Table 1. Conceptual outline including key stakeholder groups, interests, focal metrics for analysis, key confounding variables and hypothesized outcomes.

Causal relationship	Sustainability interest	Metrics	Units key	Confounding variables	Hypothesised outcome
RSPO certification results in improved sustainability outcomes	Environmental	Orangutan	Orangutan presence	Concession size	RSPO > Non-RSPO
		Fire	Reduced number of fire incidents	Concession size	RSPO > Non-RSPO
	Social	Poverty	Reduced number of households receiving government assistance	Household density	RSPO > Non-RSPO
		Health Services	Availability of rural health facilities	Population density	RSPO > Non-RSPO
	Economic	Yield	Fresh fruit bunch (FFB) produced	Land under plantation	RSPO > Non-RSPO
		Profits	Share price	Total land bank	RSPO > Non-RSPO

From here, concession names were cross-referenced with data included in RSPO Annual Communication of Progress (ACOP) reports. As supply chain certificates are given to Palm Oil Mills (POM), the majority of the RSPO certified estate layer was derived from the record of RSPO certified mills and supply estates, with several estates often falling within one nationally recognized concession (figure 1). Additional verification was undertaken using the Sustainable Palm Oil Transparency Toolkit (Zoological Society of London Sustainable Palm Oil Platform (ZSL SPOM)) and from Global Forest Watch (GFW). Ninety one RSPO certified concessions identified, belonging to 41 companies, were then mapped in in ARC GIS v10 (see supplementary table S4 for further detail). These estates cover a total area of 5733 km^2 equating to 18% of planted palm oil concessions in Kalimantan and 22% of the area covered by RSPO certification globally. We note that our dataset likely excludes many small-and medium-sized oil palm growers because these estates are rarely registered at the national level (Gaveau *et al* 2016).

Analysis

Metrics selected to measure key sustainability impacts may be influenced by confounding factors (see table 1). To account for pre-existing variation in focus metrics and minimise rival explanations that may mimic or mask a relationship between cause (certification) and effect (sustainability outcome), we employed a propensity score matching technique. Propensity score matching was undertaken using the nearest neighbour method with calliper width of 0.25 standard deviations (Wang *et al* 2013). This method allows for outcome comparison between concessions that have been treated (RSPO certified) and comparable concessions that have not been treated (non-RSPO certified) at the concession level (Andam *et al* 2008, Caliendo and Kopeinig 2008, Ferraro 2009). Many factors are likely to be associated with both participation in RSPO and associated outcomes of the program. While we could not comprehensively include all of these variables, we included key confounding variables for which adequate data are available thereby controlling for variables that have not been considered in previous studies (e.g Levin *et al* 2012, Preusser 2016) These variables included concession size as larger concessions are theoretically more likely to host wide ranging, high conservation value species, and also have a higher risk of experiencing fire events. Village population density was controlled for when examining social sustainability metrics, as the number of health facilities should theoretically be positively correlated with increased population density. Similarly, larger populations should exhibit differential accumulation of wealth. Proportion of concession under oil palm crop plantation was controlled for when examining economic sustainability, as we were concerned with yields and profits as a result of plantation productivity rather than concession size (see table 1 and associated methodologies in supplementary material for further detail).

Matched data sets were then analysed using before and after control impact (BACI) analysis for each of the sustainability metrics highlighted in table 1. Whereby, certified and non-certified concessions were compared at a uniform point in time prior to the implementation of RSPO and again after implementation, allowing for the comparison of initial differences between samples that were to be treated and those that were not, as well as changes post treatment relative to both the starting point and treatment types (Conner *et al* 2016) Further details on metric selection and treatment are provided below.

Sustainability metrics

Environment

Orangutan presence and density

A central pillar of the RSPO is to manage palm oil plantations in a way that 'maintains and/or enhances' high conservation value (HCV) species (or fully protected species) (see table 1). As one example of a HCV and fully protected species, Bornean Orangutans (*Pongo pygmaeus*) are to be monitored and protected by palm oil growers. As the species has also become a global icon for biodiversity conservation efforts in the face of the rapid and continued expansion of