54 (0.4%)	343 (0.4%)
White and Asian	258 (0.4%)	45 (0.4%)	303 (0.4%)
Mixed Other	421 (0.6%)	69 (0.6%)	490 (0.6%)
Arab	152 (0.2%)	19 (0.2%)	171 (0.2%)
Other ethnicity	739 (1.1%)	89 (0.7%)	828 (1.0%)
White British	51,518 (76.1%)	9,754 (80.4%)	61,272 (76.8%)
White Irish	528 (0.8%)	125 (1.0%)	653 (0.8%)
White Roma	378 (0.6%)	67 (0.6%)	445 (0.6%)
White Other	4,212 (6.2%)	433 (3.6%)	4,645 (5.8%)

30 Numeric variables report mean (sd) median: min-max. Qualitative variables report N (%)

Index offence

	non-RF EM (N=67,680)	RF EM (N=12,137)	Total (N=79,817)
Criminal damage	311 (0.5%)	58 (0.5%)	369 (0.5%)
Drugs	3,229 (4.8%)	705 (5.8%)	3,934 (4.9%)
Fraud	2,237 (3.3%)	367 (3.0%)	2,604 (3.3%)
Miscellaneous	2,802 (4.1%)	456 (3.8%)	3,258 (4.1%)
Weapons	110 (0.2%)	21 (0.2%)	131 (0.2%)
Public order	1,997 (3.0%)	391 (3.2%)	2,388 (3.0%)
Robbery	24 (0.0%)	5 (0.0%)	29 (0.0%)
Sex offences	619 (0.9%)	72 (0.6%)	691 (0.9%)
Summary	11,006 (16.3%)	2,147 (17.7%)	13,153 (16.5%)
Summary (motoring)	8,903 (13.2%)	1,384 (11.4%)	10,287 (12.9%)
Theft	13,326 (19.7%)	2,970 (24.5%)	16,296 (20.4%)
Violence	22,940 (33.9%)	3,504 (28.9%)	26,444 (33.1%)
Other (Breach)	111 (0.2%)	48 (0.4%)	159 (0.2%)
Other (Child offence)	65 (0.1%)	9 (0.1%)	74 (0.1%)

	non-RF EM (N=67,680)	RF EM (N=12,137)	Total (N=79,817)
History of drug offences (N)	0.13 (0.61) 0: 0.00–21.00	0.16 (0.70) 0: 0.00–18.00	0.13 (0.63) 0: 0.00–21.00
History of weapons offences (N)	0.01 (0.10) 0: 0.00–4.00	0.01 (0.10) 0: 0.00–4.00	0.01 (0.10) 0: 0.00–4.00
History of public order offences (N)	0.01 (0.18) 0: 0.00–12.00	0.02 (0.24) 0: 0.00–10.00	0.01 (0.19) 0: 0.00–12.00
History of robbery (N)	0.01 (0.12) 0: 0.00–6.00	0.01 (0.14) 0: 0.00–6.00	0.01 (0.13) 0: 0.00–6.00
History of theft (N)	1.04 (3.02) 0: 0.00–66.00	1.36 (3.36) 0: 0.00–53.00	1.09 (3.08) 0: 0.00–66.00
History of Violence (N)	0.04 (0.29) 0: 0.00–15.00	0.05 (0.32) 0: 0.00–8.00	0.04 (0.29) 0: 0.00–15.00
History of summary offences (N)	1.30 (2.54) 0: 0.00–129.00	1.55 (2.83) 1: 0.00–61.00	1.34 (2.59) 0: 0.00–129.00
Other history (N)	0.11 (0.47) 0: 0.00–12.00	0.14 (0.53) 0: 0.00–10.00	0.12 (0.48) 0: 0.00–12.00

	non-RF EM	RF EM	Total
	(N=67,680)	(N=12,137)	(N=79,817)
Prior prison sentences (N)	0.27 (1.38) 0: 0.00–36.00	0.38 (1.69) 0: 0.00–43.00	0.29 (1.43) 0: 0.00–43.00

	non-RF EM	RF EM	Total
	(N=67,680)	(N=12,137)	(N=79,817)
Number of offences in probation disposal	0.56 (1.03) 0: 0.00–18.00	0.50 (0.97) 0: 0.00–10.00	0.55 (1.02) 0: 0.00–18.00
Number of requirements	1.60 (0.77) 1: 1.00–8.00	1.64 (0.85) 1: 1.00–9.00	1.60 (0.78) 1: 1.00–9.00
Disposal length	339.81 (219.49) 364: 0.00–2,619.00	180.27 (177.39) 91: 0.00–1,793.00	315.55 (221.17) 364: 0.00–2,619.00
Electronic monitoring period		95.26 (81.65) 81: 0.00–1,456.00	

Financial quarter

	non-RF EM	RF EM	Total
	(N=67,680)	(N=12,137)	(N=79,817)
April 16	17,833 (26.3%)	3,137 (25.8%)	20,970 (26.3%)
July 16	16,732 (24.7%)	2,952 (24.3%)	19,684 (24.7%)
October 16	15,496 (22.9%)	2,901 (23.9%)	18,397 (23.0%)
January 17	17,619 (26.0%)	3,147 (25.9%)	20,766 (26.0%)

Table A.3. Balance tables for community orders (with valid OASys record only), 2014–18

CEM

	Control (mean)	Control (SD)	RF EM (mean)	RF EM (sd)	Difference	Ratio	Overlap
Age (at offence)	28.631	8.074	28.638	8.060	0.001	0.997	0.011
Gender: Male	0.923	0.266	0.923	0.266	0.000	*NA*	0.000
Ethnicity: Asian Indian	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Ethnicity: Asian Pakistani	0.001	0.034	0.001	0.034	0.000	*NA*	0.000
Ethnicity: Asian Bangladeshi	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Ethnicity: Asian Chinese	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Ethnicity: Asian Other	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Ethnicity: Black Caribbean	0.001	0.034	0.001	0.034	0.000	*NA*	0.000
Ethnicity: Black African	0.001	0.034	0.001	0.034	0.000	*NA*	0.000
Ethnicity: Black Other	0.000	0.020	0.000	0.020	0.000	*NA*	0.000
Ethnicity: White and Black Caribbean	0.002	0.044	0.002	0.044	0.000	*NA*	0.000
Ethnicity: White and Black African	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Ethnicity: White and Asian	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Ethnicity: Mixed Other	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Ethnicity: Arab	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Ethnicity: Other ethnicity	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Ethnicity: White British	0.992	0.090	0.992	0.090	0.000	*NA*	0.000
Ethnicity: White Irish	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Ethnicity: White Roma	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Ethnicity: White Other	0.002	0.048	0.002	0.048	0.000	*NA*	0.000

	Control (mean)	Control (SD)	RF EM (mean)	RF EM (sd)	Difference	Ratio	Overlap
Index offence: Criminal damage	0.000	0.020	0.000	0.020	0.000	*NA*	0.000
Index offence: Drugs	0.013	0.112	0.013	0.112	0.000	*NA*	0.000
Index offence: Fraud	0.002	0.044	0.002	0.044	0.000	*NA*	0.000
Index offence: Miscellaneous	0.008	0.092	0.008	0.092	0.000	*NA*	0.000
Index offence: Weapons	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Index offence: Public order	0.012	0.110	0.012	0.110	0.000	*NA*	0.000
Index offence: Robbery	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Index offence: Sex offences	0.000	0.020	0.000	0.020	0.000	*NA*	0.000
Index offence: Summary	0.199	0.399	0.199	0.399	0.000	*NA*	0.000
Index offence: Summary (motoring)	0.044	0.205	0.044	0.205	0.000	*NA*	0.000
Index offence: Theft	0.348	0.476	0.348	0.476	0.000	*NA*	0.000
Index offence: Violence	0.373	0.484	0.373	0.484	0.000	*NA*	0.000
Index offence: Other (Breach)	0.000	0.020	0.000	0.020	0.000	*NA*	0.000
Index offence: Other (Child offence)	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
History of drug offences (N)	0.063	0.247	0.053	0.228	-0.013	0.851	0.015
History of weapons offences (N)	0.001	0.028	0.001	0.028	0.000	1.000	0.092
History of public order offences (N)	0.005	0.068	0.006	0.078	0.006	1.337	0.023
History of robbery (N)	0.001	0.039	0.001	0.039	0.000	1.000	0.047
History of theft (N)	1.181	2.047	1.140	2.072	-0.010	1.024	0.033
History of Violence (N)	0.005	0.071	0.005	0.071	0.000	1.000	0.021
History of summary offences (N)	1.278	1.761	1.330	1.704	0.015	0.936	0.057
Other history (N)	0.025	0.187	0.025	0.187	0.000	1.000	0.001

	Control (mean)	Control (SD)	RF EM (mean)	RF EM (sd)	Difference	Ratio	Overlap
Prior prison sentences (N)	0.190	0.653	0.160	0.622	-0.015	0.907	0.015
Number of offences in probation disposal	0.293	0.519	0.261	0.505	-0.030	0.946	0.035
Number of requirements	1.577	0.737	1.577	0.737	0.000	1.000	0.003
Disposal length	174.622	148.283	166.161	148.709	-0.047	1.006	0.126
OASys flags: Accommodation	0.534	0.499	0.534	0.499	0.000	*NA*	0.000
OASys flags: Employment	0.690	0.463	0.690	0.463	0.000	*NA*	0.000
OASys flags: Relationships	0.814	0.389	0.814	0.389	0.000	*NA*	0.000
OASys flags: Lifestyle & Associates	0.789	0.408	0.789	0.408	0.000	*NA*	0.000
OASys flags: Drug Misuse	0.606	0.489	0.606	0.489	0.000	*NA*	0.000
OASys flags: Alcohol Misuse	0.358	0.480	0.358	0.480	0.000	*NA*	0.000
OASys flags: Thinking & Behaviour	0.782	0.413	0.782	0.413	0.000	*NA*	0.000
OASys flags: Attitudes	0.787	0.409	0.787	0.409	0.000	*NA*	0.000
Year (requirement started): 2014	0.240	0.427	0.240	0.427	0.000	*NA*	0.000
Year (requirement started): 2015	0.231	0.422	0.231	0.422	0.000	*NA*	0.000
Year (requirement started): 2016	0.235	0.424	0.235	0.424	0.000	*NA*	0.000
Year (requirement started): 2017	0.154	0.361	0.154	0.361	0.000	*NA*	0.000
Year (requirement started): 2018	0.139	0.346	0.139	0.346	0.000	*NA*	0.000
PSM (distance)	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
Average					0.002		

	Control (mean)	Control (SD)
All (Effective Sample Size)	108,743	21,410
All (Unweighted)	108,743	21,410
Matched (Effective Sample Size)	2,393.11	2,598
Matched (Unweighted)	3,706	2,598
Unmatched	105,037	18,812
Match rate		12%

PSM

	Control (mean)	Control (SD)	RF EM (mean)	RF EM (sd)	Difference	Ratio	Overlap
Age (at offence)	31.668	9.746	31.805	9.850	0.014	1.021	0.016
Gender: Male	0.861	0.346	0.861	0.346	0.000	*NA*	0.000
Ethnicity: Asian Indian	0.006	0.078	0.007	0.084	0.012	*NA*	0.001
Ethnicity: Asian Pakistani	0.018	0.132	0.019	0.135	0.006	*NA*	0.001
Ethnicity: Asian Bangladeshi	0.006	0.077	0.006	0.075	-0.004	*NA*	0.000
Ethnicity: Asian Chinese	0.000	0.007	0.000	0.007	0.000	*NA*	0.000
Ethnicity: Asian Other	0.004	0.065	0.005	0.071	0.011	*NA*	0.001
Ethnicity: Black Caribbean	0.021	0.145	0.023	0.150	0.010	*NA*	0.001
Ethnicity: Black African	0.015	0.122	0.016	0.126	0.008	*NA*	0.001
Ethnicity: Black Other	0.006	0.077	0.006	0.078	0.001	*NA*	0.000
Ethnicity: White and Black Caribbean	0.017	0.131	0.020	0.139	0.016	*NA*	0.002
Ethnicity: White and Black African	0.004	0.067	0.004	0.066	-0.001	*NA*	0.000
Ethnicity: White and Asian	0.005	0.070	0.005	0.070	-0.001	*NA*	0.000
Ethnicity: Mixed Other	0.005	0.069	0.005	0.068	-0.003	*NA*	0.000
Ethnicity: Arab	0.001	0.033	0.001	0.032	-0.001	*NA*	0.000
Ethnicity: Other ethnicity	0.004	0.064	0.004	0.066	0.006	*NA*	0.000
Ethnicity: White British	0.855	0.352	0.848	0.359	-0.021	*NA*	0.008
Ethnicity: White Irish	0.008	0.091	0.008	0.090	-0.001	*NA*	0.000
Ethnicity: White Roma	0.006	0.077	0.006	0.080	0.006	*NA*	0.001
Ethnicity: White Other	0.017	0.129	0.017	0.130	0.002	*NA*	0.000
Index offence: Criminal damage	0.006	0.080	0.006	0.077	-0.006	*NA*	0.000

	Control (mean)	Control (SD)	RF EM (mean)	RF EM (sd)	Difference	Ratio	Overlap
Index offence: Drugs	0.052	0.221	0.054	0.226	0.010	*NA*	0.002
Index offence: Fraud	0.017	0.130	0.017	0.129	-0.002	*NA*	0.000
Index offence: Miscellaneous	0.040	0.197	0.040	0.196	0.000	*NA*	0.000
Index offence: Weapons	0.002	0.045	0.002	0.043	-0.004	*NA*	0.000
Index offence: Public order	0.034	0.182	0.035	0.184	0.005	*NA*	0.001
Index offence: Robbery	0.001	0.026	0.001	0.025	-0.004	*NA*	0.000
Index offence: Sex offences	0.005	0.071	0.004	0.065	-0.014	*NA*	0.001
Index offence: Summary	0.161	0.368	0.169	0.374	0.020	*NA*	0.008
Index offence: Summary (motoring)	0.078	0.268	0.082	0.274	0.014	*NA*	0.004
Index offence: Theft	0.315	0.464	0.306	0.461	-0.019	*NA*	0.009
Index offence: Violence	0.285	0.451	0.280	0.449	-0.010	*NA*	0.004
Index offence: Other (Breach)	0.003	0.054	0.004	0.061	0.013	*NA*	0.001
Index offence: Other (Child offence)	0.001	0.031	0.001	0.028	-0.007	*NA*	0.000
History of drug offences (N)	0.196	0.741	0.208	0.769	0.015	1.077	0.010
History of weapons offences (N)	0.013	0.174	0.013	0.138	0.000	0.629	0.093
History of public order offences (N)	0.024	0.332	0.023	0.276	-0.003	0.692	0.023
History of robbery (N)	0.016	0.181	0.014	0.153	-0.014	0.717	0.048
History of theft (N)	2.052	3.943	2.021	4.117	-0.008	1.090	0.028
History of Violence (N)	0.088	0.456	0.087	0.447	-0.002	0.961	0.022
History of summary offences (N)	1.874	3.580	1.910	3.409	0.011	0.907	0.046
Other history (N)	0.208	0.669	0.206	0.648	-0.004	0.937	0.005
Prior prison sentences (N)	0.601	1.848	0.582	1.999	-0.010	1.171	0.017

	Control (mean)	Control (SD)	RF EM (mean)	RF EM (sd)	Difference	Ratio	Overlap
Number of offences in probation disposal	0.614	1.108	0.587	1.091	-0.025	0.969	0.023
Number of requirements	1.897	0.997	1.820	1.046	-0.073	1.101	0.120
Disposal length	187.224	157.784	184.182	179.185	-0.017	1.290	0.174
OASys flags: Accommodation	0.438	0.496	0.426	0.495	-0.023	*NA*	0.011
OASys flags: Employment	0.609	0.488	0.605	0.489	-0.010	*NA*	0.005
OASys flags: Relationships	0.706	0.456	0.701	0.458	-0.011	*NA*	0.005
OASys flags: Lifestyle & Associates	0.760	0.427	0.755	0.430	-0.013	*NA*	0.005
OASys flags: Drug Misuse	0.550	0.497	0.545	0.498	-0.010	*NA*	0.005
OASys flags: Alcohol Misuse	0.355	0.479	0.356	0.479	0.001	*NA*	0.001
OASys flags: Thinking & Behaviour	0.686	0.464	0.680	0.467	-0.012	*NA*	0.006
OASys flags: Attitudes	0.739	0.439	0.732	0.443	-0.017	*NA*	0.008
Year (requirement started): 2014	0.164	0.370	0.163	0.369	-0.003	*NA*	0.001
Year (requirement started): 2015	0.174	0.379	0.170	0.375	-0.011	*NA*	0.004
Year (requirement started): 2016	0.228	0.420	0.233	0.423	0.013	*NA*	0.005
Year (requirement started): 2017	0.214	0.410	0.214	0.410	-0.001	*NA*	0.000
Year (requirement started): 2018	0.220	0.414	0.220	0.414	0.000	*NA*	0.000
PSM (distance)	0.249	0.124	0.249	0.124	0.000	1.001	0.000
Average					0.009		

	Control (mean)	Control (SD)
All (Effective Sample Size)	108,743	21,410
All (Unweighted)	108,743	21,410
Matched (Effective Sample Size)	21,408	21,408
Matched (Unweighted)	21,408	21,408
Unmatched	87,335	2
Match rate		100%

Table A.4. Descriptive statistics for suspended sentence orders (with valid OASys record only), 2014–18[31]

	non-RF EM (N=63,562)	RF EM (N=10,041)	Total (N=73,603)
Age (at offence)	31.36 (9.74) 30: 18.00–84.00	30.34 (9.91) 28: 18.00–80.00	31.22 (9.77) 29: 18.00–84.00

Gender

	non-RF EM (N=63,562)	RF EM (N=10,041)	Total (N=73,603)
Female (reference)	7,225 (11.4%)	1,036 (10.3%)	8,261 (11.2%)
Male	56,337 (88.6%)	9,005 (89.7%)	65,342 (88.8%)

Ethnicity

	non-RF EM (N=63,562)	RF EM (N=10,041)	Total (N=73,603)
Asian Indian	610 (1.0%)	103 (1.0%)	713 (1.0%)
Asian Pakistani	1,232 (1.9%)	227 (2.3%)	1,459 (2.0%)
Asian Bangladeshi	519 (0.8%)	66 (0.7%)	585 (0.8%)
Asian Chinese	30 (0.0%)	2 (0.0%)	32 (0.0%)
Asian Other	437 (0.7%)	56 (0.6%)	493 (0.7%)
Black Caribbean	2,142 (3.4%)	324 (3.2%)	2,466 (3.4%)
Black African	1,441 (2.3%)	223 (2.2%)	1,664 (2.3%)
Black Other	528 (0.8%)	75 (0.7%)	603 (0.8%)
White and Black Caribbean	1,591 (2.5%)	243 (2.4%)	1,834 (2.5%)
White and Black African	295 (0.5%)	53 (0.5%)	348 (0.5%)
White and Asian	290 (0.5%)	39 (0.4%)	329 (0.4%)
Mixed Other	450 (0.7%)	60 (0.6%)	510 (0.7%)
Arab	112 (0.2%)	21 (0.2%)	133 (0.2%)
Other ethnicity	471 (0.7%)	50 (0.5%)	521 (0.7%)
White British	50,572 (79.6%)	8,109 (80.8%)	58,681 (79.7%)
White Irish	559 (0.9%)	91 (0.9%)	650 (0.9%)
White Roma	427 (0.7%)	83 (0.8%)	510 (0.7%)
White Other	1,856 (2.9%)	216 (2.2%)	2,072 (2.8%)

31 Numeric variables report: mean, (sd), median, min-max. Qualitative variables report N (%)

Index offence

	non-RF EM	RF EM	Total
	(N=63,562)	(N=10,041)	(N=73,603)
Criminal damage	401 (0.6%)	55 (0.5%)	456 (0.6%)
Drugs	4,399 (6.9%)	850 (8.5%)	5,249 (7.1%)
Fraud	1,247 (2.0%)	191 (1.9%)	1,438 (2.0%)
Miscellaneous	3,294 (5.2%)	418 (4.2%)	3,712 (5.0%)
Weapons	184 (0.3%)	25 (0.2%)	209 (0.3%)
Public order	3,187 (5.0%)	540 (5.4%)	3,727 (5.1%)
Robbery	463 (0.7%)	108 (1.1%)	571 (0.8%)
Sex offences	744 (1.2%)	89 (0.9%)	833 (1.1%)
Summary	6,273 (9.9%)	905 (9.0%)	7,178 (9.8%)
Summary (motoring)	5412 (8.5%)	1,074 (10.7%)	6,486 (8.8%)
Theft	14,961 (23.5%)	2,465 (24.5%)	17,426 (23.7%)
Violence	22,749 (35.8%)	3,281 (32.7%)	26030 (35.4%)
Other (Breach)	115 (0.2%)	23 (0.2%)	138 (0.2%)
Other (Child offence)	133 (0.2%)	17 (0.2%)	150 (0.2%)

	non-RF EM	RF EM	Total
	(N=63,562)	(N=10,041)	(N=73,603)
History of drug offences (N)	0.17 (0.67) 0: 0.00–30.00	0.21 (0.75) 0: 0.00–11.00	0.18 (0.68) 0: 0.00–30.00
History of weapons offences (N)	0.02 (0.17) 0: 0.00–12.00	0.02 (0.17) 0: 0.00–6.00	0.02 (0.17) 0: 0.00–12.00
History of public order offences (N)	0.02 (0.28) 0: 0.00–35.00	0.02 (0.30) 0: 0.00–22.00	0.02 (0.28) 0: 0.00–35.00
History of robbery (N)	0.02 (0.17) 0: 0.00–7.00	0.02 (0.22) 0: 0.00–9.00	0.02 (0.17) 0: 0.00–9.00
History of theft (N)	1.67 (3.79) 0: 0.00–72.00	1.65 (3.74) 0: 0.00–42.00	1.67 (3.79) 0: 0.00–72.00
History of Violence (N)	0.10 (0.48) 0: 0.00–21.00	0.11 (0.50) 0: 0.00–15.00	0.10 (0.48) 0: 0.00–21.00
History of summary offences (N)	1.61 (3.09) 1: 0.00–163.00	1.73 (3.21) 1: 0.00–79.00	1.63 (3.11) 1: 0.00–163.00
Other history (N)	0.19 (0.62) 0: 0.00–14.00	0.22 (0.70) 0: 0.00–13.00	0.20 (0.64) 0: 0.00–14.00

	non-RF EM	RF EM	Total
	(N=63,562)	(N=10,041)	(N=73,603)
Prior prison sentences (N)	0.59 (1.97) 0: 0.00–41.00	0.62 (2.12) 0: 0.00–42.00	0.59 (1.99) 0: 0.00–42.00

	non-RF EM	RF EM	Total
	(N=63,562)	(N=10,041)	(N=73,603)
Number of offences in probation disposal	0.90 (1.46) 0: 0.00–26.00	0.87 (1.41) 0: 0.00–19.00	0.90 (1.46) 0: 0.00–26.00
Number of requirements	1.97 (0.94) 2: 1.00–14.00	2.34 (1.17) 2: 1.00–11.00	2.02 (0.98) 2: 1.00–14.00
Disposal length	353.44 (229.46) 364: 0.00–2,720.00	326.84 (251.54) 324: 0.00–1,666.00	349.81 (232.77) 364: 0.00–2,720.00
Electronic monitoring period		173.85 (152.47) 122: 0.00–1,830.00	

Year (requirement started)

	non-RF EM	RF EM	Total
	(N=63,562)	(N=10,041)	(N=73,603)
2014	14,353 (22.6%)	1,562 (15.6%)	15,915 (21.6%)
2015	14,350 (22.6%)	1,815 (18.1%)	16,165 (22.0%)
2016	12,669 (19.9%)	2,569 (25.6%)	15,238 (20.7%)
2017	12,447 (19.6%)	2,341 (23.3%)	14,788 (20.1%)
2018	9,743 (15.3%)	1,754 (17.5%)	11,497 (15.6%)

OASys flags

	non-RF EM	RF EM	Total
	(N=63,562)	(N=10,041)	(N=73,603)
Accommodation	0.44 (0.50): 0.00–1.00	0.39 (0.49): 0.00–1.00	0.44 (0.50): 0.00–1.00
Employment	0.55 (0.50): 0.00–1.00	0.57 (0.49): 0.00–1.00	0.55 (0.50): 0.00–1.00
Relationships	0.70 (0.46): 0.00–1.00	0.66 (0.47): 0.00–1.00	0.70 (0.46): 0.00–1.00
Lifestyle & Associates	0.73 (0.45): 0.00–1.00	0.76 (0.43): 0.00–1.00	0.73 (0.44): 0.00–1.00
Drug Misuse	0.49 (0.50): 0.00–1.00	0.52 (0.50): 0.00–1.00	0.49 (0.50): 0.00–1.00

	non-RF EM	RF EM	Total
	(N=63,562)	**(N=10,041)**	**(N=73,603)**
Alcohol Misuse	0.33 (0.47): 0.00–1.00	0.32 (0.47): 0.00–1.00	0.33 (0.47): 0.00–1.00
Thinking & Behaviour	0.68 (0.47): 0.00–1.00	0.67 (0.47): 0.00–1.00	0.68 (0.47): 0.00–1.00
Attitudes	0.70 (0.46): 0.00–1.00	0.72 (0.45): 0.00–1.00	0.71 (0.46): 0.00–1.00

Table A.5. Descriptive statistics for suspended sentence orders (PNC cohort), April 2016 – March 2017[32]

	non-RF EM (N=34,707)	RF EM (N=6,598)	Total (N=41,305)
Age (at offence)	32.86 (10.79) 31: 18.00–88.00	31.55 (10.96) 29: 18.00–80.00	32.65 (10.83) 31: 18.00–88.00

Gender

	non-RF EM (N=34,707)	RF EM (N=6,598)	Total (N=41,305)
Female (reference)	4,765 (13.7%)	918 (13.9%)	5,683 (13.8%)
Male	29,942 (86.3%)	5,680 (86.1%)	35,622 (86.2%)

Ethnicity

	non-RF EM (N=34,707)	RF EM (N=6,598)	Total (N=41,305)
Asian Indian	503 (1.4%)	81 (1.2%)	584 (1.4%)
Asian Pakistani	915 (2.6%)	169 (2.6%)	1,084 (2.6%)
Asian Bangladeshi	331 (1.0%)	59 (0.9%)	390 (0.9%)
Asian Chinese	39 (0.1%)	6 (0.1%)	45 (0.1%)
Asian Other	388 (1.1%)	52 (0.8%)	440 (1.1%)
Black Caribbean	1,113 (3.2%)	227 (3.4%)	1,340 (3.2%)
Black African	1,032 (3.0%)	175 (2.7%)	1,207 (2.9%)
Black Other	316 (0.9%)	52 (0.8%)	368 (0.9%)
White and Black Caribbean	748 (2.2%)	161 (2.4%)	909 (2.2%)
White and Black African	154 (0.4%)	28 (0.4%)	182 (0.4%)
White and Asian	143 (0.4%)	28 (0.4%)	171 (0.4%)
Mixed Other	238 (0.7%)	44 (0.7%)	282 (0.7%)
Arab	68 (0.2%)	21 (0.3%)	89 (0.2%)
Other ethnicity	394 (1.1%)	56 (0.8%)	450 (1.1%)
White British	25,879 (74.6%)	5,118 (77.6%)	30,997 (75.0%)
White Irish	308 (0.9%)	73 (1.1%)	381 (0.9%)
White Roma	220 (0.6%)	40 (0.6%)	260 (0.6%)
White Other	1,918 (5.5%)	208 (3.2%)	2,126 (5.1%)

[32] Numeric variables report: mean, (sd), median, min-max. Qualitative variables report N (%)

Index offence

	non-RF EM	RF EM	Total
	(N=34,707)	(N=6,598)	(N=41,305)
Criminal damage	198 (0.6%)	44 (0.7%)	242 (0.6%)
Drugs	3,167 (9.1%)	756 (11.5%)	3,923 (9.5%)
Fraud	1,585 (4.6%)	314 (4.8%)	1,899 (4.6%)
Miscellaneous	2,688 (7.7%)	374 (5.7%)	3,062 (7.4%)
Weapons	151 (0.4%)	28 (0.4%)	179 (0.4%)
Public order	1,429 (4.1%)	323 (4.9%)	1,752 (4.2%)
Robbery	199 (0.6%)	60 (0.9%)	259 (0.6%)
Sex offences	618 (1.8%)	82 (1.2%)	700 (1.7%)
Summary	3,000 (8.6%)	527 (8.0%)	3,527 (8.5%)
Summary (motoring)	4,077 (11.7%)	815 (12.4%)	4,892 (11.8%)
Theft	5,902 (17.0%)	1,183 (17.9%)	7,085 (17.2%)
Violence	11,517 (33.2%)	2,056 (31.2%)	13,573 (32.9%)
Other (Breach)	62 (0.2%)	21 (0.3%)	83 (0.2%)
Other (Child offence)	114 (0.3%)	15 (0.2%)	129 (0.3%)

	non-RF EM	RF EM	Total
	(N=34,707)	(N=6,598)	(N=41,305)
History of drug offences (N)	0.14 (0.60) 0: 0.00–17.00	0.16 (0.65) 0: 0.00–9.00	0.14 (0.61) 0: 0.00–17.00
History of weapons offences (N)	0.01 (0.14) 0: 0.00–12.00	0.01 (0.11) 0: 0.00–2.00	0.01 (0.13) 0: 0.00–12.00
History of public order offences (N)	0.01 (0.14) 0: 0.00–7.00	0.01 (0.14) 0: 0.00–6.00	0.01 (0.14) 0: 0.00–7.00
History of robbery (N)	0.01 (0.14) 0: 0.00–5.00	0.01 (0.17) 0: 0.00–4.00	0.01 (0.14) 0: 0.00–5.00
History of theft (N)	1.02 (3.03) 0: 0.00–67.00	1.02 (3.02) 0: 0.00–40.00	1.02 (3.03) 0: 0.00–67.00
History of Violence (N)	0.06 (0.33) 0: 0.00–8.00	0.06 (0.32) 0: 0.00–5.00	0.06 (0.33) 0: 0.00–8.00
History of summary offences (N)	1.24 (2.54) 0: 0.00–129.00	1.29 (2.55) 0: 0.00–56.00	1.24 (2.54) 0: 0.00–129.00
Other history (N)	0.14 (0.54) 0: 0.00–14.00	0.14 (0.53) 0: 0.00–9.00	0.14 (0.54) 0: 0.00–14.00

	non-RF EM	RF EM	Total
	(N=34,707)	(N=6,598)	(N=41,305)
Prior prison sentences (N)	0.34 (1.53) 0: 0.00–31.00	0.37 (1.67) 0: 0.00–27.00	0.35 (1.55) 0: 0.00–31.00

	non-RF EM	RF EM	Total
	(N=34,707)	(N=6,598)	(N=41,305)
Number of offences in probation disposal	0.80 (1.35) 0: 0.00–30.00	0.76 (1.34) 0: 0.00–19.00	0.79 (1.35) 0: 0.00–30.00
Number of requirements	1.70 (0.81) 2: 1.00–8.00	2.17 (1.02) 2: 1.00–8.00	1.77 (0.86) 2: 1.00–8.00
Disposal length	425.15 (227.80) 364: 0.00–2,392.00	367.57 (262.83) 364: 0.00–1,301.00	415.95 (234.70) 364: 0.00–2,392.00
Electronic monitoring period		169.46 (142.05) 122: 0.00–1,460.00	

Financial quarter

	non-RF EM	RF EM	Total
	(N=34,707)	(N=6,598)	(N=41,305)
April 16	9,182 (26.5%)	1,629 (24.7%)	10,811 (26.2%)
July 16	8,622 (24.8%)	1,598 (24.2%)	10,220 (24.7%)
October 16	7,985 (23.0%)	1,620 (24.6%)	9,605 (23.3%)
January 17	8,918 (25.7%)	1,751 (26.5%)	10,669 (25.8%)

Table A.6. Balance tables for suspended sentence orders (with valid OASys record only), 2014–18

CEM

	Control (mean)	Control (SD)	RF EM (mean)	RF EM (sd)	Difference	Ratio	Overlap
Age (at offence)	27.156	7.227	27.138	7.387	-0.002	1.045	0.018
Gender: Male	0.962	0.191	0.962	0.191	0.000	*NA*	0.000
Ethnicity: Asian Indian	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Ethnicity: Asian Pakistani	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Ethnicity: Asian Bangladeshi	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Ethnicity: Asian Chinese	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Ethnicity: Asian Other	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Ethnicity: Black Caribbean	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Ethnicity: Black African	0.008	0.089	0.008	0.089	0.000	*NA*	0.000
Ethnicity: Black Other	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Ethnicity: White and Black Caribbean	0.002	0.040	0.002	0.040	0.000	*NA*	0.000
Ethnicity: White and Black African	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Ethnicity: White and Asian	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Ethnicity: Mixed Other	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Ethnicity: Arab	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Ethnicity: Other ethnicity	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Ethnicity: White British	0.990	0.097	0.990	0.097	0.000	*NA*	0.000
Ethnicity: White Irish	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Ethnicity: White Roma	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Ethnicity: White Other	0.000	0.000	0.000	0.000	0.000	*NA*	0.000

	Control (mean)	Control (SD)	RF EM (mean)	RF EM (sd)	Difference	Ratio	Overlap
Index offence: Criminal damage	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Index offence: Drugs	0.040	0.195	0.040	0.195	0.000	*NA*	0.000
Index offence: Fraud	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Index offence: Miscellaneous	0.016	0.125	0.016	0.125	0.000	*NA*	0.000
Index offence: Weapons	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Index offence: Public order	0.014	0.119	0.014	0.119	0.000	*NA*	0.000
Index offence: Robbery	0.002	0.040	0.002	0.040	0.000	*NA*	0.000
Index offence: Sex offences	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Index offence: Summary	0.071	0.258	0.071	0.258	0.000	*NA*	0.000
Index offence: Summary (motoring)	0.029	0.167	0.029	0.167	0.000	*NA*	0.000
Index offence: Theft	0.235	0.424	0.235	0.424	0.000	*NA*	0.000
Index offence: Violence	0.594	0.491	0.594	0.491	0.000	*NA*	0.000
Index offence: Other (Breach)	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
Index offence: Other (Child offence)	0.000	0.000	0.000	0.000	0.000	*NA*	0.000
History of drug offences (N)	0.065	0.266	0.078	0.296	0.017	1.243	0.023
History of weapons offences (N)	0.000	0.000	0.000	0.000	0.000	*NA*	0.064
History of public order offences (N)	0.006	0.074	0.010	0.097	0.013	1.707	0.054
History of robbery (N)	0.000	0.000	0.000	0.000	0.000	*NA*	0.007
History of theft (N)	0.922	1.884	0.897	1.898	-0.007	1.015	0.052
History of Violence (N)	0.071	0.280	0.070	0.279	-0.002	0.991	0.020
History of summary offences (N)	1.243	1.828	1.184	1.626	-0.018	0.791	0.106
Other history (N)	0.014	0.119	0.014	0.119	0.000	1.000	0.004

	Control (mean)	Control (SD)	RF EM (mean)	RF EM (sd)	Difference	Ratio	Overlap
Prior prison sentences (N)	0.183	0.661	0.189	0.709	0.003	1.152	0.008
Number of offences in probation disposal	0.381	0.670	0.367	0.675	-0.010	1.013	0.018
Number of requirements	1.951	0.807	1.951	0.807	0.000	1.000	0.001
Disposal length	285.151	239.122	287.841	236.306	0.011	0.977	0.054
OASys flags: Accommodation	0.525	0.499	0.525	0.499	0.000	*NA*	0.000
OASys flags: Employment	0.627	0.484	0.627	0.484	0.000	*NA*	0.000
OASys flags: Relationships	0.775	0.418	0.775	0.418	0.000	*NA*	0.000
OASys flags: Lifestyle & Associates	0.754	0.431	0.754	0.431	0.000	*NA*	0.000
OASys flags: Drug Misuse	0.541	0.498	0.541	0.498	0.000	*NA*	0.000
OASys flags: Alcohol Misuse	0.338	0.473	0.338	0.473	0.000	*NA*	0.000
OASys flags: Thinking & Behaviour	0.756	0.430	0.756	0.430	0.000	*NA*	0.000
OASys flags: Attitudes	0.762	0.426	0.762	0.426	0.000	*NA*	0.000
Year (requirement started): 2014	0.260	0.439	0.260	0.439	0.000	*NA*	0.000
Year (requirement started): 2015	0.244	0.430	0.244	0.430	0.000	*NA*	0.000
Year (requirement started): 2016	0.238	0.426	0.238	0.426	0.000	*NA*	0.000
Year (requirement started): 2017	0.159	0.365	0.159	0.365	0.000	*NA*	0.000
Year (requirement started): 2018	0.098	0.298	0.098	0.298	0.000	*NA*	0.000
PSM (distance)	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
Average					0.001		

	Control (mean)	Control (SD)
All (Effective Sample Size)	63,462	10,034
All (Unweighted)	63,462	10,034
Matched (Effective Sample Size)	623.6403	630
Matched (Unweighted)	830	630
Unmatched	62,632	9,404
Match rate		6%

PSM

	Control (mean)	Control (SD)	RF EM (mean)	RF EM (sd)	Difference	Ratio	Overlap
Age (at offence)	30.423	9.231	30.348	9.913	-0.008	1.153	0.060
Gender: Male	0.896	0.305	0.897	0.304	0.001	*NA*	0.000
Ethnicity: Asian Indian	0.011	0.102	0.010	0.101	-0.003	*NA*	0.000
Ethnicity: Asian Pakistani	0.021	0.143	0.023	0.149	0.013	*NA*	0.002
Ethnicity: Asian Bangladeshi	0.006	0.079	0.007	0.081	0.004	*NA*	0.000
Ethnicity: Asian Chinese	0.000	0.017	0.000	0.014	-0.007	*NA*	0.000
Ethnicity: Asian Other	0.006	0.075	0.006	0.075	0.000	*NA*	0.000
Ethnicity: Black Caribbean	0.034	0.180	0.032	0.177	-0.007	*NA*	0.001
Ethnicity: Black African	0.021	0.144	0.022	0.147	0.008	*NA*	0.001
Ethnicity: Black Other	0.007	0.083	0.007	0.086	0.007	*NA*	0.001
Ethnicity: White and Black Caribbean	0.027	0.162	0.024	0.154	-0.018	*NA*	0.003
Ethnicity: White and Black African	0.004	0.064	0.005	0.071	0.014	*NA*	0.001
Ethnicity: White and Asian	0.005	0.070	0.004	0.062	-0.018	*NA*	0.001
Ethnicity: Mixed Other	0.006	0.076	0.006	0.077	0.001	*NA*	0.000
Ethnicity: Arab	0.002	0.041	0.002	0.046	0.009	*NA*	0.000
Ethnicity: Other ethnicity	0.004	0.061	0.005	0.070	0.017	*NA*	0.001
Ethnicity: White British	0.812	0.391	0.808	0.394	-0.010	*NA*	0.004
Ethnicity: White Irish	0.010	0.099	0.009	0.095	-0.008	*NA*	0.001
Ethnicity: White Roma	0.007	0.086	0.008	0.091	0.010	*NA*	0.001
Ethnicity: White Other	0.019	0.135	0.021	0.145	0.020	*NA*	0.003

	Control (mean)	Control (SD)	RF EM (mean)	RF EM (sd)	Difference	Ratio	Overlap
Index offence: Criminal damage	0.006	0.077	0.005	0.074	-0.007	*NA*	0.000
Index offence: Drugs	0.083	0.276	0.085	0.278	0.006	*NA*	0.002
Index offence: Fraud	0.022	0.145	0.019	0.137	-0.018	*NA*	0.002
Index offence: Miscellaneous	0.041	0.198	0.042	0.200	0.004	*NA*	0.001
Index offence: Weapons	0.002	0.045	0.002	0.050	0.010	*NA*	0.000
Index offence: Public order	0.053	0.224	0.054	0.226	0.004	*NA*	0.001
Index offence: Robbery	0.011	0.103	0.011	0.103	0.000	*NA*	0.000
Index offence: Sex offences	0.009	0.096	0.009	0.093	-0.005	*NA*	0.000
Index offence: Summary	0.087	0.282	0.090	0.286	0.009	*NA*	0.002
Index offence: Summary (motoring)	0.105	0.306	0.107	0.309	0.007	*NA*	0.002
Index offence: Theft	0.251	0.434	0.246	0.430	-0.012	*NA*	0.005
Index offence: Violence	0.328	0.469	0.327	0.469	-0.002	*NA*	0.001
Index offence: Other (Breach)	0.002	0.041	0.002	0.048	0.013	*NA*	0.001
Index offence: Other (Child offence)	0.001	0.032	0.002	0.041	0.017	*NA*	0.001
History of drug offences (N)	0.218	0.808	0.213	0.753	-0.007	0.869	0.022
History of weapons offences (N)	0.019	0.169	0.022	0.175	0.014	1.070	0.064
History of public order offences (N)	0.020	0.261	0.020	0.302	0.003	1.348	0.054
History of robbery (N)	0.023	0.196	0.022	0.223	-0.004	1.303	0.008
History of theft (N)	1.731	3.675	1.653	3.742	-0.021	1.037	0.040
History of Violence (N)	0.104	0.511	0.107	0.502	0.007	0.964	0.021
History of summary offences (N)	1.692	2.968	1.725	3.205	0.010	1.166	0.093

	Control (mean)	Control (SD)	RF EM (mean)	RF EM (sd)	Difference	Ratio	Overlap
Other history (N)	0.225	0.677	0.218	0.696	-0.010	1.059	0.011
Prior prison sentences (N)	0.668	2.096	0.621	2.124	-0.022	1.027	0.018
Number of offences in probation disposal	0.888	1.418	0.875	1.409	-0.010	0.989	0.012
Number of requirements	2.314	1.107	2.338	1.151	0.020	1.081	0.099
Disposal length	319.946	222.056	326.749	251.556	0.027	1.283	0.132
OASys flags: Accommodation	0.390	0.488	0.394	0.489	0.008	*NA*	0.004
OASys flags: Employment	0.566	0.496	0.572	0.495	0.011	*NA*	0.005
OASys flags: Relationships	0.658	0.474	0.658	0.474	0.001	*NA*	0.000
OASys flags: Lifestyle & Associates	0.760	0.427	0.759	0.428	-0.003	*NA*	0.001
OASys flags: Drug Misuse	0.522	0.500	0.517	0.500	-0.010	*NA*	0.005
OASys flags: Alcohol Misuse	0.312	0.463	0.318	0.466	0.012	*NA*	0.005
OASys flags: Thinking & Behaviour	0.670	0.470	0.666	0.472	-0.008	*NA*	0.004
OASys flags: Attitudes	0.719	0.449	0.718	0.450	-0.002	*NA*	0.001
Year (requirement started): 2014	0.152	0.359	0.155	0.362	0.009	*NA*	0.003
Year (requirement started): 2015	0.181	0.385	0.181	0.385	-0.001	*NA*	0.000
Year (requirement started): 2016	0.261	0.439	0.256	0.436	-0.011	*NA*	0.005
Year (requirement started): 2017	0.232	0.422	0.233	0.423	0.004	*NA*	0.002
Year (requirement started): 2018	0.175	0.380	0.175	0.380	0.000	*NA*	0.000
PSM (distance)	0.175	0.087	0.175	0.087	0.000	1.001	0.000
Average					0.009		

	Control (mean)	Control (SD)
All (Effective Sample Size)	63,462	10,034
All (Unweighted)	63,462	10,034
Matched (Effective Sample Size)	10,024	10,024
Matched (Unweighted)	10,024	10,024
Unmatched	53,438	10
Match rate		100%