Appendix A: Methodology

Data structure

We analyse overseas entities (OEs) using the Companies House Basic Data Product, and registered beneficial owners (RBOs) using the Persons (PSC) Register Data Product, both published by Companies House.[1] We analyse titles via the Overseas Companies Ownership Database (OCOD), published by HM Land Registry.

Variables are constructed from the source data at three levels, as follows:

1. RBO-OE pairs: the unit of observation is RBO-OE pair because some RBO characteristics (e.g. capacity) can vary by OE.
2. OEs
3. Titles (converted to properties on a one-to-many basis using the Bourne-McKenzie algorithm.[2]

Matching procedures

To begin our analysis, we match the source datasets together. The main difficulty is that there is no single unique identifier that allows to match overseas entities across all three data sources.

Step 1: OCOD to Basic Data Product. We match entities in OCOD to CH Basic Data Product based on name and country of incorporation. It is worth noting that this is done in several rounds, in order to minimise the number of false matches while having a maximum number of matches. The first round is the most restrictive, and matched companies in OCOD are removed once matched. There can be more than one perfect match between OCOD and CH, usually when the same overseas entity has registered both as a foreign company (CH) and as an overseas entity (OE). When this is the case, we always keep the entry as an overseas entity. The matching involves name cleaning, because the same company may be recorded either as 'XXX ltd.', 'XXX ltd' or 'XXX limited', and fuzzy string matching.

Step 2: Basic Data Product to (and from) PSC register. It is straightforward to go from OEs in the Basic Data Product to RBOs in the PSC Register, as the PSC Register is already structured in OE-RBO pairs, and the company number of the OE is used as an identifier in both datasets. However, matching corporate entities that are registered as RBOs to entities in the Basic Data product is not straightforward. All corporate RBOs are required to report a company identifier, but this identifier only works for companies registered in the UK. OEs do not usually provide their CH company number, but instead the one in their original

---

[1] The ROE does not exist as a separate database. It is incorporated into these existing Companies House data products.
[2] Bourne, Ingianni & McKenzie (2023).

jurisdiction. Consequently, they must be matched on name and jurisdiction, like in step 1, which creates the risk of false matches, and reduces the scope. For UK companies, the CH company number is usually provided, but identifying which companies are registered in CH is not necessarily straightforward, as the relevant variable in the PSC register (data.identification.legal_authority) is not standardised, and there are 789 different values taken for just 3098 different companies, and 57 different ones just for the UK (we use character string analysis to identify them).

Step 3. PSC register to (and from) Orbis. We match corporate RBOs to BVD Orbis – a commercial dataset – based on name and country. This allows us to know if the entities are listed, which is one of the conditions on which a company becomes an RBO and exempts the OE from providing details of the RBO's individual beneficial owners.

RBO level

Type

- RBO_TYPE
    - Categorical variable describing statutory type of RBO
    - Values:
        - 1=individual
        - 2=other legal entity
        - 3=government/public authority
    - From PSC Register data.kind:
        - 1='individual-beneficial-owner' OR 'individual-person-with-significant-control' OR 'super-secure-beneficial-owner' OR 'super-secure-person-with-significant-control'
        - 2='corporate-entity-beneficial-owner' OR 'corporate-entity-person-with-significant-control'
        - 3='legal-person-beneficial-owner' OR 'legal-person-person-with-significant-control'
- RBO_PROTECTED_FLAG
    - 1 if individual covered by statutory protection regime
    - PSC Register data.kind
        - 1='super-secure-beneficial-owner' OR 'super-secure-person-with-significant-control'

Interest

- RBO_BENEFICIARY_FLAG
    - 1 if RBO has ownership of shares or other beneficial interest
    - PSC Register data.natures_of_control contains "ownership-of-shares" or "share-surplus-assets"
- RBO_CONTROLLER_FLAG
    - 1 if RBO has significant influence or control
    - PSC Register data.natures_of_control contains "right-to-appoint" or "voting-rights" or "significant-influence-or-control"

Capacity

- RBO_TRUSTEE_FLAG
  - 1 if RBO is acting as trustee
  - PSC Register data.natures.of.control contains 'as-trust'
- RBO_PARTNER_FLAG
  - 1 if RBO exercises significant influence or control over the activities of a partnership,
    unincorporated association or other entity (Condition 5).
  - Notes: applies to both individuals and legal entities
  - PSC Register data.natures.of.control contains 'as-firm'
- RBO_GENERALPARTNER_NAME_FLAG
  - 1 if RBO is acting as a corporate general partner
  - Notes: limited partners are excluded because these cannot reliably be distinguished from limited *partnerships* (LP is ambiguous). Does not apply to individuals as these cannot be identified from name.
  - 1 if RBO_TYPE=2 (other legal entity) and Company Name includes any of the following terms: 'general partner', gp, g.p
- RBO_NOMINEE_NAME_FLAG
  - 1 if RBO_OE_TYPE=2 (other legal entity) and Company Name includes any of the following terms: nominee, nominees
- RBO_OWNCAPACITY_FLAG
  - 1 if RBO_TRUSTEE_FLAG=0 and RBO_PARTNER_FLAG=0 and RBO_NOMINEE_NAME_FLAG=0 and RBO_GENERALPARTNER_NAME_FLAG = 0

Corporate BOs

- RBO_IS_OE_FLAG
  - 1 if RBO is a registered OE
  - 1 if RBO_TYPE=2 and Company Number has OE prefix
- RBO_PSC_SCOPE_FLAG
  - 1 if RBO is a (non-OE) legal entity within scope of PSC Register
  - 1 if RBO_TYPE=2 and RBO_IS_OE_FLAG=0 and company name is matched in CH basic data product and is within qualifying CompanyCategory (see online appendix)
- RBO_PSC_MATCHED_FLAG
  - 1 if RBO is a legal entity and matched as (non-OE) owned entity in PSC Register
  - 1 if RBO_TYPE=2 and RBO_PSC_SCOPE_FLAG=1 and RBO is matched in PSC Register (=matching company name to company number in CH data, then by company number in PSC Register)
- RBO_REGOVERSEASTRUSTEE_FLAG
  - 1 if RBO is (likely) regulated overseas trust service provider
  - 1 if RBO_TYPE=2 and RBO_IS_OE_FLAG=0 and RBO_PSC_SCOPE_FLAG=0 and RBO_TRUSTEE_FLAG=1
- RBO_IS_OE_UBOIDENTIFIED_FLAG

- o 1 if RBO is a registered OE and it has been possible (within 10 search iterations) to identify at least one individual owner BO acting as own capacity OR a listed company
- o 1 if RBO_IS_OE_FLAG=1 and somewhere up the chain: [RBO_OWNCAPACITY_FLAG = 1 and RBO_BENEFICIARY_FLAG = 1] OR RBO_QUAL_LISTED_FLAG = 1

- RBO_QUAL_LISTED_FLAG
  - o 1 if RBO is listed on qualifying market
  - o 1 if RBO_TYPE=2 and RBO_IS_OE_FLAG=0 and RBO_PSC_SCOPE_FLAG=0 and RBO_REGOVERSEASTRUSTEE_FLAG=0 and entity is matched in Orbis and shown as listed in qualifying market (see online appendix)
- RBO_CORPRESIDUAL_FLAG
  - o 1 if RBO is legal entity but appears not to meet any of criteria for "subject to own disclosure requirements"
  - o 1 if RBO_TYPE=2 and RBO_IS_OE_FLAG=0 and RBO_PSC_SCOPE_FLAG=0 and RBO_REGOVERSEASTRUSTEE_FLAG=0 and RBO_QUAL_LISTED_FLAG=0

Validity

- RBO_SUFFICIENT_FLAG
  - o 1 if the report identifies at least one individual with a beneficial interest who is acting in their own capacity, or the RBO is listed on a qualifying market or is a government/public authority.
  - o 1 if either:
    - ▪ RBO_TYPE=1 and RBO_BENEFICIARY_FLAG=1 and RBO_OWNCAPACITY_FLAG=1
    - ▪ RBO_TYPE=2 and RBO_BENEFICIARY_ FLAG =1 and RBO_OWNCAPACITY_FLAG=1 and [RBO_IS_OE_UBOIDENTIFIED_FLAG=1 or RBO_QUAL_LISTED_FLAG=1]
    - ▪ RBO_TYPE=3
  - o Otherwise set to 0
- RBO_MISSINGBO_FLAG
  - o 1 if the information for the RBO indicates that there may be missing beneficial owners of the land whose identities are not available
  - o 1 if either:
    - ▪ RBO_TYPE=1 and RBO_OWNCAPACITY_FLAG=0
    - ▪ RBO_TYPE=2 and RBO_OWNCAPACITY_FLAG=0 or
    - ▪ RBO_TYPE=2 and RBO_IS_OE_UBOIDENTIFIED_FLAG=0 and RBO_QUAL_LISTED_FLAG=0
  - o Otherwise set to 0

OE level

Type

- OE_NON_OE_PREFIX_FLAG
    - 1 if OE matched with proprietor in LR data does not have 'OE' as prefix, which means that it is not an 'Overseas Entity', and therefore does not have to report any information on trusts.

Location

- OE_COUNTRY_INCORP
    - Country of incorporation.
    - CH data 'CountryOfOrigin'
    - Ideally country name is cleaned and converted to ISO-3 country code, if not already clean
- OE_COUNTRY_ADDRESS
    - Country of registered address.
    - CH data 'RegAddress.Country'
    - Ideally country name is cleaned and converted to ISO-3 country code, if not already clean
- OE_UK_ADDRESS_FLAG
    - 1 if OE_COUNTRY_ADDRESS is UK

Ownership chain

- OE_MATCHED_TITLE_COUNT
    - Count of titles matched with OE (via fuzzy-matching to proprietor on name)
    - 0 if Company Name not matched to any proprietors
- OE_MATCHED_RBO_COUNT
    - Count of RBOs matched with OE (linked by company number in PSC Register)
    - 0 if OE is not found in PSC Register
- OE_BENEFICIARY_COUNT
    - Count of beneficiaries in the PSC register (based on RBO_BENEFICIARY_FLAG)
    - 0 if OE is not found in PSC Register
- OE_OWNS_OE_COUNT
    - Count of other OEs owned by this OE
    - Number of times this OE appears as RBO in OE-RBO pair
- OE_OWNEDBY_OE_COUNT
    - Count of other OEs that own this OE
    - Number of RBOs where RBO_IS_OE_FLAG=1

Capacity

- OE_QUAL_LISTED_FLAG
    - 1 if OE reports no RBO and is listed on a qualifying regulated market.

- o OE_MATCHED_RBO_COUNT =1 and Entity is shown as listed in Orbis <u>and</u> the market is one of the qualifying regulated markets within UK/EU/EEA/USA/Japan/Switzerland/Israel (see online appendix).
- OE_TRUSTEE_NAME_FLAG
  - o 1 if Company Name includes any of the following terms: trust, trusts, trustee, trustees
- OE_TRUSTEE_ADDRESS_FLAG
  - o 1 if registered address matches registered/office address of known trust services provider (see online appendix)
- OE_GENERALPARTNER_NAME_FLAG
  - o 1 if OE is acting as a general partner
  - o Notes: limited partners are excluded because these cannot reliably be distinguished from limited *partnerships* (LP is ambiguous).
  - o 1 if Company Name includes any of the following terms: 'General Partner', GP, G.P
- OE_NOMINEE_NAME_FLAG
  - o 1 if Company Name includes any of the following terms: nominee, nominees
- OE_OWNCAPACITY_FLAG
  - o 1 if the OE owns the property in it's own capacity i.e. on trust for someone else or as a partner
  - o 1 if OE_TRUSTEE_NAME_FLAG=0 and OE_GENERALPARTNER_NAME_FLAG=0 and OE_NOMINEE_NAME_FLAG=0
  - o Note: we may want to expand these conditions based on address flag, depending on outcome of validation process

BO statement

- OE_NOBO_STATEMENT_FLAG
  - o 1 if OE states no beneficial owners
  - o PSC Register data.statement == 'no-beneficial-owner-identified' OR data.statement == 'no-individual-or-entity-with-signficant-control'
- OE_MISSINGBO_STATEMENT_FLAG
  - o 1 if OE states only some BOs identified (whether or not full information provide)
  - o PSC Register data.statement == 'at-least-one-beneficial-owner-unidentified' OR data.statement == 'at-least-one-beneficial-owner-unidentified-and-information-not-provided-for-at-least-one-beneficial-owner'
- OE_ALLBO_STATEMENT_FLAG
  - o 1 if OE states that all BOs identified and all info provided
  - o PSC Register data.statement == 'all-beneficial-owners-identified'
- OE_MISSINGINFO_STATEMENT_FLAG
  - o 1 if OE states that all BOs identified but some info missing
  - o PSC Register data.statement == 'information-not-provided-for-at-least-one-beneficial-owner'

RBO summary

- OE_INDIVIDUAL_RBO_COUNT
    - Number of RBOs who are individuals
    - Count of RBOs with RBO_TYPE=1
- OE_CORPBO_RBO_COUNT
    - Number of RBOs who are corporate beneficial owner
    - Count of RBOs with RBO_TYPE=2
- OE_GOV_RBO_COUNT
    - Number of RBOs who are government/public authority
    - Count of RBOs with RBO_TYPE=3
- OE_INDIV_TRUSTEE_RBO_COUNT
    - Number of individual trustee RBOS
    - Counts of RBOS with [RBO_TRUSTEE_FLAG=1 and RBO_TYPE=1]
- OE_CORPORATE_TRUSTEE_RBO_COUNT
    - Number of individual trustee RBOS
    - Counts of RBOS with [RBO_TRUSTEE_FLAG=1 and RBO_TYPE=2]
- OE_RBO_PARTNER_COUNT
    - Number of RBOs who are partners
    - Counts of RBOS with [RBO_PARTNER_FLAG =1]
- OE_RBO_INDIVPARTNER_COUNT
    - Number of RBOs who are individual partners
    - Counts of RBOS with [RBO_PARTNER_FLAG = and RBO_TYPE=1]
- OE_RBO_CORPPARTNER_COUNT
    - Number of RBOs who are individual partners
    - Counts of RBOS with [RBO_PARTNER_FLAG = and RBO_TYPE=2]
- OE_RBO_GOVPPARTNER_COUNT
    - Number of RBOs who are individual partners
    - Counts of RBOS with [RBO_PARTNER_FLAG = and RBO_TYPE=3]
- OE_RBO_GENERALPARTNER_COUNT
    - Number of RBOs who have 'general partner' or equivalent in their name
    - Counts of RBOs with [RBO_GENERAL_PARTNER_NAME_FLAG=1]
- OE_RBO_OE_UNIDENTIFIED_BO_COUNT
    - Number of OE RBOs with no identified UBOs
    - Counts of RBOS with [RBO_IS_OE_FLAG=1 and RBO_IS_OE_UBOIDENTIFIED_FLAG = 0]
- OE_RBO_LISTED_COUNT
    - Number of listed RBOS
    - Counts of RBOS with RBO_QUAL_LISTED_FLAG =1
- OE_RBO_REGOVERSEASTRUSTEE_COUNT
    - Number of RBOS that are regulated overseas trusts
    - Counts of RBOs with RBO_REGOVERSEASTRUSTEE_FLAG =1

Validity

- OE_SUFFICIENTBO_FLAG
  - 1 if the OE or an RBO is listed on a qualifying market or is a government/public authority, or at least one RBO identifies at least one individual with a beneficial interest who is acting in their own capacity.
  - 1 if OE_QUAL_LISTED_FLAG=1 or at least one RBO with RBO_SUFFICIENT_FLAG=1 and OE_GENERALPARTNER_NAME_FLAG= 0
- OE_MISSINGBO_FLAG
  - 1 if the information provided indicates that there may be missing beneficial owners of the land whose identities are not available
  - 1 if there is at least one RBO with RBO_MISSINGBO_FLAG=1 and/or OE_OWNCAPACITY_FLAG=0 and/or OE_MATCHED_RBO_COUNT = 0
- OE_VALID_REPORT_FLAG
  - 1 if OE has at least one sufficient beneficial owner or is listed on a qualifying market, and has no missing beneficial owners
  - 1 if [OE_SUFFICIENTBO_FLAG=1 or and OE_MISSINGBO_FLAG=0] or OE_QUAL_LISTED_FLAG=1

Source of invalidity

- OE_INVALID_OENOBO_FLAG
  - 1 if OE has no sufficient beneficial owners and is not listed on a qualifying market
  - 1 if OE_SUFFICIENTBO_FLAG=0 and OE_QUAL_LISTED_FLAG=0
- OE_INVALID_OENOMINEE_FLAG
  - 1 if OE is nominee
  - 1 if OE_NOMINEE_NAME_FLAG=1
- OE_INVALID_OETRUST_FLAG
  - 1 if OE is trustee
  - 1 if OE_TRUSTEE_NAME_FLAG=1
- OE_INVALID_OEPARTNER_FLAG
  - 1 if OE is partner
  - 1 if OE_GENERALPARTNER_NAME_FLAG=1
- OE_INVALID_RBOTRUST_FLAG
  - 1 if at least one RBO is acting as trustee
  - 1 if at least one RBO has RBO_TRUSTEE_FLAG=1
- OE_INVALID_RBOPARTNER_FLAG
  - 1 if at least one RBO is acting as partner
  - 1 if at least one RBO has RBO_PARTNER_FLAG=1 or RBO_GENERALPARTNER_NAME_FLAG
- OE_INVALID_RBOPSC_FLAG
  - 1 if at least one RBO is within scope of PSC Register (meaning that it is not possible to ascertain whether the ownership structure involves a trust)
  - 1 if at least one RBO has RBO_PSC_SCOPE_FLAG=1
- OE_INVALID_RBOCORPRESIDUAL_FLAG
  - 1 if at least one RBO does not appear to have any grounds for qualifying as 'subject to own disclosure requirements'
  - 1 if at least one RBO has RBO_CORPRESIDUAL_FLAG=1

Title level

Proprietor counts

- TITLE_TOTAL_OE_COUNT
  - Count of OE proprietors of title
  - Total number of proprietors shown in OCOD
- TITLE_MATCHED_OE_COUNT
  - Count of OE proprietors of title that are matched in CH data
  - Based on fuzzy-matching Company Name
- TITLE_UNMATCHED_OE_COUNT
  - Count of OE proprietors of title that are not matched in CH data
  - Based on fuzzy-matching Company Name
- TITLE_MATCHED_LISTED_OE_COUNT
  - Count of OE proprietors that are listed on a qualifying market
  - Based on OE_QUAL_LISTED_FLAG
- TITLE_MATCHED_MISSINGBO_COUNT
  - Count of OE proprietors that report missing BO statements
  - Based on OE_MISSINGBO_STATEMENT_FLAG
- TITLE_ADDITIONAL_PROP_FLAG
  - Title has additional (non-OE) proprietors
  - 1 if OCOD 'Additional Proprietor Indicator'=Y (and equivalent for COCR if available)
- TITLE_MATCHED_BENEFICIARY_COUNT
  - Number of owners matched to a beneficiary (based on OE_BENEFICIARY_COUNT)
- TITLE_MATCHED_CORPO_BENEFICIARY_COUNT
  - Number of owners with a corporate BO
- TITLE_RBO_FLAG
  - 1 if there is at least one RBO
- TITLE_RBO_PARTNER_COUNT
  - Sum of proprietors where there is at least one RBO declaring to be a partner
  - Based on OE_RBO_PARTNER_COUNT
- TITLE_RBO_GENERALPARTNER_COUNT
  - Sum of proprietors where there is at least one RBO with a general partner name
  - Based on OE_RBO_GENERALPARTNER_COUNT
- TITLE_OE_RBO_COUNT
  - Sum of proprietors with at least one OE RBO
  - Based on OE_OWNS_OE_COUNT
- TITLE_OE_RBO_NO_UBOIDENTIFIED_COUNT
  - Sum of proprietors with OE RBOs with no identified OEs
  - Based on OE_RBO_OE_UNIDENTIFIED_BO_COUNT
- TITLE_OE_RBO_LISTED_COUNT
  - Sum of proprietors with at least one listed RBO
  - Based on OE_RBO_LISTED_COUNT
- TITLE_OE_RBO_REGOVERSEASTRUSTEE_COUNT
  - Sum of proprietors with at least one overseas regulated trustee RBO
  - Based on OE_RBO_REGOVERSEASTRUSTEE_COUNTTitle characteristics

- TITLE_TRANSACTION_DATE
  - OCOD 'Date Proprietor Added' (and equivalent for COCR)
- TITLE_LONDON_FLAG
  - 1 if Region = 'Greater London'
  - 0 by construction for COCR
- TITLE_PRICEINFO_FLAG
  - 1 if Price Paid is non-missing
- TITLE_VALUE
  - Price paid value in the data. When prices are duplicated; which happens when multiple titles are acquired in one transaction, the total value is attributed equally between the titles.
  - Where price paid is missing, NA value

Validity

- TITLE_VALID_OE_COUNT
  - Total number of matched OEs with valid information
  - Count of OEs with OE_VALID_REPORT_FLAG=1
- TITLE_INVALID_OE_COUNT
  - Total number of matched OEs with invalid information
  - Count of OEs with OE_VALID_REPORT_FLAG=0
- TITLE_VALID_REPORT_FLAG
  - 1 if title has at least one valid OE and no invalid OEs or unmatched OEs
  - 1 if TITLE_VALID_OE_COUNT>0 and TITLE_INVALID_OE_COUNT=0 and TITLE_UNMATCHED_OE_COUNT=0

Source of invalidity

- TITLE_UNMATCHED_OE_FLAG
  - 1 if title has at least one OE proprietor that is not matched in CH data
  - 1 if TITLE_UNMATCHED_OE_COUNT>0
- TITLE_INVALID_OENOBO_FLAG
  - 1 if title has at least one OE that has no sufficient beneficial owners and is not listed on a qualifying market
  - 1 if at least one OE has OE_INVALID_OENOBO_FLAG=1
- TITLE_INVALID_OENOMINEE_FLAG
  - 1 if title has at least one OE that is nominee
  - 1 if at least one OE has OE_INVALID_OENOMINEE_FLAG=1
- TITLE_INVALID_OETRUST_FLAG
  - 1 if title has at least one OE that is trustee
  - 1 if at least one OE has OE_INVALID_OETRUST_FLAG=1
- TITLE_INVALID_OEPARTNER_FLAG
  - 1 if title has at least one OE that is partner
  - 1 if at least one OE has OE_INVALID_OEPARTNER_FLAG=1
- TITLE_INVALID_RBOTRUST_FLAG
  - 1 if title has at least one OE with at least one RBO acting as trustee
  - 1 if at least one OE has OE_INVALID_RBOTRUST_FLAG=1
- TITLE_INVALID_RBOPARTNER_FLAG
  - 1 if title has at least one OE with at least one RBO acting as partner
  - 1 if at least one OE has OE_INVALID_RBOPARTNER_FLAG=1

- TITLE_INVALID_RBOPSC_FLAG
    - 1 if title has at least one OE with at least one RBO that is within scope of PSC Register
    - 1 if at least one OE has OE_INVALID_RBOPSC_FLAG=1
- TITLE_INVALID_RBOCORPRESIDUAL_FLAG
    - 1 if title has at least one OE with at least one RBO that does not appear to have any grounds for qualifying as 'subject to own disclosure requirements'
    - 1 if at least one OE has OE_INVALID_RBOCORPRESIDUAL_FLAG