Bank of Finland Research Discussion Papers
12 • 2025

Ville Voutilainen

Estate owners' ensemble— Mapping commercial real estate concentration using Finnish firm ownership

Bank of Finland
Research

Bank of Finland Research Discussion Papers
Editor-in-Chief Esa Jokivuolle

Bank of Finland Research Discussion Papers 12/2025
9 December 2025

Ville Voutilainen
Estate owners' ensemble— Mapping commercial real estate concentration using Finnish firm ownership

ISSN 1456-6184, online

Bank of Finland
Research Unit

PO Box 160
FIN-00101 Helsinki

Phone: +358 9 1831

Email: research@bof.fi
Website: www.suomenpankki.fi/en/research/research-unit/

Estate Owners’ Ensemble — Mapping Commercial Real Estate Concentration using Finnish Firm Ownership Network

Ville Voutilainen *

Bank of Finland
ville.voutilainen@bof.fi

First version: July 30, 2025
This version: March 4, 2026

Abstract

The commercial real estate (CRE) market is an important source of financial stability risks, yet ownership structures remain opaque. This paper uses comprehensive Finnish register data to construct firm-level ownership network and identify owners of CRE firms. We document that government entities are the most important ultimate owners, holding about 10% of the sector’s balance sheet. We show that government ownership predicts lower interest rate spreads on CRE bank loans, consistent with creditors perceiving such firms as less risky. Our results highlight the need to incorporate ownership structures into financial stability assessments and credit risk models.

Keywords: *CRE; networks; firm ownership; loan pricing; financial stability.*
JEL Classification: *R33; G10; C63.*

1 Introduction

The real estate market is among the core topics of financial stability: of the 50 systemic banking crises in recent decades, more than two-thirds were preceded by boom-bust patterns in house prices (IMF, 2019). Real estate can be divided into *residential real estate* (RRE) and *commercial real estate* (CRE). The latter is considered the more cyclical of the two and can affect the economy through several channels.[1] Despite CRE’s importance, ownership

*I thank Maija Järvenpää for the support in data access and for many insightful discussions. I am also grateful to Essi Eerola, Esa Jokivuolle, Kimmo Koskinen, Antti Ripatti, Juhana Siljander, and the participants of the Bank of Finland’s research seminars for their valuable suggestions. Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the Bank of Finland or the Eurosystem. During the preparation of this work, I used Microsoft Copilot to reword and rephrase text. After using this tool/service, I reviewed and edited the content as needed and take full responsibility for the content of the publication. All mistakes are my own.

[1]First, CRE can represent a significant portion of assets held by different institutions. Swings in asset valuations can have major effects through the collateral channel. Second, as construction activity constitutes large shares of countries’ GDPs, sharp downturns in the CRE market often have consequences for the real economy. Third, CRE constitutes a non-trivial share of banks’ lending portfolios. Defaults in the CRE market can trigger increases in non-performing loan shares. Fourth, changes in CRE market prices can impact stock markets through real estate investment funds and trusts (REIFs and REITs).

structures remain largely opaque. This opacity matters for financial stability: when prices collapse, who ultimately bears the losses? And do implicit guarantees, such as government ownership, influence credit markets?

We'll provide answers to these questions by identifying owners of CRE firms in the Finnish economy. We show that while other CRE firms are the most important direct owners of CRE firms, government entities are the single most important owner group at the ultimate-owner level. Further, a surprisingly large share of the total CRE can be identified as being owned either by government or non-profit entities. This suggests that a non-trivial share of the CRE sector in Finland has an implicit backstop. To demonstrate the usefulness of this result, we investigate the association between the pricing of CRE bank loans and the government ownership of the debtors using cross-sectional regressions. We show that government ownership predicts lower loan rates compared to a case without government ownership.

The analysis rests on comprehensive Finnish microdata. While there has been some recent advancement in untangling the complex CRE exposure landscape (Daly, Ryan, & Schwartz Blicke, 2024), this study is, to my knowledge, the first to offer a sector-wide deep dive into CRE ownership.[2] By focusing on a single country with rich register data, it is possible to carefully define commercial real estate firms without the complications arising from jurisdictional differences. Further, it allows us to capture a major part of inter-firm ownership within the economy. Our hope is that this work can help address the daunting data gaps related to CRE (European Systemic Risk Board, 2023) and spur similar endeavors in other jurisdictions.

More concretely, we use firm background information to cast firms into groups of interest: *CRE firms*, *Financial firms*, *Government firms*, and *Holding companies*. Using tools from network analysis, we construct a network describing owner-owned links between Finnish firms and derive ownership chains that reveal the entire ownership structure. The chains are used to identify and analyze owners of CRE firms. Further, we investigate whether government ownership of a CRE debtor predicts the pricing of its bank loans by employing several regression models in which bank rates of CRE loans are regressed on the ownership status of the debtor.

The constructed ownership network covers a lion's share of the Finnish firm landscape, measured by share of total balance sheet amount captured. While the network is sparse, it is far from trivial, with about 40% of the firms in the network belonging to the same giant substructure. This observation underlines the usefulness of a network-based approach. Of the total CRE firm balance sheet amount of EUR 159.4 billion, a direct firm-owner can be identified for 45.0%. About half of this is covered by other CRE firms, making them the most important direct owner group of CRE firms. An ultimate owner (at the beginning of an ownership chain) can be identified for about a third of the CRE firms' balance

[2]There are three types of exposure to CRE (Daly et al., 2024): i) exposure through ownership of physical estate; ii) exposure through equity investments (i.e., holding shares in a company); and iii) exposure through funding (banks and bond investors). This study focuses on the second type. As explained in Section 3, neglecting direct estate ownership is not considered a major limitation. Exposures through funding are outside the scope of this work.

sheet amount. Government entities emerge as the most important owner group at the ultimate-owner level, with a 28.2% share of the identified ownerships. This translates to the government owning 10.3% of the total CRE balance sheet amount in Finland. When the set of government entities is broadened to include all public and non-profit entities, the attributed ownership share of the total CRE balance sheet amount increases to 14.0%. These observations highlight that a non-negligible portion of CRE firms in the Finnish economy is (directly or indirectly) owned by entities that are backed, at least implicitly, by taxpayer money or benefit from tax exemptions due to their social benefit role.

Results from the bank loan rate regressions provide robust evidence that government ownership of CRE firms is associated with the price such firms are charged on their bank loans. The effect in our preferred model is -33 basis points, on average—a statistically and economically significant result. In words: government-owned CRE firms are charged, on average, one-third of a percentage point lower rates than CRE debtors with non-government owners. The result is robust to including debtors without an identified owner in the comparison group. A plausible explanation for the finding is that creditors perceive government ownership as a factor that increases a firm's creditworthiness, likely due to expectations of implicit guarantees or bailouts.

The rest of the paper is structured as follows. Section 2 provides institutional context and reviews related literature on real estate analysis, network analysis, as well as bank loan pricing. Section 3 describes the data. Section 4 constructs the ownership network and uses it to analyze owner of CRE firms. Section 5 uses the identified ownerships to analyze the association between government ownership and the pricing of bank loans to CRE firms. Finally, Section 6 concludes with a discussion of the findings and possible extensions.

2 Institutional setting and related literature

2.1 Real estate analysis in macroeconomics

Despite the recognized importance of real estate markets—particularly the CRE market—for the functioning and stability of the financial system, the related literature remains relatively scarce. We categorize the existing literature into the following streams:

Real estate market and the business/financial cycle. Papers in this category investigate links between real estate market and the business and/or financial cycles. Herring and Wachter (1999) provide evidence that real estate markets are vulnerable to optimism and thus booms often end in banking busts. M. A. Davis and Heathcote (2005) attribute the high volatility of residential investments to being construction intensive and to the fact that the residential structures depreciate very slowly. Leamer (2007), in turn, argues that housing is the single-most critical part of the U.S. business cycle. Jaccard (2021) finds that policies enhancing risk-sharing between lenders and borrowers reduce the magnitude of boom-bust cycles in real estate prices. Finally, Duca, Muellbauer, and Murphy (2021) provide an overview of literature focusing on real estate price cycles and its implications on economic activity. In comparison, this paper adopts a more focused approach: it examines

CRE specifically, rather than real estate as a whole, and identifies CRE owners, which can serve as a building block for further macroeconomic analyses.

CRE market, leverage, and bank risk. Findings by E. P. Davis and Zhu (2011) suggest that bank lending is closely linked to commercial property prices, and that CRE cycles are largely driven by dynamic interactions between the commercial property sector, bank credit, and the broader macroeconomy. Kragh-Sørensen and Solheim (2014) argue that CRE lending is the most important source of losses for banks' lending books during crises. Relatedly, Shibut and Singer (2015) compare the loss-given-default (LGD) values of loans from failed banks and find that LGDs are higher for construction and development loans than for other CRE and commercial loans. Antoniades (2015) claims that the primary driver of commercial bank failures during the Great Recession was exposure to the real estate sector, especially exposure to non-household real estate borrowers. Our paper relates to this literature by examining whether banks price CRE loans to government-owned debtors more favorably than without government ownership, entailing how risky banks perceive such debtors.

CRE and financial stability policy analysis. CRE markets are of particular interest to central banks, supervisory bodies as well as macroprudential authorities (e.g., European Central Bank, 2022). The CRE market is less understood than the RRE market due to its opaque nature, complexity, and data gaps.[3] A related contribution comes from the European Central Bank (Daly et al., 2024). It lays out a system-wide mapping of CRE exposures in the euro area, including both investment (asset holdings and equity investments) and financing (bank loans) exposure channels. Authors identify real estate companies, real estate investment funds and real estate investment trusts as having a particularly large CRE exposure. While demonstrating clear need for such analysis, the multinational (euro area) focus and reliance only on commercial data vendors can provide only a shallow analysis of all possible key players. In particular, it excludes most public (tax payer backed) institutions. Instead, by leveraging high-quality registry data from a single country, our paper provides a more comprehensive view of the investment (ownership) exposures of different sectors to the CRE market.

2.2 Network analysis

A *network* is a set of interconnected entities, such as people or institutions. *Network analysis* represents networks as graphs—that is, structures composed of objects where pairs of objects are connected in a specific way—and applies methods from graph theory to study these relationships. Owing to its generality, network analysis has a wide range of applications across many disciplines.

When applied to economics, network analysis examines the relationships among economic agents and how these interactions influence economic outcomes. It has been widely used to study a broad range of economic phenomena, such as financial market structures

[3]Dagrain et al. (2019) outlines these data gaps in the European context and proposes methodologies for CRE policy analysis.

via interbank flows (e.g., Boss, Elsinger, Summer, and Thurner, 2004; Soramäki, Bech, Arnold, Glass, and Beyeler, 2007; Craig and von Peter, 2014), money markets (Iori, De Masi, Precup, Gabbi, & Caldarelli, 2008), international trade (e.g., Chaney, 2016; Baqaee and Farhi, 2024) as well as contagion between financial market players (e.g., Allen and Gale, 2000; Battiston, Puliga, Kaushik, Tasca, and Caldarelli, 2012; Acemoglu, Ozdaglar, and Tahbaz-Salehi, 2015). Literature reviews include Engel, Nardo, and Rancan (2021), Bardoscia et al. (2021), (Jackson, 2016), and Bramoullé, Galeotti, and Rogers (2016). The use of network data in econometrics is discussed in Graham (2020).

Papers adopting a network approach to investigation of firm ownership and control include Engel et al. (2021), who demonstrate how analyzing network structure can uncover crucial concentrations of ownership power; Borsos and Stancsics (2020), who construct an ownership network and use it to provide interesting descriptive statistics; and Vitali, Glattfelder, and Battiston (2011), who identify a global economic "super-entity" of corporate control. Our paper is closely related to these efforts, applying ownership information specifically to CRE markets.

2.3 Bank loan pricing

There is ample research investigating how specific factors are associated with corporate bank loan spreads. These include, for example, Santos, 2010 (greater increases in loan spreads for firms borrowing from more distressed banks following the Subprime Crisis), Santos and Winton, 2019 (bank capital and borrower bargaining power), Gambacorta and Mistrulli, 2014 (bank–firm relationship characteristics), Mattes, Steffen, and Wahrenburg, 2010 (information rents), and Ashraf, 2021 (economic uncertainty). Yet, papers taking a broader view of the general determinants of loan pricing are scarce. A notable exception is a working paper by Beim (1996), which systematically maps factors influencing bank loan pricing using a sample of U.S. corporate loans. Our exercise continues the tradition of papers focusing on a single pricing factor by examining the impact of government ownership of the debtor.

There are a few papers that examine ownership structure and lending. First, Aslan and Kumar (2012) provide evidence that greater control concentration among debtors positively affects debt pricing, although their analysis does not explicitly focus on government-owned borrowers. Second, focusing explicitly on government ownership and bank lending, Sapienza (2004) examines the ownership of banks—rather than debtors—and its influence on bank lending behavior. Perhaps the closest related work is Borisova and Megginson (2011), who explore whether government ownership affects the cost of firms' bond financing rather than bank lending. In contrast, this paper focuses explicitly on the relationship between government ownership of debtors and the pricing of their bank loans.

3 Data

Our paper uses novel register data on Finnish firms. There are four key data components. First three, firm population and background data, firm ownership data, and firm balance sheet data, come from the ready-made research datasets by Research Services of Statistics Finland. The fourth is the proprietary and confidential corporate credit registry of Finnish corporate bank loans, that is, the Finnish implementation of the Analytical Credit Database (AnaCredit).[4]

First, dataset FIRM BASE is used to define the total population of Finnish firms for year 2021 as well as to provide their background characteristics. We select firms whose business start date was on or before 2021 and that were still operating at year-end 2021. If a characteristic value of a firm changed between year-end 2021 and the data sourcing date (January 17, 2022), this change is reflected in the data used. However, since background characteristics are largely fixed and the time gap is very short, this is not considered a problem. We filter the sample by excluding *entrepreneurs* (see Appendix Subsection A.1), as the focus in this study is on firms owning other firms.[5]

Firms operating in the CRE market are of special interest in this work. We refer to them as *CRE firms*. These are firms that contribute to the same phenomenon of interest: commercial use or development of estates. We focus on firms since notable physical assets (i.e., estates) in Finland used for commercial purposes, such as offices or shopping malls, are typically encapsulated in dedicated company structures. We construct a definition of CRE firms that is practical to apply and is consistent with the CRE definition used by the European Systemic Risk Board (Dagrain et al., 2019). To this end, we distinguish four mutually exclusive CRE subgroups: (CRE-1) *Real estate activities excl. housing corporations*, (CRE-2) *Housing corporations excl. private and foreign housing companies*, (CRE-3) *Private and foreign housing companies*, and (CRE-4) *Construction firms*. Union of the four subgroups forms the definition of *CRE firms*.[6] We further define firm groups that are hypothesized to be important owners of CRE firms: i) *Government firms*, containing *Local government firms* and *Other government firms*; ii) *Holding companies*, containing *Financial holding companies* and *Head offices*; iii) *Financial firms*, containing *Funds*, *Banks*, *Insurance firms*, *Pension firms*, and *Other financial firms*. All the top-level groups are mutually exclusive. The group *Other firms* includes all remaining uncategorized firms. Finally, we also use an alternative grouping of firms based on whether they can be categorized among *Public & non-profit firms* ("P&NP firms" for short). This group includes firms that are either backed, at least implicitly, by taxpayer money or benefit from tax exemptions due to their social benefit role.[7] The group may overlap with the previously defined firm groups. Detailed definition of each firm group is found in Subsection B.4 of the Appendix.

[4]In the Appendix, Subsection A.1 provides a description of Finnish firm types and Section B short descriptions of the datasets.

[5]The definition of base population includes firms that are classified as *non-active*. Consequently, the sample may be somewhat larger than the set of going-concern firms.

[6]*Construction firms* are part of the definition used in this paper, whereas they are excluded in the ESRB definition.

[7]That is, "public" in this context does not refer to "publicly traded" firms.

Second, dataset FLOWN OWNER provides shareholder information of corporations[8] and FLOWN PARTNER provides partner information of business partnerships. The data are reported by corporations and business partnerships as part of their business taxation. The two datasets are merged together to construct a harmonized dataset as of year-end 2021 describing ownership relations between firms, i.e., owner-owned links. All links where the owner (either a shareholder or a partner) is a natural person are removed from the harmonized dataset, as are links where the owned entity is an *entrepreneur*.[9] Ownership shares are defined for shareholders as relative number of shares held, and for partners as relative shares out of partnership's wealth or income accredited to the partner. The harmonized ownership dataset will be used to construct the ownership network studied in this paper (see Section 4). Although the FLOWN datasets contain some information on foreign owners, the quality of these data is considered poor. Hence, we do not explicitly distinguish foreign owners.[10] Further, owner-owned links where the owned entity is either a *mutual real estate investment company* (subset of CRE-1), a *public housing company* (subset of CRE-2), or a *private or foreign housing company* (entire CRE-3) are missing because such firms do not report their owners to FLOWN data. Details about the harmonized ownership data can be found in Appendix Section B.3.

Third, the balance sheet size of firms is obtained from the FIRM FSS dataset and merged with the network and base population where needed. FIRM FSS covers a subset of firms from the base population. The balance sheet size is used to provide a "volume" weighting for individual firms or groups of firms. Some specific sectors are missing from the FIRM FSS data, with *public authority units* being the most important for our purposes (see Appendix Section B.1). This means that firms in the group *Government firms* do not acquire any balance sheet amounts.

Fourth, the Finnish implementation of AnaCredit is used to extract a sample of the bank loan stock of Finnish firms as of October 2021 for the regressions in Section 5.[11] The following filters are applied to construct the common sample underlying each individual regression sample. Debtors are limited to Finland-domiciled entities operating in institutional sectors S.11 *non-financial corporations*—excluding sector 1121 *housing companies*[12]—and to entities classified under any of the four CRE subgroups listed above. Rare cases of debtors with conflicting CRE status information between AnaCredit and FIRM BASE data are excluded. Similarly, rare cases of loans involving multiple debtors are removed from the

[8] A corporation must report all its shareholders if there are at maximum 10 shareholders. If there are more than 10 shareholders, it must report shareholders owning at least 10% of the shares as well as all shareholders who have been granted a shareholder loan.

[9] Excluding natural persons and *entrepreneurs* reduces size of the harmonized dataset significantly: the amount of firms drops by 88%, from about 576 thousand to 68 thousand. This is not surprising, as eventually all firms are owned by natural persons. In this work, we are not interested in this tautological fact, but rather on which type of firms own CRE firms.

[10] There may be a handful of foreign owners, who in the results are categorized either as *Other firms* or as missing.

[11] Appendix Subsection B.2 shortly describes the AnaCredit dataset. October was the latest available data point and is thus used instead of December 2021.

[12] As noted above, housing companies do not have identified owners and are therefore excluded from the analysis in Section 5.

sample. Loan instruments are euro-denominated, non-revolving credit, non-syndicated loans with the instrument type *Other loans* and a positive outstanding nominal amount. Additional sample selection criteria for individual regressions are presented in Subsection 5.1.

4 Ownership network

In this section, an ownership network along with ownership chains is derived. The network is used to address the question: *Who are the owners of CRE firms?* Subsection 4.1 details the methodology, Subsection 4.2 displays descriptive statistics of the network, and Subsection 4.3 provides results.

4.1 Methodology

Construction of the ownership network begins with the harmonized owner-owned dataset described in Section 3. We construct a directed network from the bilateral ownership links, where nodes represent firms and an edge from source node j to target node i indicates an ownership relation—node j owns (a part of) node i. Derived ownership shares are used as edge *weights* w_{ji}. The following conditions apply for the weights: $\sum_i w_{ji} \leq 1$, meaning each owned firm is owned at most 100%; max$(w_{ji}) \leq 1$, meaning a single ownership share is at most 100%; and min$(w_{ji}) > 0$, meaning all edges with a derived ownership share of zero have been removed from the network. See Panel (a) in Figure 1 for an illustration.

Our main goal is to identify the owners of CRE firms and quantify the extent of their ownership across these firms. In addition to direct ownership, we also aim to capture indirect ownership through chains of direct ownership links. An *ownership chain* is defined as a *path*[13] between an *ultimate owner* firm (a firm in the network that is not owned by any other firm, i.e., has zero in-degree) and an *ultimate owned* firm (a firm in the network that owns no other firm, i.e., has zero out-degree). Nodes with strictly positive in- and out-degree are called *intermediate* nodes. We are particularly interested in ultimate owners, as they may act as "backstops"—for example, a parent company of a conglomerate that owns multiple companies. All nodes in the network belong to at least one ownership chain and may belong to multiple chains.

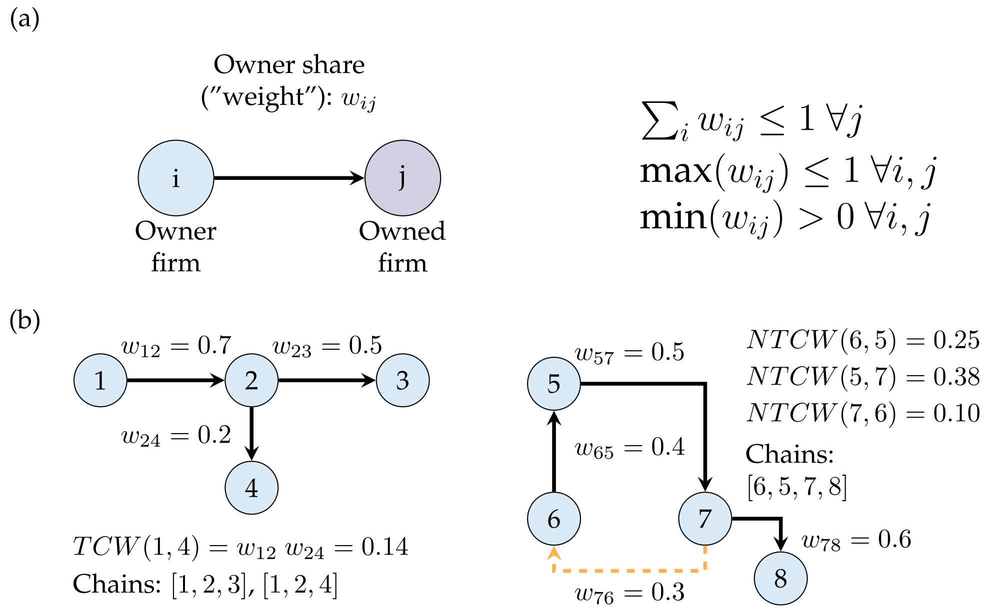
As an example, consider the left-hand plot in Panel (b) of Figure 1. There are three direct ownership links: node #1 owns node #2, node #2 owns node #3, and node #2 owns node #4.[14] Nodes #3 and #4 have a *direct owner* (node #2) but also an *indirect owner* (node #1). There are two ownership chains in this example: $[1, 2, 3]$ and $[1, 2, 4]$. Node #1 is an ultimate owner, and nodes #3 and #4 are ultimate owned nodes.

The original network contains *cycles*, i.e., trails in which only the first and last node are

[13] In graph theory, a *walk* is a sequence of edges that joins a sequence of nodes. A *trail* is a walk in which all edges are distinct. A *path* is a trail in which all nodes are distinct.

[14] Notice that none of the weights equals 100%, meaning that some ownership shares in the example are not captured by the network.

Figure 1: Ownership network illustration



Notes: (a) Schematic representation of firm i's ownership share (also referred to as *weight*) in firm j. (b) Illustrations of the derivation of ownership chains. The left-hand component contains no cycles, making the calculation of chains and cumulative ownership shares (*total cumulative weights*, *TCW*) straightforward. The right-hand component requires unwinding: the orange dashed edge is removed based on netted ownership shares (*net total cumulative weights*, *NTCW*), as explained in the text.

equal, making the determination of ownership chains nontrivial.[15] As an example, consider the right-hand plot in Panel (b) of Figure 1. Node #8 is clearly an ultimate owned node. Nodes #5, #6, and #7 can be considered to own node #8 either directly or indirectly, but which of these nodes is the ultimate owner? To obtain well-defined ownership chains, we first need to break the cycles and transform the original network into an "uncycled" version. Sun, Ajwani, Nicholson, Sala, and Parthasarathy (2017) present several approaches to this problem. We adopt a domain-specific algorithm, where the decision for the unwinding process is based on the nature of the problem.[16] Specifically, we break cycles by removing ownership links with the weakest ownership share.

Consider any two nodes j and i in the network that are connected by at least one path. For example, in the left-hand plot of Panel (b) in Figure 1, there exists one path between node #1 and #3: $[1, 2, 3]$. If an additional edge from node #1 to node #3 were present, there would be two paths: $[1, 3]$ and $[1, 2, 3]$. A *cumulative weight* along a directed path l from j to i is defined as $CW_l(j, i) = w_{jk_1} w_{k_1k_2} \dots w_{k_{N-1}k_N} w_{k_N i}$, where $k_1, \dots, k_N$ denote the nodes in the directed path between j and i. The *total cumulative weight* between nodes is the sum of cumulative weights over all directed paths from node j to i: $TCW(j, i) = \sum_l CW_l(j, i)$.

[15]Cycles in ownerships structures may be related to tax optimization or similar objectives.

[16]Other solutions discussed in Sun et al. (2017) are: i) simple heuristics based on breadth-first or depth-first searches; ii) the minimum feedback arc set; and iii) a hierarchy-preserving approach proposed by the authors themselves.

TCW thus represents the total (cumulative) ownership share that firm j holds in firm i. Next, the *net total cumulative weight* is defined as $NTCW(j,i) = TCW(j,i) - TCW(i,j)$.[17] When firms do not own each other (directly or indirectly), the NTCW for the owner firm coincides with TCW. NTCW becomes relevant when two firms own each other. A positive (negative) $NTCW(j,i)$ indicates that firm j owns cumulatively more (less) of firm i than vice versa. We unwind cycles by removing the direct edge with the smallest corresponding NTCW—i.e., edge (j,i) is removed if $NTCW(j,i)$ is the smallest among all node pairs connected by an edge in the cycle. In the right-hand example of Panel (b) in Figure 1, this means that among the three edges in the cycle, edge $(7,6)$ is removed. With the uncycled network at hand, finding ownership chains becomes straightforward. An algorithm for this procedure and details about the unwinding process are provided in Subsection C.1 of the Appendix. Further statistics on the unwinding process are reported in Section D of the Appendix.

By utilizing the identified ownership chains, we can determine the owners of a given target firm at different *owner levels*. This refers to the position of an owner node relative to a target node within an ownership chain. As a simple example, consider the left-hand structure in Panel (b) of Figure 1. Let node #3 be the target node. It appears at the end of the ownership chain $[1,2,3]$. Node #2 is a direct owner of node #3, as there is a directed edge between the nodes. It is also an intermediate owner, as it is itself owned by another node (node #1) and thus appears in the middle of the chain. Node #1 is an ultimate owner of node #3, as it appears at the start of the chain (in this example, node #1 is the only ultimate owner of node #3). Node #1 is also an ultimate owner of nodes #2 and #4. Finally, the owner levels of nodes #2 and #1 for target node #3 are 1 and 2, respectively.

We use the term *ownership unit mass* to denote the amount of ownership in a given firm (at most 100%). The corresponding weighted measure—ownership unit mass multiplied by the balance sheet size of the target firm—is referred to as *ownership volume mass*. The uncycled network, together with the calculated weights, can then be used to track how much of the target firm's mass each owner holds when moving up the ownership chains. We are interested in quantifying how much ownership mass is captured at different owner levels. Particularly interesting levels are ultimate owners and first-level (i.e., direct) owners. The ownership unit mass of an owned node i captured by an ultimate owner j is defined as $TCW(j,i)$. In contrast, the ownership unit mass of an owned node i captured by a first-level owner j is defined to be the direct ownership weight w_{ji}.[18] We distinguish four types of *ownership mass flows* between owned and owner nodes: A) mass of observations that are not in the network (and hence have no identified owner); B) mass of observations that are in the network but for which no owners are found; C) mass of target nodes lost due

[17] If there are no directed paths from j to i, then j is not an owner of i and $TCW = 0$. This does not preclude the existence of paths from i to j. If there are no directed paths between j and i in either direction, then $NTCW(j,i) = NTCW(i,j) = 0$.

[18] This is because, for first-level owners, we aim to capture only the truly direct exposures and ignore the indirect exposures resulting from additional, longer ownership chains between j and i. For example, consider a structure formed by the chains $[k,i]$ and $[k,l,i]$. In this case, the unit mass of i captured by k at the first-owner level comes solely via direct the edge $k \to i$.

to missing ownerships (arising from nodes for which at least one, but not all, owners are identified); and D) mass of target nodes that can be linked to owners.[19]

With the tools described above, we can conduct a detailed inspection of the owners of CRE firms—how many there are, who they are, and how concentrated the ownership structures are. Results from these inspections are presented in Section 4.3. In Figure 5, statistics are reported for three different sets of owners: ultimate owners, first-level owners, and *all owners*, which comprises unique owners from all owner levels; that is, for a given target node, it includes every node from which there is a directed path to the target node. The weight used to define acquired ownership mass for the all owners set is TCW.[20] Note that summing TCWs over all owners may overestimate the total identified ownership unit mass of the target nodes, as TCWs are "double-counted" for consecutive owners in an ownership chain.[21] Nevertheless, TCWs for the all owners set provide a quantifiable measure of the *adjacency* of owners to CRE firms—i.e., how "close" the owners in a group are to the firms.

4.2 Descriptive statistics

This section presents descriptive statistics on the ownership network and the total firm population. The uncycled ownership network comprises 66,141 firms (Table 1), representing 10% of all firms in the base population.[22] In other words, we identify 10% of Finnish firms as either owner entities (with at least one outgoing edge), owned entities (with at least one incoming edge), or both. At first glance, this may seem like a modest share, but the picture changes when we consider balance sheet sizes. As noted in Section 3, balance sheet data are available for a subset of 213,000 firms—about one-third of all firms. The aggregate balance sheet size, referred to as "volume," amounts to approximately EUR 1.8 trillion. Measure by this EUR volume, firms in the uncycled network account for 89% of all firms.

Table 2 and Figure 2 present descriptive statistics for the uncycled network. The network contains 61,389 edges, implying an average degree (the ratio of incoming and outgoing edges to the number of nodes) of nearly two.[23] The top-left subplot in Figure 2

[19]For an example of different ownership mass flows, consider the right-hand structure in Panel (b) of Figure 1. Node #1 has no identified owners, so its unit mass of 1.0 contributes to flow B. Node #3 has one direct owner, capturing unit mass of 0.5, which contributes to flow D. For node #4, 0.8 units of mass are lost due to missing ownerships, contributing to flow C.

[20]This implies that for the subset of first-level owners within the all owners set, the weights may differ from those used for the same owners in the first-level owners set.

[21]For example, consider the left-hand structure in Panel (b) of Figure 1. Let the target node be node #3. The all owners set of node #3 include nodes #1 and #2. Now $TCW(2,3) = 0.5$ (equal to the direct weight) and $TCW(1,3) = 0.35$. Summing these yields 0.85, which obviously exceeds the identified ownership mass for node #3 because node #2's share is partly "double-counted" in $TCW(1,3)$.

[22]For a valid comparison, network observations are counted only for firms that also appear in the base population. An additional 2,090 firms exist in the uncycled network but are not part of the base population. These firms typically appear as owners in the results of this subsection. However, analyses focusing on owned CRE firms include only firms present in both samples, as it is otherwise impossible to determine whether a firm is a CRE firm without base population attributes.

[23]The number of network nodes in 2 is slightly higher than in Table 1 due to the reason explained in Footnote 22.

Table 1: Firm counts and volume amounts

Panel 1

	Observations	Share of total
Network	66,141	10%
All firms	641,477	100%

Panel 2

	Obs. with volume	Volume amount (B)	Share, obs. with volume	Share, volume
Network	47,816	1,632	22%	89%
All firms	213,000	1,836	100%	100%

Notes: Panel 1 shows the number of firms in the uncycled network and in the base population. Network observation counts are based on firms that also appear in the base population. Additionally, there are 2,090 observations in the network that are not found among base population firms. Panel 2 shows the corresponding volumes measured by balance sheet size. The number of observations with available balance sheet information is reported separately. Sources: Statistic Finland's research data and author's calculations.

shows that the edge weight distribution is somewhat bimodal: weights most frequently cluster near 1.0, meaning that the most typical identified ownership share is (nearly) one. However, many edges also have weights close to zero. The top-right subplot indicates that target nodes (zero out-degree) are more common than source nodes (zero in-degree), accounting for 47,7% and 42,8% of all nodes, respectively, while approximately 9.5% of nodes act as intermediaries (non-zero in- and out-degree). The network is sparse: its density (the ratio of edges to the maximum possible number of edges) is virtually zero, and clustering coefficients are zero in the vast majority of cases (bottom-left subplot of Figure 2).

Degree assortativity values in Table 2 measure the similarity of connected nodes based on their degrees.[24] In a directed network, four types of degree assortativity can be distinguished: out-in, in-out, in-in, and out-out. For example, out-in degree assortativity indicates whether owners with high out-degree (many ownerships) tend to own firms with high in-degree (owned by many firms). In our network, the assortativity values are generally low, but nevertheless suggest a tiered structure. Owners with many ownerships tend to own targets that themselves have many ownerships, which is reflected in positive out-out assortativity, while they also tend to own targets that are not widely owned by others, resulting in negative out-in assortativity. Similarly, owners that are themselves widely owned tend to own targets with many ownerships, indicating positive in-out assortativity. Conversely, owners that are not widely owned tend to own targets that are owned by many, which corresponds to negative in-in assortativity.

The middle-row subplots in Figure 2 illustrate the distributions of node degrees and node strengths in the network. The numbers of incoming and outgoing edges per node are generally very low (middle-left plot), as indicated by 90th percentile being only 2 for both in-degree and out-degree. However, some nodes have multiple connections: the 99th

[24]Degree assortativities are Pearson correlation coefficients of degrees between pairs of nodes connected by a directed edge.

Table 2: Descriptive statistics of the uncycled network

# of nodes	# of edges	Density	Avg. degree	# of DC	DA out-in	DA in-out	DA in-in	DA out-out
68,231	61,389	0.0%	1.8	13,424	-2.8%	2.7%	-3.7%	3.1%

Notes: "DC" refers to disconnected components, and "DA" to degree assortativity (see text). Sources: Statistic Finland's research data and author's calculations.

percentile reaches 7 for in-degree and 8 for out-degree. The corresponding node strength distributions (sum of incoming and outgoing weights; middle-right plot) show that more than 30% of the in-strength distribution is concentrated at 1.0—the maximum possible value—indicating that for a relevant proportion of target nodes, we capture all of their owners.

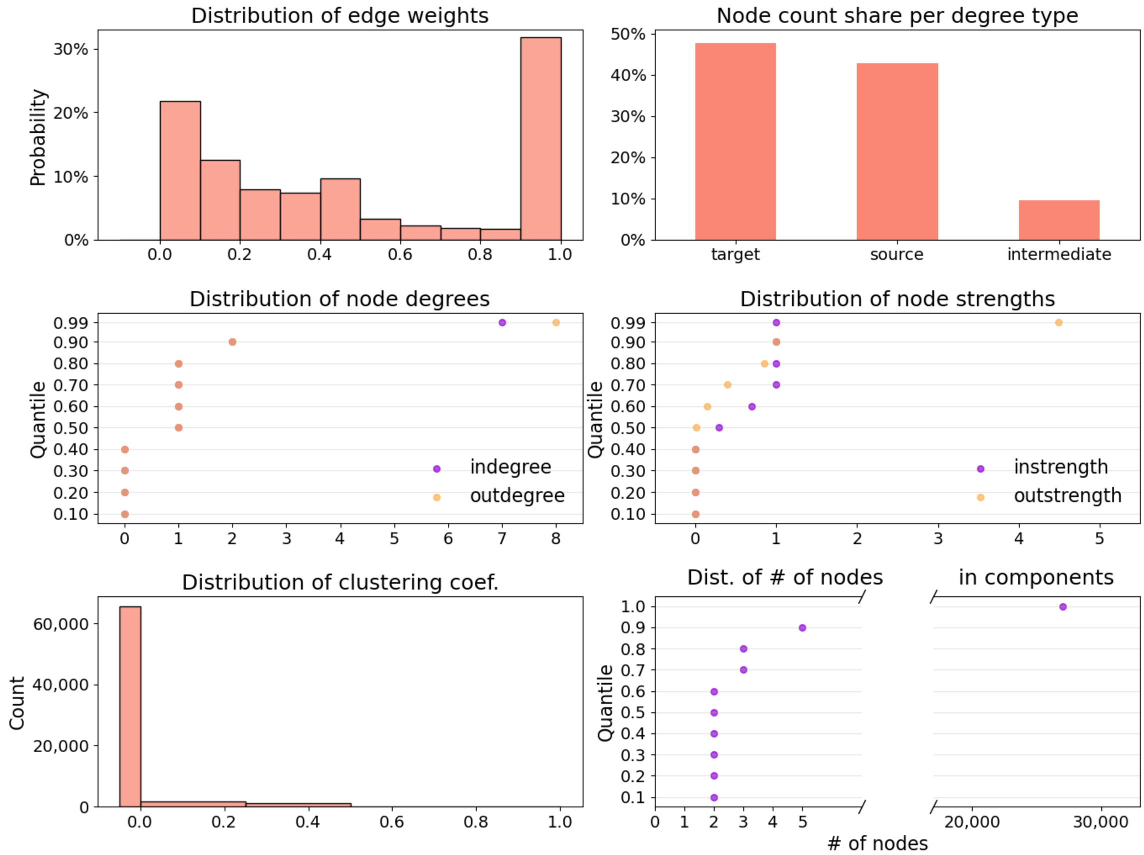
There are 13,424 disconnected components[25] in the uncycled network. On average, each component contains approximately five nodes. However, the distribution of component sizes is highly uneven. The most common motif is a two-node structure consisting of an owner node and a single owned node. As shown in the bottom-right subplot of Figure 2, more than half of the network's disconnected components follow this pattern. A particularly notable observation is the presence of a giant connected component comprising 26,981 nodes. In other words, 40% of all nodes in the network are concentrated in a single substructure. This finding underscores the value of a network-based approach for analyzing ownership relations.

To conclude this subsection, we examine statistics for the firm groups of interest defined in Section 3. Table 3 reports the number of observations and corresponding volume amounts (i.e., balance sheet sizes) for these groups and their subgroups found within the uncycled network. Panel 1 shows that there are 220,536 CRE firms in the base population, representing roughly one-third of all firms. Thus, in terms of counts, CRE firms constitute a major segment of the Finnish firm landscape, highlighting their importance. As expected, *Government firms* (Panel 2) and *Financial firms* (Panel 3) firms account for a much smaller share of the total firm population. When considering relative sizes in volume terms, however, *Financial firms* dominate, accounting for more than half of the total volume (EUR 1,046 billion out of EUR 1,836 billion). *Government firms* have no volume values because public authority units are not included in the FIRM FSS dataset.

Housing companies—a Scandinavian peculiarity (see Appendix Subsection A.1)—constitute the largest subgroup (CRE-3) of CRE firms in terms of observation counts. Their corresponding volume share is negligible because balance sheet data are mostly unavailable for housing companies. Furthermore, housing companies are largely absent from the ownership network for the reason noted in Section 3. *Government firms* and *Holding companies* are relatively well represented in the network, with ratios of network firms to base population firms of 53% and 57%, respectively. In terms of volume, *Holding companies* are

[25] Disconnected components are network structures in which no edges exist between two components, but within each component there is an undirected path between every node.

Figure 2: Descriptive distributions of the uncycled network



Notes: The top-left subplot shows the distribution of weights for direct edges. The top-right subplot displays the share of nodes by type: targets (zero out-degree), sources (zero in-degree), and intermediates (non-zero in- and out-degrees). The middle-left (middle-right) subplot presents quantiles of node in- and out-degrees (in- and out-strengths). Minimum and maximum values are omitted for confidentiality reasons. The bottom-left subplot shows the distribution of node correlation coefficients, while the bottom-right subplot presents quantiles of the number of nodes in disconnected components. Sources: Statistic Finland's research data and author's calculations.

well represented in the network (yet, it is to be noted that the availability of balance sheet data for such firms may be incomplete; see Appendix Section B.1). Table 4 examines the firm groups in more detail by tabulating the number of observations classified as owners, owned entities, or both. Importantly, for our focus on owners of CRE firms, Panel 1 shows that none of the none of the relatively few housing companies in the network are identified as owned entities. This reflects the data limitations discussed above. Interestingly, 2,161 housing companies are reported as owners of other firms.

Among other CRE subgroups, firms are more often identified as owned entities rather than owners, which is expected. Nevertheless, it is not uncommon for CRE firms to also act as owners. In contrast, *Government firms* and *Financial firms* are predominantly owners, again as anticipated.

Table 3: Amounts in firm groups, uncycled network vs. all firms

Panel 1

			Amount	Share of total	Vol. amount (B)	Share of total vol.	Amount obs. with vol.	Share of total obs. with vol.
CRE firms	**In network**	**Construction**	4,231	9%	30.5	73%	3,567	13%
		Hous. comp.	2,161	2%	0.0	23%	9	24%
		Other hous. corp.	1,963	5%	37.2	91%	1,447	27%
		RE activities	7,801	17%	48.8	64%	6,087	25%
		Total	16,156	7%	116.4	73%	11,110	19%
	All firms	**Construction**	47,961	100%	41.8	100%	28,119	100%
		Hous. comp.	90,548	100%	0.1	100%	37	100%
		Other hous. corp.	36,300	100%	40.8	100%	5,335	100%
		RE activities	45,727	100%	76.7	100%	24,836	100%
		Total	220,536	100%	159.4	100%	58,327	100%

Panel 2

			Amount	Share of total	Vol. amount (B)	Share of total vol.	Amount obs. with vol.	Share of total obs. with vol.
Government	**In network**	**Local government**	505	75%	-	-	-	-
		Other government	69	17%	-	-	-	-
		Total	574	53%	-	-	-	-
	All firms	**Local government**	674	100%	-	-	-	-
		Other government	416	100%	-	-	-	-
		Total	1,090	100%	-	-	-	-

Panel 3

			Amount	Share of total	Vol. amount (B)	Share of total vol.	Amount obs. with vol.	Share of total obs. with vol.
Financials	**In network**	**Funds**	580	30%	15.5	89%	462	72%
		Banks	139	46%	792.0	94%	130	60%
		Ins. & Pension	69	14%	90.9	95%	47	42%
		Other fin.	6,086	40%	69.4	76%	4,098	49%
		Total	6,874	38%	967.9	93%	4,737	51%
	All firms	**Funds**	1,921	100%	17.3	100%	642	100%
		Banks	299	100%	841.6	100%	217	100%
		Ins. & Pension	502	100%	95.7	100%	111	100%
		Other fin.	15,312	100%	91.1	100%	8,393	100%
		Total	18,034	100%	1,045.6	100%	9,363	100%

Panel 4

			Amount	Share of total	Vol. amount (B)	Share of total vol.	Amount obs. with vol.	Share of total obs. with vol.
Holding companies	**In network**	**Total**	1,033	57%	33.3	96%	594	69%
	All firms	**Total**	1,802	100%	34.6	100%	867	100%

Notes: The table shows the number of firm observations and the corresponding balance sheet amounts across different firm groups, both in the uncycled network and in the base population. Missing values indicate that there are too few underlying observations to publish. Sources: Statistic Finland's research data and author's calculations.

Table 4: Owner and owned amounts by firm groups in the uncycled network

Panel 1

		Amount			Vol. amount (B)			Amount obs. with vol.		
		Both	**Owned**	**Owner**	**Both**	**Owned**	**Owner**	**Both**	**Owned**	**Owner**
CRE firms	**Construction**	412	2,126	1,693	16.9	5.9	7.6	389	1,713	1,465
	Hous. comp.	0	0	2,161	0.0	0.0	0.0	0	0	9
	Other hous. corp.	208	1,234	521	30.2	5.9	1.1	181	919	347
	RE activities	559	5,401	1,841	15.5	15.6	17.6	465	4,075	1,547
	Total	1,179	8,761	6,216	62.6	27.5	26.3	1,035	6,707	3,368

Panel 2

		Amount			Vol. amount (B)			Amount obs. with vol.		
		Both	**Owned**	**Owner**	**Both**	**Owned**	**Owner**	**Both**	**Owned**	**Owner**
Government	**Local government**	30	52	423	-	-	-	-	-	-
	Other government	11	17	41	-	-	-	-	-	-
	Total	41	69	464	-	-	-	-	-	-

Panel 3

		Amount			Vol. amount (B)			Amount obs. with vol.		
		Both	**Owned**	**Owner**	**Both**	**Owned**	**Owner**	**Both**	**Owned**	**Owner**
Financials	**Banks, Ins., Pension**	22	17	169	206.0	67.3	609.7	21	15	141
	Funds	196	338	46	5.9	8.5	1.1	171	271	20
	Other fin.	974	1,311	3,801	27.1	21.7	20.5	715	779	2,604
	Total	1,192	1,666	4,016	239.1	97.5	631.3	907	1,065	2,765

Panel 4

		Amount			Vol. amount (B)			Amount obs. with vol.		
		Both	**Owned**	**Owner**	**Both**	**Owned**	**Owner**	**Both**	**Owned**	**Owner**
Hold. comp.	**Total**	212	199	622	20.6	1.2	11.5	132	97	365

Notes: The table shows the number of firm observations and the corresponding balance sheet amounts across different firm groups, categorized by whether a firm is an owner, owned, or both. All values are reported for both the uncycled network and the base population. *Banks*, *Insurance*, and *Pension* firms are combined into a single subgroup. *Government firms* have no volume values. Sources: Statistic Finland's research data and author's calculations.

4.3 Owners of CRE firms

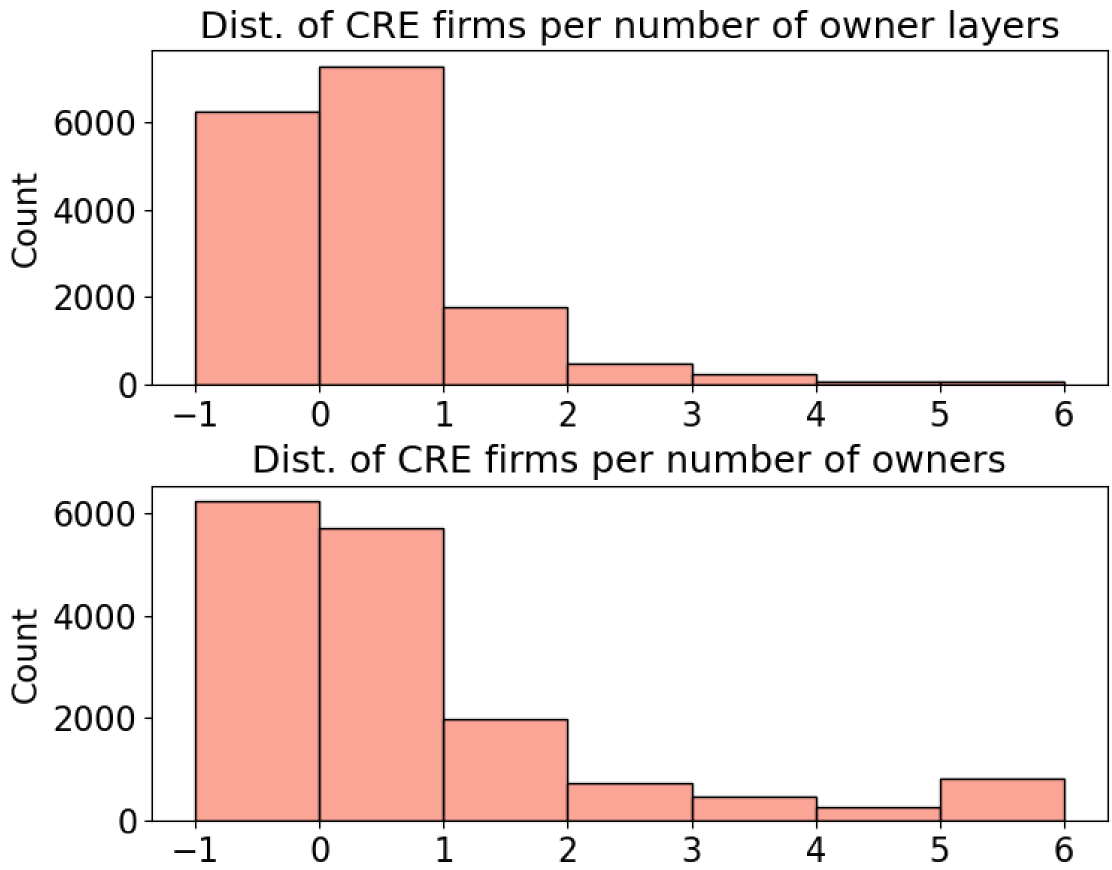
This subsection addresses the central theme of the paper: the owners of commercial real estate firms—how many there are, who they are, and the structures in which they appear. In what follows, we use the terms *owned CRE firm* and *target node* interchangeably. We emphasize volume-weighted results, as these highlight the significance of firms and their ownership linkages.

There are 16,156 target nodes in the uncycled network (Table 3). Figure 3 breaks down this number by the number of owners—whether direct, intermediate, or ultimate owner—associated with each target node, as well as the depth of the corresponding ownership chains. From the bottom subplot, we observe that 6,216 target nodes have no owner; in other words, they "own themselves" while acting as owners of other firms. The remaining 9,940 CRE firms have one or more owners (sum of bars excluding the leftmost bar). When at least one owner exists, the most common case is a single owner per target node, although multiple owners are not entirely uncommon.

The top subplot of Figure 3 shows that, in most cases, an ownership chain for a given CRE firm is only one owner level deep.[26] That is, as noted in Section 4.2, a large share of

[26] The maximum owner level is calculated from the shortest paths. Starting from a target node and moving

Figure 3: Breakdown of CRE firm counts by owners and owner levels



Notes: The top subplot shows the distribution of the number of CRE firms by maximum owner level (i.e., the maximum number of steps from the target, calculated via shortest paths). The bottom subplot shows the distribution of the number of CRE firms by the number of owners (intermediate or ultimate) they have. Bucket bins are right-inclusive. Sources: Statistic Finland's research data and author's calculations.

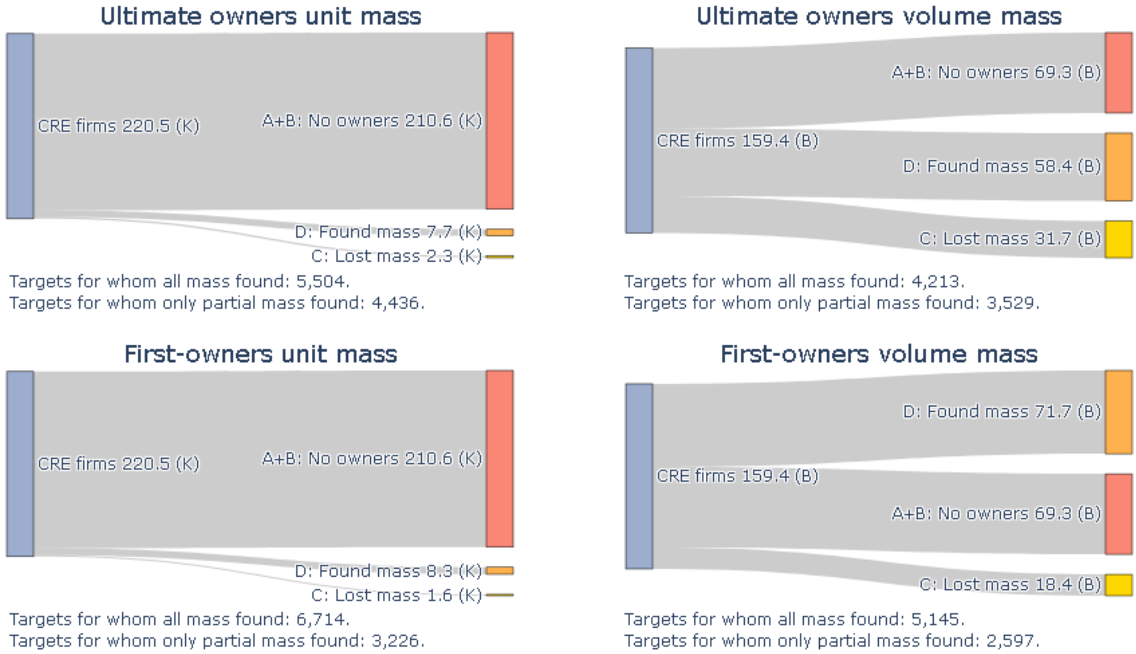
ownership links for target CRE firms consists of simple pairs of an owner and an owned firm. However, more complex structures also occur. Compared to the number of owner nodes (bottom subplot), the number of owner levels declines more rapidly, indicating that multiple owners are more often found at the same level rather than forming longer chains. Still, longer chains do appear, with 429 target nodes having owners from four or more levels.

Figure 4 illustrates how effectively the ownership network identifies owners of CRE firms by mapping ownership mass flows A-D, as described in Subsection 4.1, between owned CRE firms and their ultimate as well as first-level owners.[27] Inspection of ownership volume mass flows in the right-hand plots shows that, at the first-owner level, 45.0%

upstream in the network, if an owner node is reached via two paths of lengths 1 and 2, the maximum level is 1 (rather than 2, which would result if calculated from longest paths).

[27] By definition, more ownership unit mass can be linked to first-level owners than to ultimate owners; the farther we move from the target in the ownership chain, the more ownership mass is potentially lost. In the plots, flows A and B are aggregated for simplicity. Summing unit masses across the four flows returns the total number of CRE firms by definition, and similarly, summing ownership volume masses across the flows yields the total balance sheet amount of CRE firms.

Figure 4: Ownership mass flows in CRE firms



Notes: The figure illustrates ownership mass flows (A–D) from target CRE firms to owners. The top-left (bottom-left) plot shows the ownership unit mass of CRE firms linked to ultimate (first-level) owners. The top-right (bottom-right) plot presents ownership volume mass of CRE firms linked to ultimate (first-level) owners. Flows A and B are combined in the plots. Sources: Statistic Finland's research data and author's calculations.

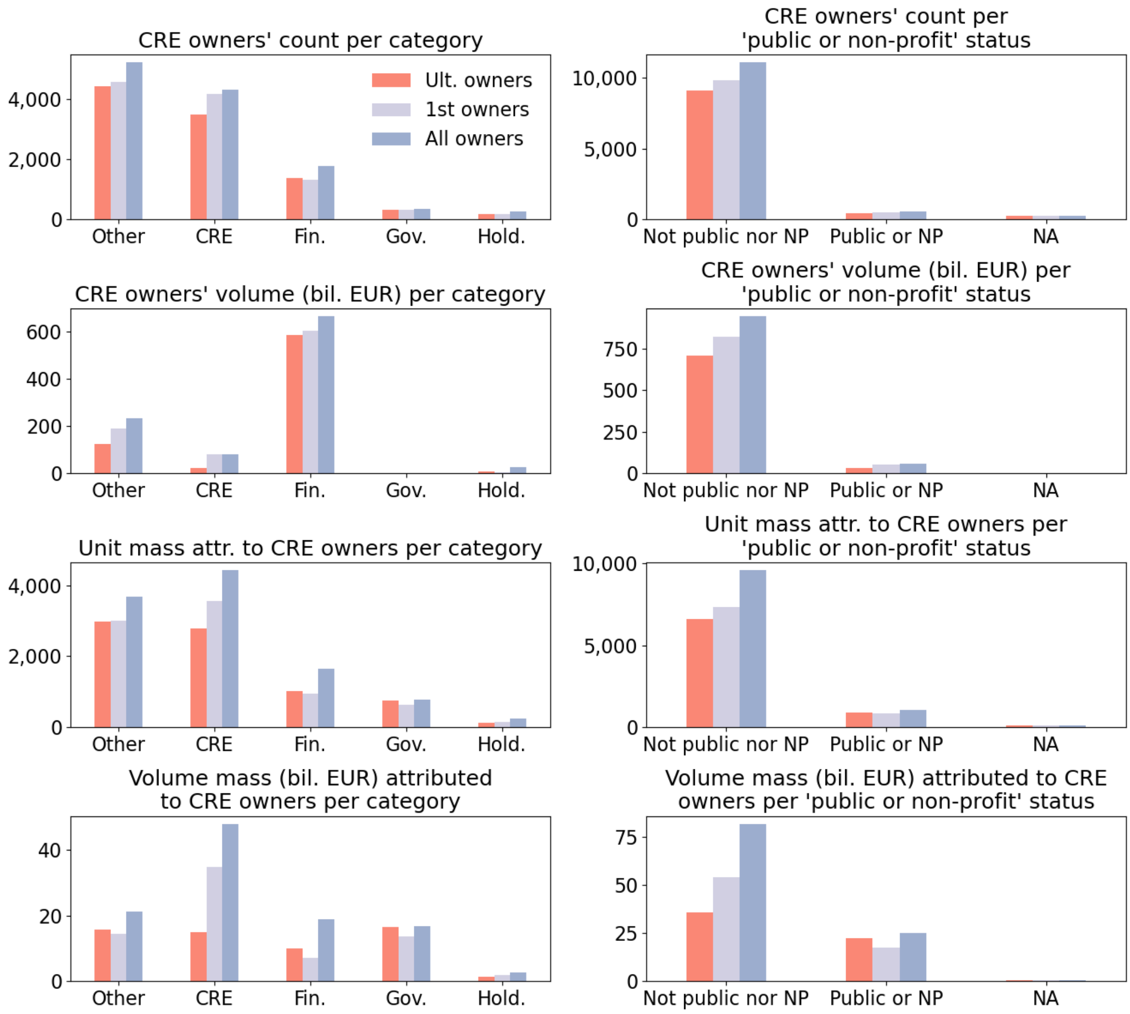
of the volume mass can be linked to owners. At the ultimate-owner level, the corresponding share is 36.7%.[28] In other words, when the volume size of target nodes is considered, the network maps a substantial share of CRE firms to owners: nearly half at the first-owner level and more than one-third at the ultimate-owner level. This indicates that a significant portion of the most important (i.e., larger) CRE firms have identifiable firm owners. In contrast, examining non-weighted unit masses in the left-hand plots shows that the network cannot identify owners for most CRE firms. This outcome is expected, as most of these firms are small and directly owned by one or more natural persons.[29] In short, the network captures a substantial share of owners of most important CRE firms, making the examination of the firm-level ownership network a worthwhile exercise.

The key result of this section, presented in Figure 5, examines who the owners of CRE firms are. Owners are categorized by firm groups (left-hand column) and by whether they belong to the group *Public & non-profit firms* (right-hand column). Statistics reported for the three owner sets (ultimate owners, first-level owners, and all owners) are each represented by separate bars. Focusing first on the four owner groups of interest (left-hand column),

[28]Note that the number of observations for which full or partial ownership volume mass is found decreases compared to the left-hand plots (unit masses). This is because not all firms have balance sheet information.

[29]Recall that all person-owners have been removed from the network, as this study focuses on firms owning other firms. Other reasons for missing ownership mass may include cases where a firm has a corporate owner but its ownership share is too small to be reported (see Footnote 8), or where ownership information is absent due to reporting deficiencies.

Figure 5: Who are CRE firm owners?



Notes: The figure shows statistics for owners of CRE firms, broken down by firm group (left-hand column)—*CRE firms*, *Financial firms*, *Government firms*, *Holding companies*, and *Other firms*—and by classification as a *Public & non-profit firm* (right-hand column). In the right-hand column, the group "NA" refers to owners not found in the base population, so no group can be derived. In the left-hand column, such observations are grouped under *Other firms*. Statistics are provided for three owner sets. The first two–ultimate and first-level owners—are the same as in Figure 4, while the third set includes owners from all levels. The first-row subplots show the number of owners. The second-row subplots report the corresponding total balance sheet amounts. The third and fourth rows display, respectively, how the ownership unit and volume masses (from Figure 4) of target CRE firms are distributed among owners. Target CRE firms that own themselves have been excluded. Sources: Statistic Finland's research data and author's calculations.

the fourth subplot displays the ownership volume mass from Figure 4 attributable to each group. At the first-owner level, nearly half (48.4%, or EUR 34.7 billion) of the identified volume-weighted ownership mass (EUR 71.7 billion) is attributable to other CRE firms. In other words, when a direct firm owner is identified for a target CRE firm, it is most often another CRE firm. The amounts attributable to other owner groups are smaller: EUR 13.7 billion (19.0%) for *Government firms*, EUR 7.2 billion (10.0%) for *Financial firms*, and EUR 1.9 billion (2.7%) for *Holding companies*, while uncategorized *Other firms* account for EUR 14.3 billion (19.9%). At the ultimate-owner level, EUR 42.7 billion (73.1%) of the identified ownership mass (EUR 58.4 billion) can be attributed to the four groups of interest. Notably, *Government firms* become the most significant owner group, with an attributable share of EUR 16.5 billion (28.2%). *CRE firms* account for EUR 14.9 billion (25.6%), *Financial firms* for EUR 10.0 billion (17.2%), and *Holding companies* for EUR 1.3 billion (2.2%). The remaining is EUR 15.7 billion (26.9%) is attributable to uncategorized *Other firms*.

Comparing these shares to the non-weighted case of unit masses (third subplot), the role of *Government firms* appears less pronounced relative to other owner groups. This indicates that government owners tend to hold larger CRE firms. The first subplot in the left-hand column simply counts the owners in each group rather than examining attributable ownership masses and reveals a picture broadly similar to that of unit masses. The second subplot shows, unsurprisingly, that in terms of balance sheet size, *Financial firms* are by far the largest owner group.[30] When inspecting the bars for all owners, other *CRE firms* emerge as the most adjacent to target CRE firms, as their attributable ownership masses are the highest in both weighted and unweighted cases (fourth and third subplot, respectively). In simple owner counts (first subplot), the uncategorized *CRE firms* group is slightly more prevalent. This suggests that ownership links tend to be stronger when CRE firms own each other.

Turning to the right-hand column, the fourth subplot shows that *Public & non-profit firms* account for 38.1% (EUR 22.2 billion) of the ownership volume mass at the ultimate-owner level, while at the first-owner level the share is 24.1% (EUR 17.3 billion). Aggregating non-profits with government owners thus identifies an even more prominent owner group, although their count is very small compared to other owners (first subplot). The second subplot, unsurprisingly, shows that because *Financial firms* are typically not categorized as public or non-profit, *Public & non-profit firms* constitute a relatively small group in volume terms as well.

We draw three main conclusions from Figure 5. First, the most important direct owners of CRE firms are other CRE firms: they account for nearly half of identified first-level owners, corresponding to 21.8% of the total CRE balance sheet amount (159.4 billion). When aggregating across all ownership levels, other CRE firms are also the most adjacent to target CRE firms. Second, *Government firms* constitute the single most important ultimate-owner group, with an attributed ownership share of 10.3% of the total CRE balance sheet amount. Third, when the group *Government firms* is expanded to include all *Public & non-*

[30] As noted in Section 3, *Government firms* do not obtain balance sheet amounts. The low volume of *Holding companies* reflects both their small numbers and the nature of their operations.

profit firms, the attributed ownership share at the ultimate-owner level increases to 14.0%.

5 CRE government ownership and loan pricing

This section investigates the association between CRE debtors' government ownership and bank loan rates. Subsection 5.1 details the research question and methodology. Subsection 5.2 displays descriptive statistics of the loan sample. Subsection 5.3 provides the results.

5.1 Methodology

We use the measures of CRE firm ownership from Section 4 to investigate whether government ownership of CRE firms is associated with the price these firms are charged for their bank loans. Intuitively, government ownership in a firm may be seen as a factor that increases its creditworthiness for the following reasons: the government may provide more support than other types of owners in times of distress; such firms may be considered politically and/or strategically important (increasing the willingness not to let them fail); or they may receive preferential treatment, such as regulatory advantages. Increased creditworthiness of a government-owned firm versus a non-government-owned firm would then materialize as lower bank loan spreads for the former.

A hypothesis we want to test is that *government-owned CRE debtors have lower observed bank loan rates*, or, in a "continuous" version, *the greater the extent of government ownership in CRE debtors, the lower the observed rates charged on their bank loans*. We focus on ownership at the ultimate-owner level, as *Government firms* are essentially always ultimate owners (see Figure 5). Using the October 2021 bank loan stock described in Section 3 as our data, we employ a regression framework of the following type:

$$R_{ldb} = \Phi_b + \beta_1 Z + \beta_2 C + \beta_3 X + \beta_4 GO + \varepsilon_{ldb} \tag{1}$$

where R_{ldb} is the loan spread of loan l of debtor d granted by bank group b, and ϵ_{ldb} is the error term of the regression. Standard errors are clustered at the *bank-group $\times$ debtor* level. Φ_b represents bank-group fixed effects, Z debtor-level features, C loan collateral, and X other loan characteristics. GO represents debtor's government ownership at the ultimate-owner level. Given our hypothesis, we expect β_4 in Equation (1) to be negative. GO may take two different forms, depending on whether we want to test the "discrete" or "continuous" version of the hypothesis. Let $\mathcal{G}$ represent the set of *Government firms*, $\mathcal{C}$ the set of *CRE debtor firms*, and UO_d the ultimate owners of CRE debtor $d \in \mathcal{C}$. For each d, we define the *degree of government ownership at the ultimate-owner level* as

$$DGO_d^{UO} = \begin{cases} 0 & \text{if } UO_d \text{ is empty} \\ \sum_{k \in UO_d} \mathbb{1}_{k \in \mathcal{G}} \, TCW(kd) & \text{otherwise,} \end{cases} \tag{2}$$

where $\mathbb{1}_{k \in \mathcal{G}}$ denotes an indicator function that equals 1 when owner k belongs to the set of *Government firms*. The term $TCW(kd)$ represents the total cumulative weight between owner k and debtor d, as defined in Subsection 4.1. Now, GO_d will have the form

$$GO_d = \begin{cases} \mathbb{1}_{DGO_d^{UO} > 0} & \text{"Discrete" version} \\ DGO_d^{UO} & \text{"Continuous" version.} \end{cases} \tag{3}$$

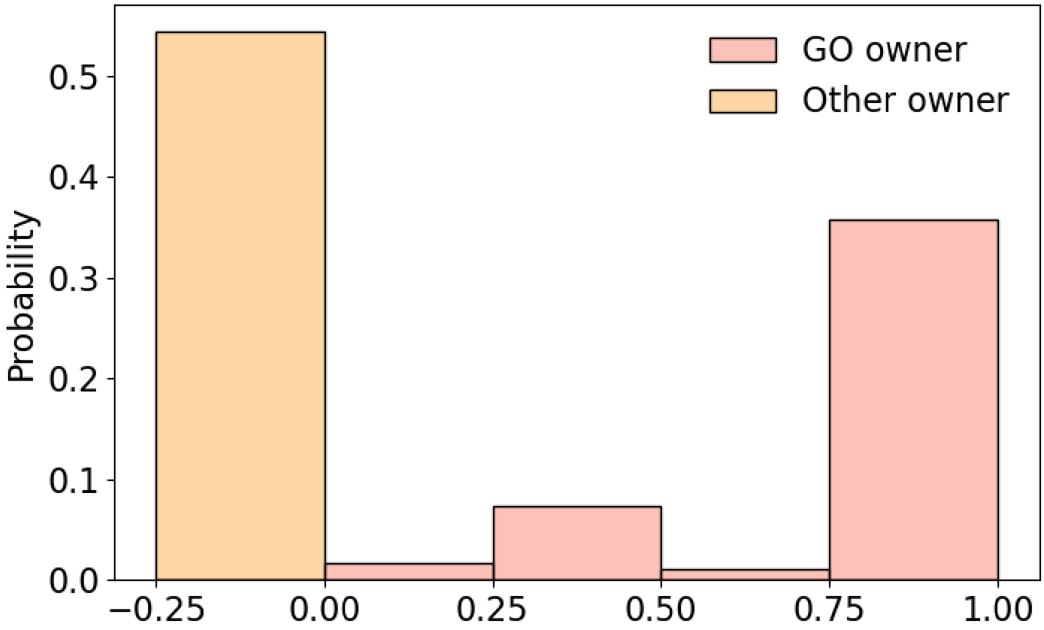
The "discrete" version indicates whether a debtor is government-owned (even to a small degree), and the "continuous" version specifies the share of government ownership.

Our aim is to construct Equation (1) so that it can quantify the association between government ownership of CRE debtors and their bank loan rates, while controlling for other relevant determinants of bank loan pricing. Risk premium (default risk of a borrower) and term premium (term structure of the risk-free rate) are the two main factors in pricing of corporate bonds. As noted by Beim (1996), corporate bank loan pricing is influenced by a much larger number of variables, and term structure, in fact, does not play a significant role. The author identifies particularly important determinants of bank loan pricing, including reference rate type (whether loan is priced against Prime Rate or market index), debtor's size and location, lender's identity, and lending relationship between the bank and the debtor. Motivated by this, we include all these variables as controls except for the lending relationship, for which we do not have a suitable proxy. The full set of control variables in our preferred model are as follows. Bank group fixed effects Φ_b are aimed to capture creditor-related idiosyncratic features, such as differences in risk appetite. Z includes debtor features that are thought to influence creditworthiness: debtor size (number of employees and balance sheet size), region, age, as well as CRE sub-category to distinguish different CRE firms from each other. C is the ratio of allocated collateral to the loan size. X includes reference rate type, loan age, maturity, and loan purpose. Formal definitions of control variables are given in Appendix Subsection B.5.

The baseline regressions employ five different models, each incorporating a distinct combination of control variables. Model *M1* excludes all control variables; *M2* includes Z; *M3* includes Z and X; *M4* includes Z and Φ_b; *M5*, our preferred model, includes Z, X, and Φ_b; finally *M6* the a full set of controls: Z, X, Φ_b, and C. The rationale for excluding C from the preferred specification is discussed in Section 5.3 alongside the presentation of the results. Building on the loan data described in Section 3, the baseline regression sample further excludes loans with fixed interest rates, as the loan spread may be undefined in such cases. Further, the baseline sample is constructed so that the comparison group consists of loans whose debtor has *some* owner, but not a government owner. This approach allows us to focus specifically on the influence of ownership by the government, rather than also capturing the question of having and owner versus not having one.[31] The models concentrate on estimating predictive associations rather than asserting definitive

[31] The inclusion of loans from debtors without an identified owner is covered in the robustness analysis.

Figure 6: Distribution of DGO_d^{UO} in the baseline model *M5* sample



Notes: Bins are right-inclusive. Bars sum to 1. Sources: Statistic Finland's research data, Bank of Finland (Anacredit), and author's calculations.

causal effects. Nevertheless, Appendix Subsection C.2 discusses whether and under what conditions our models could warrant a causal interpretation of the relationship between government ownership and loan spreads.

We assess the robustness of the baseline specification using several alternative specifications. These specifications employ the preferred model *M5*, except for minor deviations in individual cases. The first robustness specification, *R1*, considers only "new loans," defined as loans that are at most two years old. Since the owner network is constructed for the year 2021, this robustness check ensures that loans granted in earlier periods—possibly under different ownership structures or influenced by macroeconomic shocks that induce correlation among loans issued within the same period—are excluded. We exclude loan age from the set of control variables in this specification. *R2* excludes loans for which the debtor belongs to the subgroup CRE-2 (*Housing corporations excl. private and foreign housing companies*). This aims to minimize the potential biasing effect of the so-called "ARA loans" (interest-subsidized loans), although, as argued in Subsection A.2 of the Appendix, this should not pose a problem in the first place. *R3* outlines results from a weighed least-squares regression, with loan amounts as weights. *R4* uses the definition of the group *Public & non-profit firms* as a basis for the ownership variable instead of the group *Government firms*. This robustness specification examines whether the results remain consistent when the pool of government owners is extended to include non-profits. *R5* considers a scenario in which the comparison group also includes loans whose debtors have no identified owner. *R6* tests whether the results are robust to changing the response variable from loan spread to the contractual loan rate. Fixed-rate loans are included in this specification.

Finally, it may be the case that government-owned CRE debtors constitute a distinct

Table 5: Descriptive statistics of the baseline *M5* model sample

Government-owned, N = 4,512.

	Mean	**SD**	**25p**	**50p**	**75p**
Spread (bp)	73	49	44	75	95
Debtor age	28.7	14.5	17.0	30.0	39.0
# of employees	24	40	1	4	23
Debtor BS (mil. EUR)	501.1	1,078.2	10.2	41.8	215.2
P&NP debtor	0.85	0.35	1.00	1.00	1.00
Orig. maturity	25.6	10.4	19.6	25.0	32.9
Loan age	9.8	5.6	5.3	9.4	14.0
Loan size (mil. EUR)	1.99	2.75	0.21	0.80	2.46
CTL	0.90	0.36	1.00	1.00	1.00

Non-government-owned, N = 5,270.

	Mean	**SD**	**25p**	**50p**	**75p**
Spread (bp)	153	99	90	140	200
Debtor age	17.4	13.0	7.0	14.0	28.0
# of employees	26	50	0	1	23
Debtor BS (mil. EUR)	296.8	709.2	0.8	3.5	60.6
P&NP debtor	0.01	0.07	0.00	0.00	0.00
Orig. maturity	15.9	13.3	5.1	10.0	27.5
Loan age	5.0	4.6	1.5	3.6	7.0
Loan size (mil. EUR)	1.43	2.35	0.08	0.32	1.62
CTL	1.08	0.45	1.00	1.00	1.01

Notes: The table presents descriptive statistics for the sample underlying the baseline Model 5 regression. "SD" refers to the standard deviation. "25p", "50p", and "75p" denote the 25th, 50th, and 75th percentile values. Sources: Statistic Finland's research data, Bank of Finland (Anacredit), and author's calculations.

segment of the CRE market in the sense that they are firms with certain public-interest functions. Such firms may exhibit different risk profiles compared to other CRE firms, which in turn could bias the estimates of the association under study. Specifications *R7* and *R8* examine whether the *Public & non-profit* status of the debtors affects the results. In *R7*, we use the baseline sample but include the debtors' P&NP status as an additional control variable. In *R8*, we explicitly exclude all debtors with a positive P&NP status from the sample.

5.2 Descriptive statistics

The baseline regression sample described in the previous subsection has a total of 9,782 loan observations with an outstanding amount of 20.0 billion euros. Figure 6 displays the distribution of DGO_d^{UO} from Equation 2 in the baseline sample. We observe that in most cases where a government (ultimate) owner for a debtor is identified, the ownership degree is 1 or close to it, indicating a tight—possibly direct—ownership. However, more diluted government ownership degrees are also observed.

If we split the sample based on whether DGO^{UO} is greater than zero vs. equals zero, we obtain an almost equal division of observations. Descriptive statistics for the resulting groups—*government-owned* (with a government owner) and *non-government-owned* (with at least one non-government owner but no government owner)—are presented in Table 5. There are 4,512 observations in the government-owned group and 5,270 observations in the non-government-owned group. We notice that debtors in the government-owned group are, on average, older and larger in balance sheet size, yet slightly smaller in terms of number of people employed. Loans in the government-owned group have been outstanding for a longer period, have longer original maturities, and have larger outstanding amounts. Loans in both groups are highly collateralized, as the collateral-to-loan (CTL) ratios in both groups equal 1.0 in the 25th percentile. Interestingly, the average CTL is lower in the government-owned group than in the non-government-owned group (this observation is discussed in the next subsection). The average loan spread is noticeably lower in the government-owned group, suggesting that the hypotheses proposed in the previous subsection may hold true.

The dummy variable *P&NP debtor* indicates how common it is for CRE debtors themselves to be classified among *Public & non-profit firms*. The first panel shows that the majority (85%) of loans in the government-owned group are obtained by CRE debtors classified as P&NP firms. However, some CRE debtors are government-owned but not classified as P&NP firms (15% in terms of loan observations). Thus, being government-owned does not necessarily imply that the debtor is a public or non-profit entity. In contrast, the second panel shows that almost all loans in the non-government-owned group are issued to CRE debtors that are *not* classified as P&NP firms. In fact, this pattern arises because *all* debtors with P&NP status in the baseline sample have been classified as government-owned entities in the firm background characteristics data.[32] The fact that we observe a tiny number of P&NP cases also within the non-government-owned group most likely reflects minor inconsistencies between our derived ownership network and the firm background data.[33] We conclude that the P&NP status of the debtors—at least in our baseline regression sample—functions merely as an administrative label, arising solely from the fact that the debtors are government-owned.

5.3 Results

Tables 6 and 7 present results from the baseline regression specification with the discrete and continuous versions of government ownership variable, respectively. Focusing first on Table 6, in the model without controls (*M1*), the coefficient (-82 basis points) is, as expected, very close to the difference between the mean loan spreads from Table 5.[34] Adding

[32] Appendix Section B.4 describes in detail how the group *Public & non-profit firms* is constructed from the FIRM BASE data. It turns out that a positive P&NP status for each observation is always determined by the same criterion: the "is public corporation" (*Ohi*) dummy equals 1. This can occur in the following cases: (i) the debtor's deduced owner is either a municipality or the general government; (ii) the debtor's institutional sector is 13 (i.e., it is itself a government entity); or (iii) the debtor is classified as a public corporation for some other reason (via the variable *JulkisYhtTyyppi*). For the observations at hand, the first case is the sole reason behind the P&NP status.

[33] The ownership network is constructed from the FLOWN data, whereas the firm background information is drawn from the FIRM BASE data.

[34] The minor discrepancy arises from the slight difference in the number of observations between Models *M1* and *M5*.

Table 6: Baseline regression results for discrete GO

	IR Spread					
	M1	M2	M3	M4	M5	M6
GO	-81.580*** (5.839)	-56.848*** (4.337)	-35.824*** (3.376)	-46.897*** (3.194)	-33.342*** (3.017)	-32.059*** (3.069)
Debtor age		-0.608*** (0.141)	-0.453*** (0.101)	-0.530*** (0.100)	-0.415*** (0.091)	-0.408*** (0.089)
# of employees		-0.087 (0.090)	-0.063 (0.089)	-0.074 (0.081)	-0.076 (0.089)	-0.078 (0.089)
Debtor BS (mil. EUR)		-0.003 (0.005)	-0.001 (0.004)	-0.001 (0.005)	0.001 (0.004)	0.001 (0.004)
Orig. maturity			-1.452*** (0.144)		-0.909*** (0.136)	-0.898*** (0.135)
Loan age			-2.417*** (0.246)		-2.984*** (0.241)	-2.952*** (0.240)
Loan size (mil. EUR)			-3.201*** (0.472)		-2.991*** (0.443)	-3.029*** (0.443)
CTL						8.276*** (2.803)
CRE sub-group	No	Yes	Yes	Yes	Yes	Yes
Region	No	Yes	Yes	Yes	Yes	Yes
BG	No	No	No	Yes	Yes	Yes
Ref. rate type	No	No	Yes	No	Yes	Yes
Purpose	No	No	Yes	No	Yes	Yes
Observations	9,973	9,826	9,782	9,826	9,782	9,782
R2	0.200	0.315	0.439	0.511	0.573	0.574
Adj. R2	0.200	0.314	0.438	0.510	0.572	0.573

Notes: The table shows baseline regression results with a discrete choice for the GO variable. Statistical significances: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sources: Statistic Finland's research data, Bank of Finland (AnaCredit), and author's calculations.

control variables shrinks the estimated coefficient of interest; in the preferred model *M5*, the coefficient estimate is -33 basis points with a standard error of 3.02. Nevertheless, this result is both economically and statistically significant (at the 1% risk level). The explained variation, measured by the adjusted R^2, increases from 0.20 in *M1* to 0.57 in *M5*. Given the result in Figure 6, which shows that most government ownerships are concentrated around 1.0, it is not surprising that the results in Table 7 appear similar. Indeed, in the preferred model (*M5*), the estimated coefficient of interest for the continuous case is -36 basis points, with a standard error of 2.92 and overall adjusted R^2 of 0.57—very close to the discrete case.

Model *M6* is similar to *M5* but additionally includes CTL as a control variable. We exclude CTL from the preferred specification because it is considered a mediator, and controlling for it could introduce overcontrol bias. This reasoning is discussed in Appendix Section C.2. Interestingly, the coefficient of CTL in *M6* is positive—approximately 8 basis points—and statistically significant at the 1% level. A negative sign would be expected, as

Table 7: Baseline regression results for continuous GO

	IR Spread					
	M1	M2	M3	M4	M5	M6
GO	-84.753*** (5.646)	-62.560*** (4.062)	-40.231*** (3.218)	-50.736*** (3.073)	-36.466*** (2.918)	-35.141*** (2.941)
Debtor age		-0.533*** (0.138)	-0.397*** (0.100)	-0.476*** (0.097)	-0.370*** (0.089)	-0.364*** (0.088)
# of employees		-0.032 (0.088)	-0.029 (0.089)	-0.027 (0.081)	-0.043 (0.089)	-0.047 (0.089)
Debtor BS (mil. EUR)		-0.010** (0.004)	-0.006 (0.003)	-0.008* (0.004)	-0.004 (0.004)	-0.003 (0.003)
Orig. maturity			-1.458*** (0.139)		-0.933*** (0.133)	-0.920*** (0.133)
Loan age			-2.404*** (0.244)		-2.978*** (0.237)	-2.944*** (0.237)
Loan size (mil. EUR)			-3.150*** (0.485)		-2.961*** (0.451)	-2.999*** (0.451)
CTL						8.414*** (2.762)
CRE sub-group	No	Yes	Yes	Yes	Yes	Yes
Region	No	Yes	Yes	Yes	Yes	Yes
BG	No	No	No	Yes	Yes	Yes
Ref. rate type	No	No	Yes	No	Yes	Yes
Purpose	No	No	Yes	No	Yes	Yes
Observations	9,973	9,826	9,782	9,826	9,782	9,782
R2	0.189	0.318	0.442	0.512	0.574	0.575
Adj. R2	0.189	0.317	0.440	0.511	0.572	0.574

Notes: The table shows baseline regression results with a continuous choice for the GO variable. Statistical significances: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sources: Statistic Finland's research data, Bank of Finland (AnaCredit), and author's calculations.

greater pledged collateral should make the loan safer for the creditor, thereby reducing the required risk premium. One possible explanation is that our model does not account for all relevant determinants of creditworthiness.[35] Regardless of whether our reasoning for CTL holds, the coefficient for GO differs by only about one basis point between *M5* and *M6* in both tables—a negligible amount. Thus, in practice, the choice between models *M5* and *M6* has little impact, and we adopt *M5* as the preferred model.

Tables D.1 and D.2 in Appendix Section D provide results from the robustness specifications using the discrete and continuous versions of the variable of interest, respectively. As with the baseline results, the coefficients of interest across the tables are broadly simi-

[35]Explained with the help of Figure C.1 in Section C.2 of the Appendix: variable set O in the right-hand side subplot might not be empty. In this setting, CTL may partly capture that less creditworthy debtors are required to pledge more collateral, creating a positive association between CTL and the response variable. Further supporting this view, if the true data-generating process was accurately represented by the left-hand subplot of Figure C.1, controlling for CTL should make the estimate on GO *more* negative. This is because the blocked path $GO \rightarrow CW \rightarrow C \rightarrow R$ is expected to have a combined sign of $(+)(-)(-) = +$. However, this is not observed: the coefficient for GO in *M6* is *less* negative than in M5. This suggests that a more complex causal structure may be at play, like the one in the right-hand subplot.

lar, although the estimates in the continuous specification are slightly more negative than in the discrete case. We focus on discussing the results in Table D.1 and comparing them against the discrete baseline model *M5* in Table 6. First, in robustness specification *R1* (young loans), the estimated coefficient of interest is -41 basis points, with high statistical significance—that is, more negative than in the corresponding baseline case. This suggests that the hypothesized effect is stronger when focusing on newer loans, for which other confounding factors are arguably less problematic. Similarly, excluding loans in subgroup CRE-2 (*R2*) does not diminish the finding; instead, it makes it stronger (-41 bp). In the weighted least-squares specification (*R3*), the estimated coefficient of interest is the smallest in absolute value, at -26 basis points, and remains statistically significant. This suggests that when weighting by loan importance (size), the hypothesized association becomes somewhat smaller. *R4* shows that extending the ownership definition to include also non-profit owners (variable PNO in the tables) does not make a substantial difference (-32 bp). Similarly, extending the non-government-owned group to include loans without any identified owner for the debtor (*R5*) does not diminish the result; in fact, the estimate in this specification is -33 basis points, with a standard error of 3.21—a result very close to the baseline *M5* model. The estimated coefficient of interest becomes somewhat more negative (-39 bp) when the response variable is changed to the loan rate (instead of the loan spread) and fixed-rate loans are included in the sample (*R6*).

Specifications *R7* and *R8* show that the observed difference in loan spreads is not primarily driven by whether the debtors themselves are public-benefit firms. In *R7*, the coefficient on the additional variable *P&NP debtor* captures the incremental difference for P&NP debtors within the government-owned group. This estimated increment is small (-6 basis points) and statistically insignificant, whereas the main coefficient of interest remains strongly negative (–29 basis points) and statistically significant. A highly similar estimate for the main coefficient is obtained when we restrict the sample to debtors that are themselves not public or non-profit entities (*R8*). We conclude that government ownership is negatively associated with loan spreads even strictly within "normal" (non-public-benefit) debtors, and that the incremental effect attributable to public-benefit status is not material. We argue that this is because the P&NP status, at least in the baseline regression sample, serves more as an administrative label than as a factor influencing risk pricing—the substantive determinant remains government ownership.

All in all, the baseline results strongly support the hypotheses set out in Subsection 5.1: government ownership of a CRE debtor predicts lower loan rate spreads, as evidenced by the economically and statistically significant negative coefficient on the variable GO in both discrete and continuous cases and across all models. Further, the various robustness specifications clearly support the conclusion: the estimated coefficients of interest are again significantly negative—both economically and statistically—in every specification.

6 Conclusion and Discussion

Due to the opaque nature of the CRE market, obtaining a comprehensive overview is challenging. This paper contributes by examining the structure and concentration of owners of CRE firms in Finland and applies the findings to investigate determinants of CRE bank loan pricing. We show that, while other CRE firms are the most common direct owners, government entities are the most important ultimate owners, holding approximately 10% of the sector's balance sheet. Extending the government owners with non-profits creates an even more prominent owner group, with an ownership share of about 14%. Furthermore, we find that government ownership predicts lower interest rate spreads on CRE bank loans.

These results underscore the need to integrate ownership structures into financial stability assessments and credit risk models. Implicit public backstops or tax exemptions tied to social benefit roles may alter risk transmission in the CRE market, making it essential for regulators to reflect this information in their analyses. Further, implicit backstops may also influence credit pricing. Indeed, our pricing regression results are consistent with creditors perceiving CRE firms with a government owner as less risky. This poses an interesting follow-up question: can such expectations lead to credit mispricing due to possible misjudgments of the capacity or willingness of the state to intervene in case of distress? The question is important from a financial stability perspective as large shares of banks' balance sheets consist of commercial real estate loans.

Beyond financial stability, ownership mapping can inform investment strategies (firms primarily owned by the government tend to pursue low-risk, low-return growth strategies), mergers and acquisitions (partnership agreements and strategic alliances), and policy design by identifying entities affected by regulation. Our findings also underscore the need for more comprehensive financial reporting, particularly ownership data, in opaque markets such as CRE.

Future research could extend this analysis by combining the ownership network with banks' credit exposures to construct a multi-layered network. Such frameworks would enable contagion modeling and stress testing of CRE-related risks.[36] Another promising avenue is to examine the previously posed question of potential credit mispricing and its implications for financial stability.

[36]A somewhat similar analysis focusing on contagion in a two-layered ownership-supply network is conducted in Tabachová, Diem, Borsos, Burger, and Thurner (2024).

References

Acemoglu, D., Ozdaglar, A., & Tahbaz-Salehi, A. (2015, February). Systemic risk and stability in financial networks. *American Economic Review*, *105*(2), 564–608. Retrieved from https://www.aeaweb.org/articles?id=10.1257/aer.20130456 doi: 10.1257/aer.20130456

Allen, F., & Gale, D. (2000). Financial contagion. *Journal of Political Economy*, *108*(1), 1–33. Retrieved 2025-01-03, from http://www.jstor.org/stable/10.1086/262109

Antoniades, A. (2015, November). *Commercial bank failures during the great recession: The real (estate) story* (BIS Working Paper No. 530). Bank for International Settlements. Retrieved from https://ssrn.com/abstract=2697112

Ashraf, B. N. (2021). Is economic uncertainty a risk factor in bank loan pricing decisions? international evidence. *Risks*, *9*(5). Retrieved from https://www.mdpi.com/2227-9091/9/5/81 doi: 10.3390/risks9050081

Aslan, H., & Kumar, P. (2012, 04). Strategic ownership structure and the cost of debt. *The Review of Financial Studies*, *25*(7), 2257-2299. Retrieved from https://doi.org/10.1093/rfs/hhs062 doi: 10.1093/rfs/hhs062

Baqaee, D. R., & Farhi, E. (2024). Networks, barriers, and trade. *Econometrica*, *92*(2), 505-541. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA17513 doi: https://doi.org/10.3982/ECTA17513

Bardoscia, M., Barucca, P., Battiston, S., Caccioli, F., Cimini, G., Garlaschelli, D., ... Caldarelli, G. (2021). The physics of financial networks. *Nature Reviews Physics*, *3*(7), 490–507. Retrieved from https://doi.org/10.1038/s42254-021-00322-5 doi: 10.1038/s42254-021-00322-5

Battiston, S., Puliga, M., Kaushik, R., Tasca, P., & Caldarelli, G. (2012). Debtrank: Too central to fail? financial networks, the fed and systemic risk. *Scientific Reports*, *2*(541). Retrieved from https://doi.org/10.1038/srep00541 doi: 10.1038/srep00541

Beim, D. (1996). *The determinants of bank loan pricing* (Working Paper). Columbia Business School. Retrieved from https://business.columbia.edu/sites/default/files-efs/pubfiles/51/Determinants_of_Bank_Loan_Pricing.doc

Borisova, G., & Megginson, W. L. (2011, 03). Does government ownership affect the cost of debt? evidence from privatization. *The Review of Financial Studies*, *24*(8), 2693-2737. Retrieved from https://doi.org/10.1093/rfs/hhq154 doi: 10.1093/rfs/hhq154

Borsos, A., & Stancsics, M. (2020). *Unfolding the hidden structure of the hungarian multi-layer firm network* (MNB Occasional Papers No. 139). Budapest: Magyar Nemzeti Bank. Retrieved from https://hdl.handle.net/10419/241068

Boss, M., Elsinger, H., Summer, M., & Thurner, S. (2004). Network topology of the interbank market. *Quantitative Finance*, *4*(6), 677–684. Retrieved from https://www.tandfonline.com/doi/abs/10.1080/14697680400020325 doi: 10.1080/14697680400020325

Bramoullé, Y., Galeotti, A., & Rogers, B. W. (2016). *The oxford handbook of the economics of networks*. Oxford University Press. Retrieved from https://doi.org/10.1093/oxfordhb/9780199948277.001.0001 doi: 10.1093/oxfordhb/9780199948277.001.0001

Chaney, T. (2016, 04). 754networks in international trade. In *The oxford handbook of the economics of networks*. Oxford University Press. Retrieved from https://doi.org/10.1093/oxfordhb/9780199948277.013.19 doi: 10.1093/oxfordhb/9780199948277.013.19

Craig, B., & von Peter, G. (2014). Interbank tiering and money center banks. *Journal of Financial Intermediation*, *23*(3), 322-347. Retrieved from https://www.sciencedirect.com/science/article/pii/S1042957314000126 doi: https://doi.org/10.1016/j.jfi.2014.02.003

Dagrain, S., Dumont, L., Ferrière, T., Jimborean, R., Kennedy, G., Duca, M., ... Zollino, F. (2019, 12). *Methodologies for the assessment of real estate vulnerabilities and macroprudential policies: commercial real estate* (Tech. Rep.). European Systemic Risk Board. Retrieved from https://www.esrb.europa.eu/pub/pdf/reports/esrb.report191217_methodologies_assessment_real_estate_vulnerabilities_macroprudential_policies~15ff09ae41.en.pdf

Daly, P., Ryan, E., & Schwartz Blicke, O. (2024). Mapping the maze: a system-wide analysis of commercial real estate exposures and risks. *ECB Macroprudential Bulletin*(25).

Davis, E. P., & Zhu, H. (2011). Bank lending and commercial property cycles: Some cross-country evidence. *Journal of International Money and Finance*, *30*(1), 1-21. Retrieved from https://www.sciencedirect.com/science/article/pii/S0261560610000847 doi: https://doi.org/10.1016/j.jimonfin.2010.06.005

Davis, M. A., & Heathcote, J. (2005). Housing and the business cycle. *International Economic Review*, *46*(3), 751–784. Retrieved from http://www.jstor.org/stable/3663492

Duca, J. V., Muellbauer, J., & Murphy, A. (2021, September). What drives house price cycles? international experience and policy issues. *Journal of Economic Literature*, *59*(3), 773–864. Retrieved from https://www.aeaweb.org/articles?id=10.1257/jel.20201325 doi: 10.1257/jel.20201325

Engel, J., Nardo, M., & Rancan, M. (2021). Network analysis for economics and finance: An application to firm ownership. In S. Consoli, D. Reforgiato Recupero, & M. Saisana (Eds.), *Data science for economics and finance: Methodologies and applications* (pp. 331–355). Cham: Springer International Publishing. Retrieved from https://doi.org/10.1007/978-3-030-66891-4_14 doi: 10.1007/978-3-030-66891-4_14

European Central Bank. (2022). *Commercial real estate: connecting the dots.* https://www.bankingsupervision.europa.eu/press/supervisory-newsletters/newsletter/2022/html/ssm.nl220817.en.html. (Darrieux, Sébastien and Banu, Elena and McDonald, Marie-Therese and Castro Quintas, Claudio. Accessed 2025-07-14.)

European Systemic Risk Board. (2023). *Esrb issues a recommendation on vulnerabilities in the commercial real estate sector in the european economic area.* https://www.esrb.europa.eu/news/pr/date/2023/html/esrb.pr230125~f97abe5330.en.html. (Accessed 2025-07-14.)

Gambacorta, L., & Mistrulli, P. E. (2014). Bank heterogeneity and interest rate setting: What lessons have we learned since lehman brothers? *Journal of Money, Credit and Banking*, *46*(4), 753-778. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/jmcb.12124 doi: https://doi.org/10.1111/jmcb.12124

Graham, B. S. (2020). Chapter 2 - network data. In S. N. Durlauf, L. P. Hansen, J. J. Heckman, & R. L. Matzkin (Eds.), *Handbook of econometrics, volume 7a* (Vol. 7, p. 111-218). Elsevier. Retrieved from https://www.sciencedirect.com/science/article/pii/S1573441220300015 doi: https://doi.org/10.1016/bs.hoe.2020.05.001

Herring, R. J., & Wachter, S. M. (1999). *Real Estate Booms and Banking Busts: An International Perspective* (Tech. Rep. No. 99-27). The Wharton School Research Paper. Retrieved from https://ssrn.com/abstract=2546407

International Monetary Fund, Monetary and Financial Systems Dept. (2019). *Chapter 2: Downside risks to house prices* (Global Financial Stability Report No. April 2019). International Monetary Fund. Retrieved from https://doi.org/10.5089/9781498302104.082

Iori, G., De Masi, G., Precup, O. V., Gabbi, G., & Caldarelli, G. (2008). A network analysis of the italian overnight money market. *Journal of Economic Dynamics and Control*, *32*(1), 259-278. Retrieved from https://www.sciencedirect.com/science/article/pii/S0165188907000474 (Applications of statistical physics in economics and finance) doi: https://doi.org/10.1016/j.jedc.2007.01.032

Jaccard, I. (2021, April). *Leveraged property cycles* (ECB Working Paper No. 2021/2539). European Central Bank. Retrieved from https://ssrn.com/abstract=3832773 doi: 10.2139/ssrn.3832773

Jackson, M. O. (2016, 04). 71the past and future of network analysis in economics. In *The oxford handbook of the economics of networks.* Oxford University Press. Retrieved from https://doi.org/10.1093/oxfordhb/9780199948277.013.2 doi: 10.1093/oxfordhb/9780199948277.013.2

Kragh-Sørensen, K., & Solheim, H. (2014). *What do banks lose money on during crises?* (Staff Memo No. 3/2014). Oslo: Norges Bank. Retrieved from https://hdl.handle.net/10419/264958 (hdl:11250/2998632)

Leamer, E. E. (2007, September). *Housing is the business cycle* (Working Paper No. 13428). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w13428 doi: 10.3386/w13428

Mattes, J., Steffen, S., & Wahrenburg, M. (2010). Do information rents in loan spreads persist over the business cycles? *Journal of Financial Services Research*, *43*, 175-195. Retrieved from https://doi.org/10.1007/s10693-012-0129-z doi:

10.1007/s10693-012-0129-z

Santos, J. A. C. (2010, 11). Bank corporate loan pricing following the subprime crisis. *The Review of Financial Studies*, *24*(6), 1916-1943. Retrieved from https://doi.org/10.1093/rfs/hhq115 doi: 10.1093/rfs/hhq115

Santos, J. A. C., & Winton, A. (2019, 02). Bank capital, borrower power, and loan rates. *The Review of Financial Studies*, *32*(11), 4501-4541. Retrieved from https://doi.org/10.1093/rfs/hhz001 doi: 10.1093/rfs/hhz001

Sapienza, P. (2004). The effects of government ownership on bank lending. *Journal of Financial Economics*, *72*(2), 357-384. Retrieved from https://www.sciencedirect.com/science/article/pii/S0304405X03002459 doi: https://doi.org/10.1016/j.jfineco.2002.10.002

Shibut, L., & Singer, R. (2015, October). *Loss given default for commercial loans at failed banks* (FDIC Center for Financial Research Paper No. 2015-06). FDIC Center for Financial Research. Retrieved from https://ssrn.com/abstract=2710844 doi: 10.2139/ssrn.2710844

Soramäki, K., Bech, M. L., Arnold, J., Glass, R. J., & Beyeler, W. E. (2007). The topology of interbank payment flows. *Physica A: Statistical Mechanics and its Applications*, *379*(1), 317-333. Retrieved from https://www.sciencedirect.com/science/article/pii/S0378437106013124 doi: https://doi.org/10.1016/j.physa.2006.11.093

Sun, J., Ajwani, D., Nicholson, P. K., Sala, A., & Parthasarathy, S. (2017). Breaking cycles in noisy hierarchies. In *Proceedings of the 2017 acm on web science conference* (p. 151–160). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3091478.3091495 doi: 10.1145/3091478.3091495

Tabachová, Z., Diem, C., Borsos, A., Burger, C., & Thurner, S. (2024). Estimating the impact of supply chain network contagion on financial stability. *Journal of Financial Stability*, *75*, 101336. Retrieved from https://www.sciencedirect.com/science/article/pii/S1572308924001219 doi: https://doi.org/10.1016/j.jfs.2024.101336

Vitali, S., Glattfelder, J. B., & Battiston, S. (2011, 10). The network of global corporate control. *PLOS ONE*, *6*(10), 1-6. Retrieved from https://doi.org/10.1371/journal.pone.0025995 doi: 10.1371/journal.pone.0025995

Appendix to *Estate Owners' Ensemble — Mapping Commercial Real Estate Concentration using Finnish Firm Ownership Network*

Ville Voutilainen

Bank of Finland

ville.voutilainen@bof.fi

First version: July 30, 2025

This version: March 4, 2026

A Finnish firm landscape

A.1 Firm types

This subsection outlines the types of Finnish firms and defines the terminology used. The definitions are based on Finnish tax law (Ministry of Finance, 2002) and Statistics Finland[1], with some simplifications. An entity that engages in business activities and fulfills legal obligations is referred to as a *legal unit* (*oikeudellinen yksikkö*). Legal units are identified by a Business ID (*y-tunnus*). We focus on *partnerships* (*yhtymä*) and *coporations* (*yhteisö*), which we'll refer to informally as *firms*. In addition, there are *entrepeneurs* (*elinkeinonharjoittaja*), a category that includes *self-employed persons* (*ammatinharjoittaja*) and *business persons* (*liikkeenharjoittaja*)[2]. We do not include *entrepreneurs* in our analysis: although they are technically classified as "firms," for practical purposes they can be regarded as natural persons.

Under partnerships, there are *business partnerships* (*elinkeinoyhtymä*), which include *general partnerships* (*avoin yhtiö*) and *limited partnerships* (*kommandiittiyhtiö*).[3] These partnerships are intended for conducting business. For example, certain *investment funds* (*sijoitusrahastot*) adopt this structure: funds investing in real estate or private equity may be organized as limited partnerships (Finnish Tax Administration, 2025).[4] In addition, there are *deemed partnerships* (*verotusyhtymä*), which are

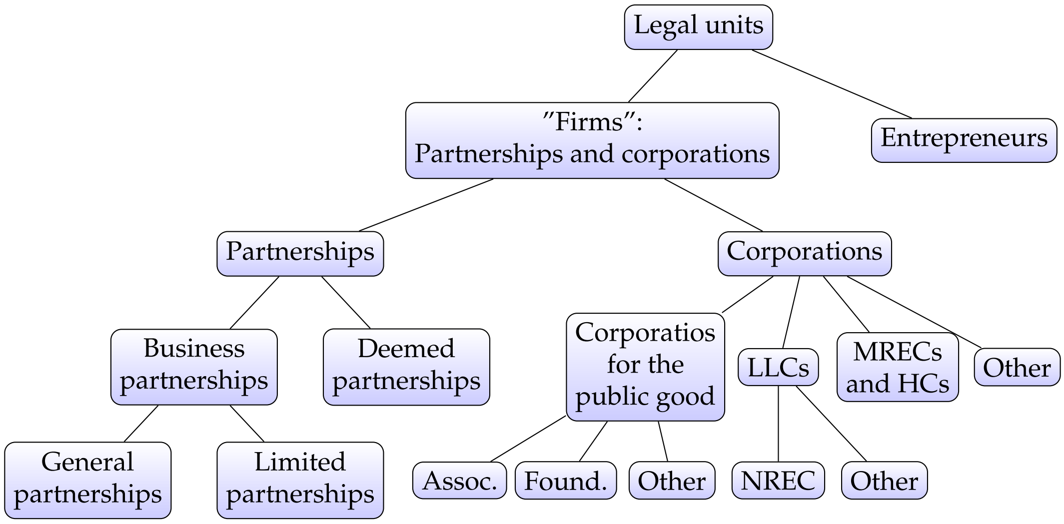
[1] `https://stat.fi/meta/kas/index_en.html`

[2] These are sometimes referred to as *sole proprietorships* (*toiminimi*).

[3] There are additional types of business partnerships, but these are excluded for the sake of simplicity.

[4] Investment funds may also adopt other legal unit types or structures. A general investment fund can be "administered by a fund management company, a limited liability company with which the unitholders in the investment fund enjoy a contractual relationship" (Finnish Tax Administration, 2025). These funds are not legal units themselves; they are managed by a separate limited liability fund management company. A real estate investment fund may also be organized as a real estate investment trust (REIT), which is a limited liability company.

Figure A.1: Finnish firm types



Notes: The figure has been adapted from its original source, which is from the perspective of taxation law. Some non-interesting categories have been dropped for clarity, and categories for *mutual real estate companies* (MREC) and *housing companies* (HC) as well as for *non-mutual real estate companies* (NREC) have been added. Sources: Ministry of Finance (2002), author's deduction.

intended for the control or cultivation of land and property.

Corporations are classified into *corporations for the public good* (*yleishyödyllien yhteisö*), *limited liability companies* (LLCs, *osakeyhtiö*), *mutual real estate companies* (MRECS; *keskinäinen kiinteistöosakeyhtiö*) and *housing companies* (HCs; *asunto-osakeyhtiö*), as well as other corporate forms[5]. Corporations for the public good include *associations* (*yhdistys*), *foundations* (*säätiö*) and other corporations for the public good, such as *municipalities* (*kunta*) and *parishes* (*seurakunta*). Limited liability companies represent what are typically considered "normal" firms, in which owners hold shares in the company. For the purposes of this paper, a particularly important category is *non-mutual real estate companies* (NREC, *ei-keskinäinen kiinteistöosakeyhtiö*). An NREC is typically used to hold or control immovable assets such as factories, offices, or other estates used for commercial purpose. As a limited liability company, its shareholders own shares in the entire company and have no direct control over individual properties or their parts (Laaka, 2023).

MRECs and housing companies are technically limited liability companies but are governed by specialized legislation. They differ from NRECs in one essential respect: in these companies, shares entitle the holder to direct control over a specific property or part of it, while the company remains the legal owner of the property. In contrast, shares in an NREC do not confer such direct control to individual owners. The key distinction between an MREC and a housing company is that, in the latter, more than half of the total floor area must be used as shareholder-occupied residential property (essentially apartments), whereas no such requirement applies to MRECs. Consequently, MRECs can be used for purposes such as commercial properties or warehouses (Laaka, 2023). In housing companies, shareholders are typically natural persons, making these entities essentially a

[5]Other corporations include, for example, *cooperatives* (*osuuskunta*), *credit unions* (*osuuspankki*), *savings banks* (*säästöpankki*), and *mutual insurance companies* (*keskinäinen vakuutusyhtiö*).

hybrid between commercial and residential real estate.

A.2 ARA loans

In Finland, the state implements a housing policy aimed at ensuring affordable and sustainable housing, particularly for low- and middle-income households as well as special groups. The state intervenes primarily through subsidies, guarantees, and interest-subsidized bank loans to support housing construction and renovation. The subsidized loans are typically called "ARA loans" due to the Finnish name of the Housing Finance and Development Centre of Finland (*Asumisen rahoitus- ja kehittämiskeskus*, ARA), which, until February 2025, was the government agency responsible of implementing state's housing policy.[6] These loans are of particular interest to us because of Section 5, where we regress the interest rate of bank loans on the government ownership of the debtor. To qualify for an interest-subsidized loan, a corporation must be classified as a "non-profit, public-benefit corporation" (*yleishyödyllinen yhteisö*), and government-ownership is one way to obtain this status. If government subsidization implied that the total contractual interest rate of a loan could not exceed a specific threshold, this might introduce bias into our regression results.

As stated on the Ministry of the Environment of Finland's website[7], the borrower's self-liability interest rate for an interest-subsidized loan is 2.3%. For the portion of the rate charged by the creditor that exceeds the self-liability rate, the borrower receives a state subsidy that decreases annually, with the first-year subsidy set at 90%. Hence, it appears unlikely that creditors would artificially cap the contractual interest rate at 2.3%. Furthermore, the Act on Interest Subsidy for Rental Housing Loans and Right of Occupancy Housing Loans section 6[8] states that the contractual rate may not exceed that of a comparable non-subsidized loan with a similar risk profile and characteristics. Moreover, creditors have no incentive to set a lower contractual rate than they otherwise would, as doing so would mean leaving money on the table. Hence, state interest-subsidized loans should have the same contractual rate as other loans, *ceteris paribus*. Consequently, such loans are unlikely to introduce bias into our results.

Non-profit, public-benefit corporation mostly fall under institutional sector 1122 *Other housing corporations* (Bank of Finland, 2023), which largely corresponds to our CRE-2 subgroup, *Housing corporations excl. private and foreign housing companies*. If ARA loans were to introduce bias, one of our robustness specifications excludes observations belonging to the CRE-2 group to ensure that any potential biasing influence is eliminated.

[6]Since then ARA's duties have been taken over by a newly established Centre for State-Subsidised Housing Construction (*Valtion tukeman asuntorakentamisen keskus*, Varke).

[7]`https://www.varke.fi/fi/yhteisot-ja-yhtiot/rahoitus-uudisrakentamiseen/korkotukilaina-asuntojen-hankintaan`

[8]`https://www.finlex.fi/fi/lainsaadanto/2001/604`

B Details on data

B.1 Ready-made research datasets of Statistics Finland

FIRM BASE

The dataset provides background information on legal units, including industry, sector, location, legal form, as well as business start and end dates. Each entity appears only once in the dataset, and the background data represent the latest known information. The data sourcing date is January 17, 2022.

FIRM FSS

The dataset contains income statement and balance sheet variables for Finnish firms at an annual frequency. Firms excluded from the data include *public authority units* (*julkisen sektorin viranomaisyksikkö*), *non-profits serving households* (Classification of Sectors 2012, code 15) as well as certain agricultural entities. Two caveats apply: 1) The following industries, defined using Standard Industrial Classification (TOL 2008) codes, may suffer from quality issues: 01-03 (Agriculture, forestry and fishing), 64-66 (Financial and insurance activities) and 701 (Activities of head offices); 2) Some information for firms in industry 68 (Real estate activities) has been imputed, as Statistics Finland may not directly obtain tax records for all firms in this industry.

FLOWN OWNER

The dataset provides information on the shareholders of a subset of corporations, based on tax filings (business tax forms 6B and 72 of the Finnish Tax Administration). Corporations reporting to the dataset include limited liability companies, cooperatives, non-mutual real estate companies, credit unions, savings banks, mutual insurance companies, and certain corporations for the public good[9]. The data are reported annually and include, *inter alia*, a pseudonymized ID for each shareholder (either a firm or a person), the number of shares owned by the shareholder, and the total number of shares in the owned entity. The dataset also contains dummy fields indicating whether the shareholder is a foreign firm or a foreign person; however, the quality of these fields is poor. A corporation must report all its shareholders if there are at most 10. If there are more than 10 shareholders, the corporation must report those owning at least 10% of the shares, as well as all shareholders who have been granted a shareholder loan.

FLOWN PARTNER

The dataset includes information on partners in business partnerships based on tax filings (business tax forms 6A and 72A of the Finnish Tax Administration). Deemed partnerships are not included. The dataset contains, *inter alia*, a dummy variable indicating whether the partnership is a

[9] *Insurance associations* (*vakuutusyhdistys*) and *pension foundations or funds* (*eläkesäätiö tai -kassa*).

Table B.1: Used Statistic Finland's datasets

Use	Dataset	Source	Description
Firm population	FIRM BASE	Statistics Finland (Documentation)	Background information on non-person legal units. Ready-made research data. ID: FIRM_BASE_jua_legalunits_001.xml.
Balance sheet info	FIRM FSS	Statistics Finland (Documentation)	Firm income statement information, yearly data 2021-2022. Ready-made research data. ID: FIRM_20132021_jua_FSS_001.xml.
Owner network	FLOWN OWNER	Statistics Finland (Documentation)	Information on shareholders of corporations. Ready-made research data. ID: FLOWN_20062021_jua_owner_001.xml.
	FLOWN PARTNER	Statistics Finland (Documentation)	Information on partners of business partnerships. Ready-made research data. ID: FLOWN_20072020_jua_partner_001.xml

general or limited partnership, a pseudonymized ID for each partner (eitehr a firm or a person), a dummy indicating whether the partner is a *silent partner* (*äänetön yhtiömies*) or a *general partner* (*vastuunalainen yhtiömies*), the country of any foreign partner, as well as partners' shares of different partnership income streams (general, agricultural, and forestry income) and wealth. The data are reported annually.

B.2 AnaCredit

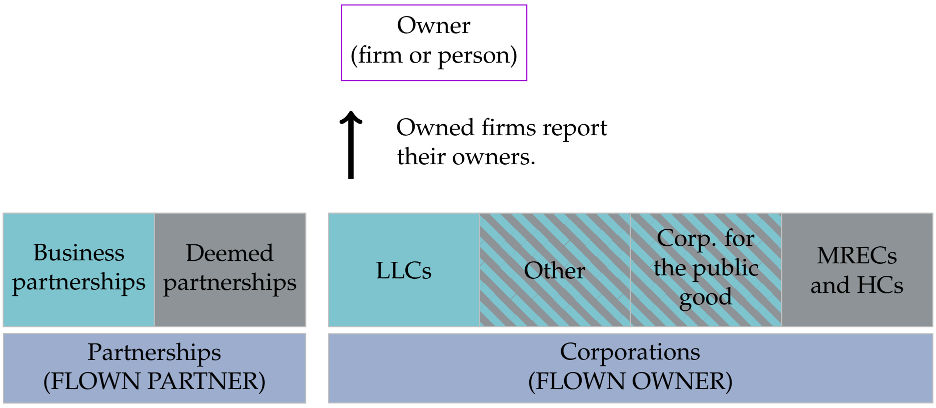
Analytical Credit Dataset (AnaCredit) is a proprietary and confidential dataset that provides detailed information on individual bank loans in the euro area[10]. In this paper, we use the Finnish implementation of AnaCredit, compiled by the Bank of Finland, which includes data on all corporate loans granted by Finnish credit institutions at a monthly frequency. These loan data have been merged with the ready-made research datasets of Statistics Finland in the FIONA remote access system (see Section E).

B.3 Harmonized ownership dataset

FLOWN PARTNER and FLOWN OWNER datasets form the foundation for the harmonized owner-owned dataset. Figure B.1 illustrates the data coverage by firm types discussed in Subsection A.1. As noted in Section B.1, corporations reporting to FLOWN OWNER do not represent the entire population of corporations. For example, most *corporations for the public good* report their owners in business tax form 6C, which is not covered by FLOWN OWNER. Furthermore—and importantly for this paper—MRECs and housing companies are subject to income taxation rather than business taxation; consequently, their owners are not included in FLOWN OWNER. Correspondingly, the striped blocks *Other* and *Corporations for the public good* in Figure B.1 indicate that not all corporations in these categories have reported their owners in the data. Firm types shown in grey blocks lack reported owners entirely. Although the FLOWN datasets include some information on foreign

[10] https://www.ecb.europa.eu/ecb-and-you/explainers/tell-me-more/html/anacredit.en.html

Figure B.1: FLOWN OWNER and FLOWN PARTNER ownership coverage



Notes: The blocks represent firm types from Figure A.1. The pastel green blocks denote entities that report their owners in FLOWN data, with the respective dataset names shown below. A striped block indicates that only a subset of entities within that type report their owners in the FLOWN OWNER data. The grey blocks denote entities that do no report their owners in the FLOWN data. Source: FLOWN data description, author's deduction.

owners, the quality of these data is considered poor. Therefore, we do not attempt to explicitly distinguish foreign owners. A handful of foreign owners may appear in the data; in the results, they are categorized either as either *Other firms* or missing.

FLOWN OWNER and PARTNER datasets are merged to create a harmonized owner-owned dataset describing ownership relations. When the same link appears in both FLOWN OWNER and FLOWN PARTNER[11], the former is preferred—except in a few cases where FLOWN OWNER is deemed unreliable, in which case information from FLOWN PARTNER is used. All owner-owned pairs where the owner is either a natural person or an *entrepreneur* are removed from the dataset, leaving only *owner firm-owned firm* links. In addition to ownership relations, the harmonized dataset includes information on the owner's share, which is derived differently depending on the data source. For links from FLOWN OWNER, the share is defined as the number of shares owned divided by the total number of shares.[12] For links from FLOWN PARTNER, the derivation is more complex, as the "owner share" of a partner in a partnership is not explicitly specified. We define owner shares based on the share of the partnership's wealth or income attributed to a given partner. These calculations also depend on the type of partnership (general or limited) and the type of partner (liable or silent).

Owner shares for partners are derived only for liable partners. Silent partners in general partnerships are treated as data errors—since, by definition, general partnerships should not to have silent partners—and are marked as liable partners. Silent partners in limited partnerships are assigned a zero owner share, as they are considered investors without control over partnership's activities. For liable partners, there are several alternative methods for defining the owner share, using either wealth, income, or agricultural income share to which the partner is entitled. We prioritize an owner share definition that satisfies certain sanity conditions as well as minimizes the

[11] Strictly speaking, this should not occur; however, small misclassifications in the data may exist.
[12] Some cleaning is applied to the denominator to ensure validity of the owner share.

number of zero shares. To achieve this, we employ a measure that we call *coherence of shares*, following a waterfall-like algorithm to select which calculated share measure is used as the owner share. *Strictly coherent shares* are defined as calculated shares for partners of a partnerships such that: i) not all shares are missing; ii) shares sum to 1, iii) the minimum share is greater than 0, and; iv) the maximum share is less than or equal to 1. *Coherent shares* are defined similarly, except that condition iii) is altered so that the minimum share equals 0. *Incoherent shares* are shares that do not meet the coherence criteria. The algorithm for deriving the owner shares of partners is:

1. First, give preference to the wealth share. For partnerships where wealth provides strictly coherent shares, use these as the owner shares. Next, among partnerships where wealth provides coherent (but not strictly coherent) shares, select those where income provides strictly coherent shares and use them. For partnerships where both wealth and income provide coherent (but not strictly coherent) shares, select those where agricultural income provides strictly coherent shares and use them. If wealth, income, and agricultural income all provide coherent (but not strictly coherent) shares, use wealth shares as the owner shares. For the remaining partnerships—those considered to have incoherent wealth shares—proceed to step 2.

2. Next, give preference to the income share. For partnerships where income provides strictly coherent shares, use these as the owner shares. Among partnerships where income provides coherent (but not strictly coherent) shares, select those where agricultural income provides strictly coherent shares and use them. For partnerships where both income and agricultural income provide coherent (but not strictly coherent) shares, use income shares as the owner shares. For the remaining partnerships—those considered to have incoherent wealth and income shares—proceed to step 3.

3. Next, give preference to the agricultural share. For partnerships where agricultural income provides strictly coherent shares, use these as the owner shares. For the remaining partnerships—those considered to have either coherent (but not strictly coherent) or incoherent agricultural shares—proceed to step 4.

4. For the remaining partnerships, calculate the owner share by dividing it equally among all partners.

Lastly, owner-owned links with a zero owner share (including those previously set to zero) are removed from the harmonized dataset.

B.4 Definitions of groups of interest

This subsection explains how the firm subgroups of interest are identified from the data.

CRE FIRMS

The top part of Figure B.2 illustrates the definition of CRE firms used in this paper. The three blocks in the top row together represent the ideal definition of a CRE firm: companies engaged in real estate activities, construction firms, and funds specializing in real estate investments. However, with the available data, it is not possible to distinguish *Real estate funds* from other types of funds. Therefore, the corresponding block is shown in grey in Figure B.2, indicating that although real estate funds would fall within the desired definition of CRE firms, they are excluded due to data limitations. These funds are instead classified under *Financial firms* (see below).

We identify four mutually exclusive CRE firm blocks from the base population, which together form our *operationalizable* definition of CRE firms.The first three blocks result from dividing real estate activities into three categories: (CRE-1) *Real estate activities excluding housing corporations*: EA 68 and IS 111.[13] This group captures a wide range of corporations operating in the real estate business, including NRECs and MRECs; (CRE-2) *Housing corporations excluding private and foreign housing companies*: EA 68 and IS 1122 or 11211; (CRE-3) *Private and foreign housing companies*: EA 68 and IS 11212 or 11213.[14] The fourth group (CRE-4) is *Construction firms*: EA 41-43 and IS 11.

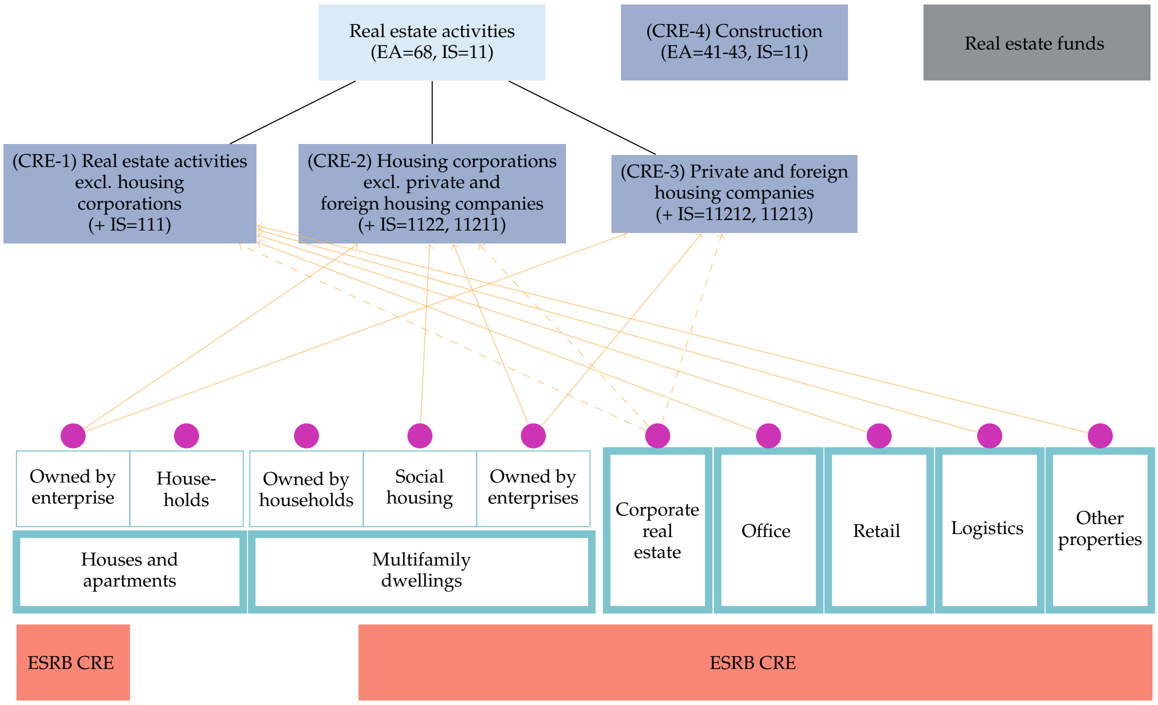
The arrows in the middle of Figure B.2 map different estate types, as reported in Dagrain et al. (2019), to our four CRE blocks. The mapping is not one-to-one; for example, the estate type *Houses and apartments owned by enterprise* maps to both CRE-2 and CRE-3, indicating that such estates may appear under both blocks. The bottom part of Figure B.2, also adapted from Dagrain et al. (2019), outlines the European Systemic Risk Board's definition of which estates and properties are classified as commercial real estate. In this definition, only *Houses and apartments owned by households* and *Multifamily dwellings owned by households* are classified as "residential" rather than "commercial" real estate, making the ESRB definition quite broad.

As shown, commercial real estate according to the ESRB's definition is largely captured by our operationalizable CRE firm definition, although aligning specific estate types with CRE blocks is not always straightforward. The main differences are: i) corporate real estates—properties owned by end users—may not always be captured by our CRE definition. Accordingly, the corresponding arrows in Figure B.2 are drawn with a dashed stroke. This reflects the fact that corporate real estate in Finland may not always be embedded within a firm structure (such as an NREC, MREC,

[13] Economic Activity (EA) codes refer to Standard Industrial Classification TOL 2008 and Industrial Sector (IS) codes refer to Classification of Sectors 2012 by Statistics Finland. When we write, for example, "IS 12", this refers to any IS code for which the two first characters are "12". IS codes used are: 11 non-financial corporations; 111 non-financial corporations, excl. housing corporations; 11211 housing companies, public; 11212 housing companies, national private; 1122 other housing corporations; 11213 housing companies, foreign controlled; 12 financial and insurance corporations; 121 central bank; 122 other monetary financial institutions; 1241 investment funds; 128 insurance corporations; 13 general government; 1313 local government; 13141 employment pension schemes. EA codes used are: 41 construction of buildings; 42 civil engineering; 43 specialized construction activities; 64 financial service activities except, insurance and pension funding; 641 monetary intermediation; 642 activities of holding companies; 643 trusts, funds, and similar financial entities; 65 insurance, reinsurance and pension funding, except compulsory social security; 651 insurance; 652 reinsurance; 653 pension funding; 66 activities auxiliary to financial services and insurance activities; 68 real estate activities; 701 activities of head offices.

[14] As explained in detail in Section B.3, the harmonized ownership dataset lacks owner information for housing companies. However, we can identify housing companies in the base population and therefore include them for completeness in the operationalizable definition of CRE firms.

Figure B.2: Commercial real estate firm definition



Notes: The three blocks on the top row together represent the desired definition of commercial real estate firms. The data does not allow us to disentangle real estate funds from other funds; therefore, they are excluded from the operationalizable definition of CRE (represented by a grey block). Real estate activities are divided into three subcategories. Blue blocks denote the CRE subgroups that are distinguished in the results, and together they form the operationalizable CRE definition. The middle row displays a breakdown of different estate types, which are mapped to the corresponding CRE blocks by orange arrows. The bottom row, shown with peach blocks, illustrates the ESRB 2019 definition of CRE. Source: Dagrain et al. (2019), author's deduction.

or housing company).[15]. In such cases, it will not be included in our operationalizable CRE firm definition; ii) *Construction firms* are included in the definition used in this paper, whereas they are excluded from the ESRB definition. We include construction firms because we consider them to contribute to the phenomenon of interest: the commercial use and/or development of estates.

OTHER GROUPS

Additional mutually exclusive firm subgroups of interest are defined as follows. *Government firms* include firms with IS code 13, excluding those code 13141. These can be divided into the

[15]Some practical examples may help illustrate different cases. As a first example, consider a barber shop operated by a *corporation*. The shop occupies premises encapsulated in an NREC, which the barber shop owns (fully or partially). In this case, the NREC is included in our CRE definition (CRE-1), and we should be able to identify the barber shop as the owner of the NREC in the ownership network. As a second example, consider the same scenario but where the barber shop is operated by an *entrepreneur*. The NREC remains within our CRE definition, but the ownership link does not appear in the ownership network because *entrepreneur* owners are excluded (see Section B.3). As a third example, suppose the barber shop is again operated by a *corporation* but occupies premises within a private housing company—i.e., a dwelling or other unit whose control is designated to the *corporation*. In this case, the housing company falls under CRE-3. However, the corresponding owner-owned link is missing from the ownership network due to data limitations regarding housing companies and MRECS (explained in Section B.3). Even without this limitation, the barber shop might not exceed the ownership reporting threshold if the housing company has more than 10 owners. As a fourth example, consider a retail firm that owns a refrigerated premise. Depending on the premise, it might be encapsulated in an NREC or be directly owned as a physical asset by the retailer (e.g., if it's a small-scale unit). In the latter case, the premise is not a "firm" and therefore it is not captured by our CRE definition.

following subgroups: *Local government* (IS 1313) and *Other government* (*Government firms* excluding *Local government*). *Holding companies* comprise two subgroups: *Financial holding companies* (EA 642 and IS not 13) and *Head offices* (EA 701 and IS not 13).[16] *Financial firms* are defined as a combination (logical OR) of the following conditions: IS 12; IS 13141; EA 64-66 and IS not 13. Additionally, any firm classified under *Holding companies* is excluded. Subgroups of *Financial firms* are defined as follows: *Funds*: EA 643 and IS not 13, or IS 1241; *Banks*: EA 641 and IS not 13, or IS 121-122; *Insurance firms*: EA 651-652 or IS 128, but excluding i) combination IS 128 & EA 653 to avoid overlap with *Pension firms*, and ii) EA 643 to avoid overlap with *Funds*; *Pension firms*: EA 653 and IS 13141; *Other financial firms*: *Financial firms* left over from the other subgroups. Finally, the group *Other firms* comprises all firms not included in any of the previously defined groups.

In some results, we examine firms based on an alternative classification: whether they can be categorized as "public and/or non-profit" firms. Note that "public" here refers to general or local government firms, not publicly traded companies. The group is similar to *Government firms* but extends the definition to include non-profit organizations—entities operating for public or social benefit (e.g., social housing providers)—which are hypothesized to be important players in the Finnish CRE market.[17] More formally, a firm is classified among *Public & non-profit firms* if any of the following criteria are met (see documentation of the dataset FIRM BASE in Table B.1): i) IS 13; ii) legal form (*VerohOikMuoto*) indicates a government agency or a religious community (codes 40-46, 48, 60-63); iii) any of the following dummies equals True: "is joint municipal authority" (*Kuntayhtyma*), "is public corporation" (*Ohi*), or "is non-profit institution" (*YleisYhteiso*); iv) has a public corporation classification (*JulkisYhtTyyppi*).

B.5 Definitions of regression variables

This subsection provides formal definitions of the variables used in the regression analysis of Section 5. Variables derived from Statistic Finland's ready-made research data (government ownership, CRE subgroup) are defined in the main text. Debtor balance sheet values are also based on Statfin's data. All other variables originate from AnaCredit. In what follows, "UTW99" denotes "upper-tail winsorized at the 99th percentile".

Response variables: *interest rate spread* is the margin added to the reference rate (UTW99); *interest rate* is the contractual total rate of the loan (typically reference rate plus spread, UTW99). Debtor-level control variables: *region* is the NUTS2-level region code; *age* is the debtor's age in years (UTW99); *number of employees* is the number of employees (UTW99); *balance sheet size* is expressed million euros (UTW99). Loan-level control variables: *original maturity* is the loan's original maturity in years (UTW99); *loan age* is the age of the loan in years (UTW99); *loan size* is the outstanding nominal amount of the loan in million euros (UTW99); *reference rate type* is a categorical variable indicating the type of reference rate (e.g., EURIBOR, index or derivative, etc.); *purpose* is a

[16] That is, general government entities are excluded to avoid overlap with *Government firms*.
[17] The group may partly overlap with the firm groups defined above.

categorical variable indicating the reported purpose of the loan (e.g., RRE purchase, CRE purchase, construction investment, etc.). *CTL* is the amount of allocated protection for the loan divided by its outstanding nominal amount, winsorized at 200%. Finally, *bank group* is a categorical variable for the creditor's bank group.

C Details on methodologies

C.1 Details on the unwinding process and ownership chains

This subsection describes the details of the network unwinding procedure and the identification of ownership chains (or simply *chains*). The process consists of three steps: first, the calculation of net total cumulative weights (NTCWs); second, the detection and breaking of cycles using NTCWs, followed by the creation of an uncycled version of the network; third, the identification of chains from the uncycled network.

Step 1 first calculates the net total cumulative weights (NTCW)—and, as an intermediate results, the cumulative weights (CW) and total cumulative weights (TCW)—for each pair of nodes (j, i), as defined in Section 4.1 of the main text. An outline of the algorithm is provided in Algorithm 1.[18] If the input network contains N nodes, the algorithm outputs an NxN matrix providing NTCWs. If there are no directed paths between j and i (in either direction), then $NTCW(j, i) = NTCW(i, j) = 0$.

Algorithm 1: Calculation of net total cumulative weights

```
Data: Original ownership network
Result: NxN matrix of NTCWs
for node j do
    Find the set L of directed paths to other nodes;
    for path l in L do
        Get weight for each edge in l;
        For each node i ≠ j in l, calculate CW_l(j, i);
    end
    Find set D of unique nodes over paths in L;
    for node i ≠ j in D do
        Calculate TCW(j, i);
    end
end
for node j do
    for node i ≠ j do
        Calculate NTCW(j, i);
    end
end
```

In Step 2, nodes forming cycles are first identified. To this end, *networkx* package provides a ready-made function, simple_cycles. Once the cycles are found, they are broken by removing the

[18] The presented algorithm is intended to illustrate the concept and is not optimized for performance. The actual implementation for finding all directed paths relies on modified versions of the *networkx* functions all_simple_edge_paths and all_simple_paths.

direct edge with the smallest corresponding NTCW, as explained in the main text. Two difficulties arise. First, multiple cycles—possibly of different lengths—may be intertwined, meaning share a set of common nodes. In such cases, breaking one cycle may affect another, and it is not straightforward to determine which cycle to break first. To resolve this, we adopt a pragmatic solution: cycles are sorted into a queue in ascending order by length, and shorter cycles are broken first. After breaking a cycle by removing the appropriate edge, we check against the remaining cycles in the queue to see if removal of the current edge already resolved them. If so, those cycles are removed from the queue. Thus, the queue ordering may affect the final result, as a different edge might have been removed had we started with, say, longer cycles. Second, ties may occur for the lowest NTCW within a cycle. Here, we again use a practical solution: edges in a cycle are sorted lexicographically by their ID (formed by concatenation of pseudonymized node IDs), and the first edge in this order is removed. There is no theoretical justification for these choices. However, as shown in Table C.1, the number of cycles is quite small, and ties are rare. Therefore, these pragmatic solutions are considered sufficient for our purpose.

Finally, in Step 3, chains are identified as described in Algorithm 2.[19] The output is a list of unique node chains, extending from ultimate source nodes (nodes with zero in-degree) to ultimate target nodes (nodes with zero out-degree).

Algorithm 2: Identification of chains

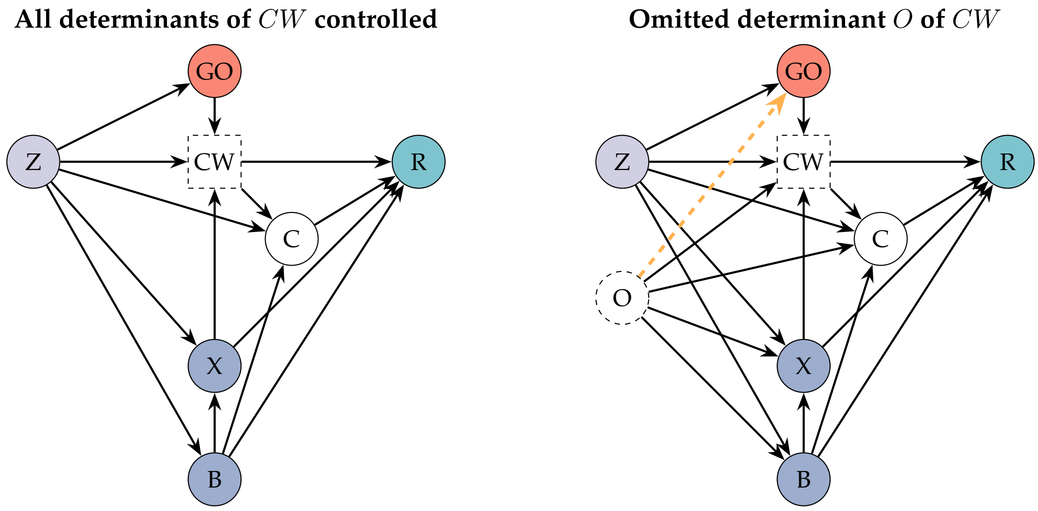
Data: Uncycled network G and its counterpart with reversed edge directions, G_r
Result: List of chains
for *node* j **do**
  In G, find the set L of directed paths to other nodes;
  In G_r, find the set L_r of directed paths to other nodes and reverse the order of nodes in each path;
  Take the Cartesian product of L and L_r to create chains and store them.
end
Among the stored chains, remove any chain that is a sub-chain of another chain.

C.2 Causal interpretation of regression results

Under what conditions could the predicted effect of government ownership on CRE bank loan spreads (Section 5) be interpreted causally? The main argument is that government ownership is mediated via higher creditworthiness to lower loan spreads. If this effect can be isolated from confounding factors, the effect may be interpreted as causal rather that merely predictive. Figure C.1 illustrates two directed acyclic graphs (DAGs) representing the posited causal relationships.[20] Variables Z, C, X, B, GO, and R correspond to those discussed alongside Equation (1) in the main text. CW is a latent variable representing the creditworthiness of a given loan-debtor pair—that is, how creditworthy a loan, with given characteristics and a debtor, is considered to be. Unlike other variables in the DAG, CW is treated as a deterministic "pass-through" variable without its

[19]The actual implementation relies on the same solution as described in Footnote 18.
[20]Fore more on DAGs, see Pearl (2009).

Figure C.1: Directed Acyclic Graph of posited causal relationships in the regressions



own (implicit) error term. The figure presents two versions: in the left-hand DAG, we assume Z and GO to fully capture all determinants of the latent creditworthiness variable—that is, CW has no variation beyond that induced by Z, GO, and X. Conversely, in the right-hand DAG, we allow for omitted determinants of creditworthiness not captured by our controls. These omitted factors are represented by O.

The reasoning behind the posited causal relations (edges in the DAG) is as follows. $GO \rightarrow CW \rightarrow R$ represents the main argument outlined above. Relations $Z \rightarrow C,\ X,\ CW$ are straightforward: debtor characteristics influence the pledged collateral, other loan characteristics, and creditworthiness. $Z \rightarrow GO$ indicates selection bias—different types of firms (debtors) are more likely to attract government ownership. Similarly, $Z \rightarrow B$ reflects that debtor characteristics affect which bank provides them lending, depending on banks' risk appetite and business models. Creditworthiness affects the amount of required collateral ($CW \rightarrow C$): lower creditworthiness translates into more collateral being required. $B \rightarrow X, R, C$ indicates that that banks offer varying loan terms and rates, as well as require varying amounts of collateral—depending on their risk appetite and lending policies. Loan characteristics may influence both creditworthiness ($X \rightarrow CW$) and loan rates directly ($X \rightarrow R$); for example, loan maturity may affect pricing, even if the term structure is not a major factor in bank loan pricing. It could be perhaps argued that a sufficient amount of pledged collateral, given creditworthiness, is a necessary condition for granting a loan, and beyond that, additional collateral does not affect pricing. However, with $C \rightarrow R$ we allow for the possibility that pledged collateral amount directly influences loan rates. Finally, if O is not empty, we assume it consists mainly of debtor characteristics that determine CW. Thus, O is allowed to have the same effects as Z.

Notably, the following relations are excluded from the DAG. First, there is no direct effect $GO \rightarrow R$. Loans to government-owned debtors could, in principle, be priced differently due to policy mandates, strategic campaigns, or other preferences favoring such loans. However, we expect this channel to be of lesser importance. Even if included this arrow, it would not affect the estimated total effect of GO on R (though it might weaken the main argument about *why* the effect exists). Second, bank-specific features are assumed not to directly influence the degree of a firm's government ownership (no arrow $B \to GO$). Third, there is no arrow $GO \to B$. If government-owned entities borrowed exclusively from specific bank groups, this effect would need to be considered. However, we do not view this as a particularly relevant channel in the present context—such considerations would matter if we were studying loans to government debtors *per se* rather than loans to debtors owned by government entities. Finally, there is no arrow $GO \to C$. Whether a debtor is government-owned is not expected to directly affect the required collateral; rather, both collateral and government ownership contribute to creditworthiness.

If the true causal structure corresponds to the left-hand DAG in Figure C.1, controlling for Z alone suffices to identify the total causal effect of GO (in this causal interpretation, the so-called *treatment variable*, colored peach) on the response variable R (colored green). In other words, the minimal adjustment set contains only Z, colored purple. Importantly, we do not *not* control for C, as doing so would block the path $GO \to CW \to C \to R$ and bias the effect of interest. Thus, our preferred model excludes C as a control. However, we include X and B (colored blue) in our preferred model. The reasoning is two-fold. First, although not required for identification in this case, controlling for these may improve model accuracy by reducing variation in the response variable, making them ”neutral controls” *à la* Cinelli, Forney, and Pearl (2024). Second, if our model suffers from omitted variable bias in the sense that Z, GO and X do not capture all determinants of CW (i.e., O is not empty), then controlling for both and X and B helps mitigate this bias by closing multiple backdoor paths of type $GO \leftarrow O \to X \to \cdots \to R$ and $GO \leftarrow O \to B \to \cdots \to R$. However, as long as O affects GO (arrow $O \to GO$, colored orange), the effect of GO on R remains biased due to backdoor paths $GO \leftarrow O \to CW \to R$, $GO \leftarrow O \to CW \to C \to R$, and $GO \leftarrow O \to C \to R$. Controlling for C in this scenario is a double-edged sword: while it would block two of these backdoor paths, it would also block the previously mentioned causal path $GO \to CW \to C \to R$. Besides the preferred model, the main text also reports results from a model where C is included as a control. If the omitted determinants of creditworthiness do not influence GO—that is, they are debtor characteristics unrelated to government ownership, so that

Table C.1: Descriptive statistics for the unwinding process

Panel 1

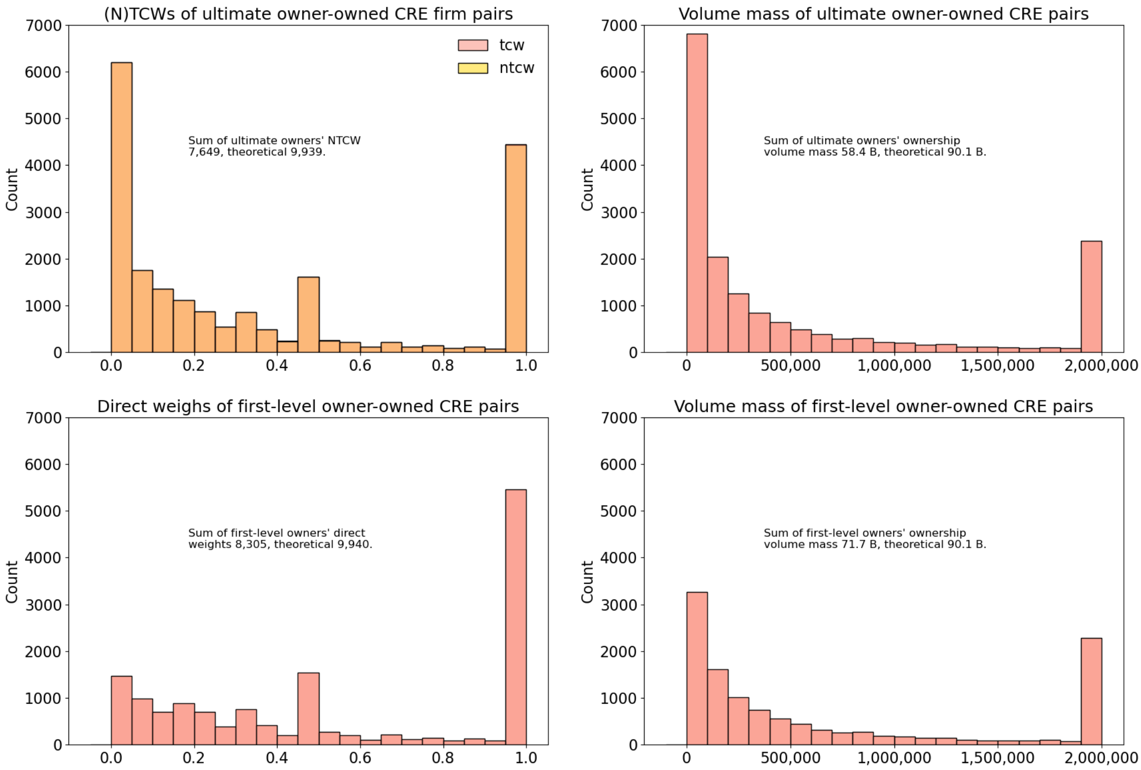
# of cycles	# of unique nodes in cycles	Avg. cycle length	# removed edges	# of ties
218	282	2.5	159	13

Panel 2

# of chains	Avg. chain length
76,710	3.2

Notes: Panel 1 provides statistics from uncycling of the original graph and Panel 2 on the resulting unwinded chains. Sources: Statistic Finland's research data and author's calculations.

Figure C.2: Ownership weight distributions for owner–CRE firm pairs



Notes: The top-left subplot depicts distributions of cumulative total weights (TCW; ownership unit mass) and net total cumulative weights (NTCW) from ultimate owner-owned CRE firm pairs. The top-right plot depicts the distribution of ultimate owners' TCW times target firm's balance sheet size (ownership volume mass). The bottom-left subplot depicts distribution of direct ownership weights (ownership unit mass) from first-level owner-owned CRE firm pairs. The bottom-right subplot depicts the distribution of first-level owners' direct weight times target firm's balance sheet size (ownership volume mass). Printed texts in each plot detail how much of the theoretical ownership masses is captured in the distributions. Bucket bins are right-inclusive. Sources: Statistic Finland's research data and author's calculations.

the orange arrow $O \rightarrow GO$ does not exist—then our preferred model would again be unbiased.

D Additional results

Table C.1 provides statistics on the unwinding process of the original ownership network and resulting unwound chains described in Section 4.1. There are 218 cycles in the original network. That is, only in a handful of cases do network motifs create cyclical patterns that require unwinding—yet this step is necessary to obtain consistently defined ownership chains. The number of unique nodes in these cycles is 282 and the average cycle length is 2.5, indicating that many cycles are simple two-node loops. 159 edges (0.3% of all edges) are removed from the network to produce an uncycled version. There are 13 ties in the unwinding process, handled as explained in Sections 4.1 and C.1. The unwinding process results in 76,710 unwound chains, which is 112% of the total number of nodes in the network. If the network consisted only of bilateral owner-owned pairs, this ratio would be 50%. This highlights that the ownership network is far from trivial, underscoring

Table D.1: Robustness regression results for discrete GO and PNO

	R1	R2	IR Spread R3	R4	R5	IR R6	IR Spread R7	R8
GO	-41.354*** (5.615)	-41.209*** (4.166)	-26.200*** (3.254)		-33.497*** (3.211)	-38.844*** (3.144)	-29.042*** (4.029)	-30.269*** (4.633)
PNO				-31.696*** (3.830)				
Debtor age	-0.439*** (0.166)	-0.452*** (0.122)	-0.247** (0.111)	-0.424*** (0.103)	-0.757*** (0.048)	-0.648*** (0.104)	-0.385*** (0.096)	-0.381*** (0.114)
# of employees	-0.317** (0.131)	-0.083 (0.189)	-0.071 (0.046)	0.003 (0.100)	-0.326*** (0.086)	-0.160** (0.077)	-0.080 (0.089)	-0.075 (0.146)
Debtor BS (mil. EUR)	0.012** (0.006)	-0.002 (0.019)	0.002 (0.002)	0.001 (0.004)	0.011*** (0.004)	0.001 (0.003)	0.001 (0.004)	-0.003 (0.008)
Orig. maturity	-1.458*** (0.323)	-0.696*** (0.205)	-0.879*** (0.171)	-0.866*** (0.134)	-1.232*** (0.108)	-1.325*** (0.149)	-0.920*** (0.136)	-1.370*** (0.193)
Loan age		-3.283*** (0.383)	-1.358*** (0.374)	-3.038*** (0.227)	-3.174*** (0.179)	-3.288*** (0.280)	-2.973*** (0.240)	-2.130*** (0.387)
Loan size (mil. EUR)	-3.868*** (0.806)	-6.282*** (0.550)	-2.543*** (0.463)	-2.889*** (0.421)	-4.538*** (0.518)	-2.659*** (0.457)	-3.004*** (0.448)	-3.048*** (0.668)
P&NP debtor							-5.773 (4.117)	
CRE sub-group	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
BG	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ref. rate type	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Purpose	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,027	5,104	9,782	9,782	32,146	11,255	9,782	5,899
R2	0.634	0.538	0.504	0.570	0.567	0.490	0.573	0.547
Adj. R2	0.628	0.535	0.502	0.568	0.566	0.488	0.572	0.544

Notes: The table shows robustness regression results with a discrete choice for the GO and PNO variables. Statistical significances: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sources: Statistic Finland's research data, Bank of Finland (AnaCredit), and author's calculations.

the need for a network-based approach. The unwound chains have an average length of 3.2 nodes.

Figure C.2 complements Figure 4 by displaying the distribution of weights used in the ownership mass calculations.[21] The top-left subplot displays histograms of the total cumulative weights (TCW) and net total cumulative weights (NTCW) between ultimate owner-owned CRE firm pairs. The former defines the ownership unit mass for such pairs. The two histograms are nearly identical, indicating that TCW and NTCW are equal in almost all cases. This observation reflects the earlier point that the network contains only a handful of cycles (in those cases, TCW can differ from NTCW). The weight distribution exhibits three distinct spikes: just above zero, around 0.5, and around 1.0, with the first being the largest. This means that, most commonly, an ultimate owner holds a highly diluted ownership share in a target CRE firm. This is reinforced by the larger distribution mass between 0 and 0.5 compared to 0.5 and 1.0, pulling the average TCW down to 0.37. However, ultimate owner holding (nearly) all or about half of a target CRE firm is not uncommon, either.

Similarly, the bottom-left subplot shows the distribution of direct weights between first-level owner-owned CRE firm pairs. This serves as the definition of ownership unit mass for such pairs. Unsurprisingly, there is much more weights mass near 1.0 compared to ultimate owners in the top-left subplot. In other words, first-level owners most commonly own the target firm fully, and

[21] The summed weights from the histograms, as well as the theoretical maximum attainable in the uncycled network, are shown next to the histograms. The former values match flow D, and the latter values equal the sum of flows C and D in Figure 4.

Table D.2: Robustness regression results for continuous GO and PNO

			IR Spread			IR	IR Spread	
	R1	R2	R3	R4	R5	R6	R7	R8
GO	-47.329*** (5.736)	-46.718*** (4.318)	-28.021*** (3.433)		-34.027*** (3.322)	-40.766*** (3.245)	-32.522*** (4.027)	-40.450*** (5.242)
PNO				-37.557*** (3.005)				
Debtor age	-0.398** (0.170)	-0.396*** (0.121)	-0.147 (0.113)	-0.372*** (0.091)	-0.753*** (0.047)	-0.609*** (0.103)	-0.348*** (0.095)	-0.391*** (0.114)
# of employees	-0.308** (0.132)	-0.083 (0.191)	-0.043 (0.044)	-0.042 (0.090)	-0.297*** (0.090)	-0.133* (0.078)	-0.050 (0.091)	-0.068 (0.147)
Debtor BS (mil. EUR)	0.007 (0.005)	-0.013 (0.015)	-0.001 (0.002)	-0.004 (0.003)	0.007* (0.004)	-0.004 (0.003)	-0.003 (0.004)	-0.004 (0.008)
Orig. maturity	-1.356*** (0.320)	-0.745*** (0.204)	-0.950*** (0.173)	-0.897*** (0.133)	-1.247*** (0.107)	-1.361*** (0.148)	-0.939*** (0.133)	-1.350*** (0.190)
Loan age		-3.164*** (0.388)	-1.341*** (0.368)	-2.942*** (0.239)	-3.206*** (0.177)	-3.323*** (0.278)	-2.966*** (0.238)	-2.126*** (0.382)
Loan size (mil. EUR)	-3.934*** (0.807)	-6.398*** (0.544)	-2.553*** (0.468)	-2.856*** (0.441)	-4.554*** (0.522)	-2.649*** (0.464)	-2.974*** (0.455)	-3.115*** (0.670)
P&NP debtor							-5.006 (4.360)	
CRE sub-group	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
BG	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ref. rate type	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Purpose	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,027	5,104	9,782	9,782	32,146	11,255	9,782	5,899
R2	0.634	0.540	0.502	0.574	0.566	0.489	0.574	0.549
Adj. R2	0.628	0.537	0.500	0.572	0.565	0.487	0.572	0.546

Notes: The table shows robustness regression results with a continuous choice for the GO and PNO variables. Statistical significances: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Sources: Statistic Finland's research data, Bank of Finland (AnaCredit), and author's calculations. Sources: Statistic Finland's research data, Bank of Finland (AnaCredit), and author's calculations.

ownership shares begin to dilute further up the ownership chain. However, 63% of the distribution mass falls outside the last bucket, indicating that smaller direct ownership shares are not uncommon. The two right-hand subplots display volume-weighted versions of the TCW and direct weight distributions, respectively, where weights are scaled by the owned firms' balance sheet size. These define ownership volume masses. The volume-weighted distributions closely resemble the raw weight distributions, although the spikes in the middle are smoothed due to variation in firm sizes.

Tables D.1 and D.2 present the results from the robustness specifications discussed in Section 5 of the main text.

E Data access and coding environment

Information about data access and delivery modes for Statistics Finland's research datasets can be found on Statistic Finland's website (Statistic Finland, 2025a). AnaCredit data are, unfortunately, not publicly available. All data processing was performed within Statistic Finland's FIONA remote access system (Statistic Finland, 2025b).

The main software tools used include Python packages *networkx* v3.1 for network analysis, (Hagberg, Schult, & Swart, 2008), *statsmodels* v0.12.2 for statistical analysis (Seabold & Perktold, 2010), *matplotlib* v3.3.4 (Hunter, 2007) and *seaborn* v0.11.1 (Waskom, 2021) for figures, as well as *pandas* v1.2.4 for data wrangling (McKinney, 2010). Extensive custom source code was developed for data extraction, data wrangling, network construction and analysis, as well as result presentation. Relevant source code will be published as the paper progresses.

References

Bank of Finland. (2023). *Kiinteistöalan luottoriskit kasvussa haastavassa toimintaympäristössä.* https://www.eurojatalous.fi/fi/2023/artikkelit/kiinteistoalan-luottoriskit-kasvussa-haastavassa-toimintaymparistossa/. (Järvenpää, Maija and Koskinen, Kimmo and Voutilainen, Ville.)

Cinelli, C., Forney, A., & Pearl, J. (2024). A crash course in good and bad controls. *Sociological Methods & Research*, *53*(3), 1071-1104. Retrieved from https://doi.org/10.1177/00491241221099552 doi: 10.1177/00491241221099552

Dagrain, S., Dumont, L., Ferrière, T., Jimborean, R., Kennedy, G., Duca, M., ... Zollino, F. (2019, 12). *Methodologies for the assessment of real estate vulnerabilities and macroprudential policies: commercial real estate* (Tech. Rep.). European Systemic Risk Board. Retrieved from https://www.esrb.europa.eu/pub/pdf/reports/esrb.report191217_methodologies_assessment_real_estate_vulnerabilities_macroprudential_policies~15ff09ae41.en.pdf

Finnish Tax Administration. (2025). *On the taxation of investment funds and the provisions of section 20a of the income tax act.* https://www.vero.fi/en/detailed-guidance/guidance/80594/on-the-taxation-of-investment-funds-and-the-provisions-of-section-20a-of-the-income-tax-act3/. (Accessed 2025-07-19.)

Hagberg, A. A., Schult, D. A., & Swart, P. J. (2008). Exploring network structure, dynamics, and function using networkx. In G. Varoquaux, T. Vaught, & J. Millman (Eds.), *Proceedings of the 7th python in science conference* (p. 11 - 15). Pasadena, CA USA.

Hunter, J. (2007). Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, *9*(3), 90–95. doi: 10.1109/MCSE.2007.55

Laaka. (2023). *Keskinäinen kiinteistöosakeyhtiö ja miten se eroaa muista kiinteistöosakeyhtiöistä.* https://laaka.fi/keskinainen-kiinteistoosakeyhtio-ja-miten-se-eroaa-muista-kiinteistoosakeyhtioista/. (Jussi Laaka, Lakimiestalo Laaka website; accessed 2025-07-18.)

McKinney, W. (2010). Data structures for statistical computing in python. In *Proceedings of the 9th python in science conference* (pp. 56–61). doi: 10.25080/Majora-92bf1922-00a

Ministry of Finance of Finland, & Prime Minister's Office of Finland. (2002). *Verosanasto - taxation glossary* (Tech. Rep.). Taloustieto Oy. Retrieved from https://valtioneuvosto.fi/

documents/10616/343825/Verosanasto_2002/a83bb463-8e34-4952-a486-6a79f1510b84/Verosanasto_2002.pdf

Pearl, J. (2009). *Causality* (2nd ed.). Cambridge University Press.

Seabold, S., & Perktold, J. (2010). statsmodels: Econometric and statistical modeling with python. In *9th python in science conference.*

Statistic Finland. (2025a). *Data and their delivery modes.* https://stat.fi/tup/tutkijapalvelut/aineistot-ja-niiden-toimitusmuodot_en.html. (Accessed 2025-07-19.)

Statistic Finland. (2025b). *Fiona remote access system.* https://stat.fi/tup/tutkijapalvelut/fiona-etakayttojarjestelma_en.html. (Accessed 2025-07-19.)

Waskom, M. L. (2021). seaborn: statistical data visualization. *Journal of Open Source Software*, *6*(60), 3021. Retrieved from https://doi.org/10.21105/joss.03021 doi: 10.21105/joss.03021

Bank of Finland Research Discussion Papers 2025

ISSN 1456-6184, online