The Family Firm: A Synthesis, Stylized Facts, and Future Research Directions

Online Appendix A: Selection step plan

1.1. Potential target companies

1.1.1. Filters Orbis

- Dutch companies
- All Active or unknown situation companies
- >10 employees
- BV and NV
- No financial, educational, social or public administration sector companies
- 56.245 companies left

1.1.2. Refine to correct Holding

- Download Global Ultimate Owner (GUO) of company list --> to gain highest level in hierarchy
- Drop duplicates in GUO's
- If GUO is missing, the company is assumed to be the GUO
- If GUO is "Stichting Administratiekantoor", then check if consists of more than 1 subsidiary in 1st level:
    o If 1 subsidiary → take this subsidiary as holding
    o If 2 or more subsidiaries → take "Stichting Administratiekantoor" as holding

1.1.3. Refine Holding / Top company

- Filter for Pension funds (insurance companies) and state/government companies
- Filter for mutual and pension funds/nominees/trusts/trustees
    o Drop GUO if the name contains: "fund/pensioen/pension/trust/participatie/participaties" (e.g. Nordian Fund I Cooperatief U.A.)
- Keep only Dutch holdings / top company (DUO) --> because of cultural differences (except Curacao(=CW))
- If GUO is family or individual ("WW* / WWP") --> no financials and no direct subsidiary information
    o Download DUO of family/individual companies
    o Look to parent company (ISH) of GUO
        ➔ If ISH = the family/individual --> take the DUO as holding
        ➔ If ISH is foreign company --> drop (but except if CW)
            o If ISH count,ry = CW → check manually what the correct top company is (4 companies)

- o   Extra: later added back one mutual and pension funds/nominees/trusts/trustees company (see above)

  ➔   If ISH country = n.a. → use DUO as top company

- If GUO is foundation (e.g. stichting administratiekantoor / StAK) --> No financial information available
  - o   Download level 1 subsidiary of GUO (except if =>1 subsidiary in 1ste level)
  - o   Delete Dutch stichting (BvD ID: NL)- foreign company structures (BvD ID: BE/LULB) -> see exemptions above
- Add all the top companies together
- 36.740 companies left (Doc: Dutch Holding List - Websites and Financial Information)
  - o   List of Dutch holdings with a subsidiary company of at least 10 employees (i.e. cancels out some holdings with more than 10> consolidated

Online Appendix B: Identification methods family businesses

1.2. Additional data to identify family businesses

1.2.1. Holding & subsidiary website (Web crawling)

Purpose: Identify family businesses by searching for specific search terms on company website

1.2.1.1 Orbis export list

- Input:
    - All holdings found in step 1.1.3
    - All subsidiaries found in step 1.2.2
- Output:
    - Company name
    - BvD ID number
    - Website address

1.2.1.2 Web Crawling method

- The methodology used for web crawling can be found at section 2.2

1.2.2. Holding plus management & subsidiaries (Name listing)

Purpose: Identify family firms by comparing board members names with company name and compare board members names with each other per company

- Input: All holdings found in step 1.1.3
- Output:
    - Company name
    - BvD ID number
    - DMC Full Name (current members)
    - DMC First Name
    - DMC Middle Name
    - DMC Last name
    - DMC Individual or company
    - DMC Corresponding BvDID (when applicable)
    - DMC Date of Birth
    - Subsidiary level
    - Subsidiary name
    - Subsidiary BvD ID number

1.2.3. Subsidiary plus management (Name listing)

Purpose: Identify family firms by comparing board members names with company name and compare board members names with each other per company.

- Input: All subsidiaries from the holdings found in step 1.2.2
- Output:
    - o Company name
    - o BvD ID number
    - o DMC Full Name (current members)
    - o DMC First Name
    - o DMC Middle Name
    - o DMC Last name
    - o DMC Individual or company
    - o DMC Corresponding BvD ID (when applicable)
    - o DMC Date of Birth

1.2.4. Companies that are a DMC in a holding or subsidiary

In this step, all the companies are listed that are a director/manager in a holding or subsidiary. The reason to get this dataset is because sometimes 3 brothers of a family have their own holding that is a director of a firm. Without this step, we only have 3 companies as a director. By extracting the directors/management of the 'company directors' in this dataset, we can compare individuals of a holding or subsidiary instead of companies.

Purpose: Identify family firms by comparing board members names with company name and compare board members names with each other per company

- Input: All DMC's that are identified as 'company' from dataset 1.2.2 & 1.2.3
- Output:
    - o Company name
    - o BvD ID number
    - o DMC First Name
    - o DMC Middle Name
    - o DMC Last name
    - o DMC Individual or company
    - o DMC Corresponding BvD ID (when applicable)

o DMC Date of Birth

2.1. Name Matching

We used in total 7 name matching methods:

1. Holding: Last name of members of board of directors - holding company name
   - Whether the last name of one of the holding directors corresponds to the holding company name (highest matching score of director of holding).

2. Holding: Last names of members of board of directors (within board)
   - The number of same last names in the board of directors of the holding company.

3. Subsidiary: Last name of members of board of directors - subsidiary company name
   - Whether the last name of one of the subsidiary directors corresponds to the subsidiary company name (highest matching score of a subsidiary is added to the holding).

4. Subsidiary: Last names of members of board of directors (within board)
   - The highest number of same last names in the board of directors of a subsidiary company is added to the holding.

5. Cross: Last name of members of board of directors holding - subsidiary company name
   - Whether the last name of one of the holding directors corresponds to one of the subsidiary company names and report the maximum score in case of multiple subsidiary matches.

6. Cross: Last name of members of board of directors' subsidiary - holding company name
   - Whether the last name of one of the subsidiary directors corresponds to the holding company name and report the maximum score in case of multiple subsidiary directors matches.

7. Cross: Last names of members of board of directors (across boards)
   - The total number of same names in the holding and subsidiary's board of directors combined.

Procedure

The name matching is performed via a pattern matching technique: Fuzzy Wuzzy. This is a pattern matching algorithm that makes use of the Levenshtein Distance[12]. The distance is a score between 0-100 (i.e. score of 100 is a full name match). We consider a name match when the distance score is 80 or higher. To match the subsidiaries with the holdings we have used the Global Ultimate Owner/ GUO BvDIDNumbers, since we than know that the holding has a majority ownership. Otherwise 2 holdings can be identified as family firm in case the subsidiary is identified as family firm and 2 holdings have a share of the subsidiary.

Then we preformed the name matching procedure with Python programs. The steps that are taken are:

1. Loading all holdings and their subsidiaries separately with the following variables:

---

[12] The Levenshtein Distance is defined to be the smallest number of edit operations (insertions, deletions and substitutions) required to change one string into another (Christen, 2006)

- Company name,
- BvDIDNumber,
- DMC (director) Full name,
- DMC Middle name,
- DMC Last name,
- DMC Individual or Company,
- DMC Corresponding BvDID (when applicable),
- DMC Date of birth,

Only for holdings:

- Subsidiary - Level,
- Subsidiary - Name,
- Subsidiary - BvDIDNumber

Unique holdings: 36.740

Number of directors that are a company: 26.962

Number of directors that are an individual: 180.598

Unique subsidiaries: 134.341

Number of directors that are a company: 95.817

Number of directors that are an individual: 361.080

2. Loading DMC companies - Because some directors are a company and these company names need to be replaced by the directors that are in the board of this DMC company. An additional dataset is download with the DMC information of these companies to replace the director information in the holding and subsidiary files. If the DMC in this file is a company, then the board member is dropped. The following variables are downloaded:

- Company name,
- BvDIDNumber,
- DMC (board) Full name,
- DMC Middle name,
- DMC Last name,
- DMC Individual or Company,
- DMC Corresponding BvDID (when applicable),

Unique DMC companies: 53.114

Number of directors that are a company: 28.633

Number of directors that are an individual: 99.849

3. Loading GUO (General Ultimate Owner) list for holdings and subsidiaries. The reason to use the GUO to match the holdings and subsidiaries is to drop the subsidiaries that the holding has a minority ownership of. Furthermore, not to identify a holding as family business if the subsidiary is a family business and the holding has minority ownership. The variables that are downloaded are:

- Company name
- BvDIDNumber
- GUO BvDIDNumber
- GUO Company name

4. Merge the DMC company data with the holding and subsidiary data in case of a director that is a company.

Unique holdings: 36.740

Unique subsidiaries: 134.022 → Less then above because duplicate Company name & BvDIDNumber are dropped now.

5. Find the GUO for each holding and subsidiary

Number of holdings with GUO BvDIDNumber: 27.441 → Not all holdings do have subsidiaries

Number of subsidiaries with GUO BvDIDNumber: 111.495

6. Find matches
   a. Subsidiaries
      i. Subsidiary company vs. Subsidiary individuals (SC_SI)
         1. Fuzzy match (between 366.764 pairs)
      ii. Director count (distinctive)
      iii. Number of directors with same name per subsidiary (26.629 subsidiaries)
   b. Holdings
      i. Holding company vs. subsidiary individuals (HC_SI)
         1. Fuzzy match (between 111.088 pairs)
      ii. Holding individuals vs. subsidiary companies (HI_SC)
         1. Fuzzy match (between 952.098 pairs)
      iii. Holding company vs. holding individuals (HC_HI)
         1. Fuzzy match (between 67.791 pairs)
      iv. Holding individuals vs. subsidiary individuals (HI_SI)
         1. Match based on:
            a. GUO - BvDIDNumber

        b. GUO - Name

        c. DMC Last name

        d. DMC Date of birth

      2. Number of possible director pairs in total: 5.424.104

    v. Number of directors with same name per holding (35.149 holdings)

7. Match all data with corresponding holding. For all values found per subsidiary the maximum score is used per holding. In case of missing subsidiary for the holding, the value 'PWC_Unknown' is placed.

8. Export a CSV file with all holdings and their corresponding value per variable.

Results

Below the amount of observations that have a value per variable:

- Number of holdings: 36.740
- Number of holdings with fuzzyMatch_HI_SC: 25.762
- Number of holdings with SameName_HOLD: 35.149
- Number of holdings with fuzzyMatch_HC_HI   35.149
- Number of holdings with #Managers_SUB: 26.629
- Number of holdings with fuzzyMatch_SC_SI: 24.907
- Number of holdings with SameName_SUB: 24.907
- Number of holdings with fuzzyMatch_HC_SI: 24.907
- Number of holdings with HI_SI: 24.210

2.2. Web Crawling

We used in total 2 name matching methods

1. Holding: web crawl (Holding_Webcrawl)
- Whether the website of the holding contains one of the following words: "familiebedrijf", "family business"
2. Subsidiary: web crawl (Sub_Webcrawl)
- Whether the website of the subsidiary contains one of the following words: "familiebedrijf", "family business"

2.2.1 Procedure

- Collect website data for the 36.740 holding companies and 132.972 subsidiary companies from Orbis. Website data is available for 51% of the holding companies and 56% of the subsidiary companies.

- In the database with all URLs keep only unique URLs to avoid web crawling of one website multiple times.

- Perform web crawling in the Microsoft Azure cloud network: assign a "worker" to search the URLs in the network via the Bing search engine, this is built in by Microsoft. The worker clicks on every webpage linked to the URL and looks whether "familiebedrijf" or "family business" can be found on the webpage. When a match is made the worker reports the corresponding URL and a snippet of the text of the webpage.

- When the worker has found 20 sub-URL hits on one main URL it stops screening the main URL to avoid long processing times and web crawling (overload) failures.

- Total companies with websites 93.583 (Holdings: 18.733, Subsidiaries: 74.850).

- Removed duplicate websites: in total 52.665 unique websites (one particular URL can link to different companies).

- 54.210 web pages (one URL may have many web pages) contain the words 'familiebedrijf'/'family business', this means that 10.229 websites (i.e. URLs) have at least one web page with the search terms.

- To increase the predictability of the web crawl measure to identify family firms we apply several filters to filter out non-family companies having a website contain the search terms. We have applied several URL-filters. That is, we dropped the observation if one of the following words can be found in the URL:
    - 'Blog', 'nieuws', 'news', & 'actueel': these words often refer to news pages mentioning to other (family) companies. When the URL contains "over-ons" we do not drop the observation, since the web page is than often written by the family business itself.
    - 'Vacature': many employment websites show job openings of family businesses and also mention this on their website. However, many family company websites only mention the family aspect of the business in the job section of their website. Therefore, we made sure to only remove vacature-URLs when the rest of the URL does not contain "werken-bij", "werkenbij","werken_bij". These words are often used by family firms themselves.

- In the end, 9.164 unique firm websites (both Holding and Subsidiaries) are left. In total 4.874 holding companies have either a holding website or subsidiary website containing the search terms. This is just a small amount of the total holding list and we suggest that this method can be improved in future research.

2.3. Family Business Lists

- Within PwC several lists exist with names of family businesses (±500). As there is a very high chance that these companies will fulfil the European Commission criteria to be a family business, we use them as dependent variable (Y=1) to build the predictive model. We use this model to identify family business that we would like to send a survey.
- The predictive model can only be built if the database also contains firms with a high likelihood to be a non-family business (Y=0). Therefore, we collected information from various sources (e.g. Elsevier, LexisNexis) to create a list with non-family businesses. Elsevier has a list of largest Dutch firms where they indicate if a firm is a family business or not.

2.4. Final Family Business Identification (Predictive model)

With the information collected in step 2.1 and 2.2 the predictive models of step 2.3 are made. The predictive models are made with help of the firms of which we know that it is a family business or not. While creating the models we used several variables that we collected in different groups, since there is a large correlation between them (which was expected). Because of the highly correlated variables we created 1 model with a relatively high R^2 and without conflicting correlated variables.

We also did a variance inflation factor (VIF) test on a linear regression of our variables as an indicator for multicollinearity. In the linear model with all the variables included the VIF value was still below 5, which is recommended by Rogerson (2001)[13]. Therefore, we used all the identification variables in our main model. Next to the main model we made 7 models to identify family firms, since we have missing value for some variables because of different company structures of the holdings.

For the different models, we did an analyse what the best value was for a Y-hat to predict if the firm is as family business. At the end, we used the following variables for our predictive models:

Model used for holdings with information for all variables:
Model 1 (#22.768): probit FF_Casper_daniel sq_n_samename_hold HC_HI_match sq_n_samename_sub HI_SC_match HC_SI_match SC_SI_match tot_webcrawl
Minimal Y-hat prediction value: E-mail = 0.6 and Letter = 0.75

Model used for holdings with all information except website information:
Model 2 (#1.439): probit FF_Casper_daniel sq_n_samename_hold HC_HI_match sq_n_samename_sub HI_SC_match HC_SI_match SC_SI_match sqrt_n_samename_hold_sub
Minimal Y-hat prediction value: E-mail = 0.6 and Letter = 0.8

Model used for holdings with all information except subsidiary information:

---

[13] Rogerson, P. A. (2001). Statistical methods for geography. London: Sage.

Model 3 (#8.839): probit FF_Casper_daniel sq_n_samename_hold HC_HI_match tot_webcrawl

Minimal Y-hat prediction value: E-mail = 0.65 and Letter = 0.75

Model used for holdings without subsidiary and website information:

Model 4 (#2.103): probit FF_Casper_daniel sq_n_samename_hold HC_HI_match

Minimal Y-hat prediction value: E-mail = 0.65 and Letter = 0.65

Model used for holdings without holding board information (hiding family ownership and wealth):

Model 5 (#639): probit FF_Casper_daniel n_samename_sub SC_SI_match HC_SI_match tot_webcrawl

Minimal Y-hat prediction value: E-mail = 0.65 and Letter = 0.65

Model used for holdings without holding board and website information (hiding family ownership and wealth):

Model 6 (#61): probit FF_Casper_daniel n_samename_sub SC_SI_match HC_SI_match tot_webcrawl

Minimal Y-hat prediction value: E-mail = 0.65 and Letter = 0.65

Model used for holdings with only web crawling information:

Model 7 (#891): tot_webcrawl

Used if variable = 1 (Yes)

Variables used in models:

FF_Casper_daniel:	Family businesses identified by previous work PwC and Casper de Nooijer
sq_n_samename_hold:	Squared number of same last names for holdings
HC_HI_match:	Highest fuzzy match score of holding company name and holding individual
sq_n_samename_sub:	Squared number of same last names for subsidiaries
HI_SC_match:	Highest fuzzy match score of holding individual and subsidiary company name
HC_SI_match:	Highest fuzzy match score of subsidiary individual and holding company name
SC_SI_match:	Highest fuzzy match score of subsidiary individual and subsidiary company name
tot_webcrawl:	Dummy variable that is 1 if holding or subsidiary company website contains family firm values

Appendix C: Financial Consolidation

3.1 Input files

Two different company files are used as input for the financials data:

- o Holding List
  - ➔ Holdings (Company name, BvD ID number)
  - ➔ Subsidiaries (Company name, BvD ID number)
  - ➔ Consolidation code
- o Subsidiary List
  - ➔ Financial variables
  - ➔ Consolidation code

3.2 Output files (Orbis)

Downloaded variables per company file:

- ○ BvD ID number
- ○ Number of Employees group information (last value)
- ○ Number Companies in corporate group
- ○ Consolidation Code
- ○ Listed / Delisted / Unlisted
- ○ Number of patents
- ○ Number of Employees per year (2009-2015)
- ○ Total Assets per year (2009-2015) EUR
  - Current Assets (2009-2015) EUR
  - Tangible fixed assets (2009-2015) EUR
  - Cash & Cash equivalent (2009-2015) EUR
- ○ Shareholders' Funds (2009-2015) EUR
- ○ Total Debt: Current + non-current liabilities (2009-2015) EUR
- ○ Non-current Liabilities (2009-2015) EUR
  - Long term debt (2009-2015) EUR
  - Other non-current liabilities (The one without capitals in name ORBIS)
- ○ Current Liabilities (2009-2015) EUR
- ○ Total Sales: Sales (2009-2015) EUR
- ○ Cost of goods sold (2009-2015) EUR
- ○ Other operating expenses (2009-2015) EUR
- ○ P/L for period [Net Income] (2009-2015) EUR

Variables with a different holding and download format

- ○ Date of incorporation
- ○ US SIC core code (3 digits)
- ○ US SIC primary code (option: first line only)

3.3 Procedure consolidate the data (output files Orbis)

We distinguish several types consolidation types of holdings:

1. C1: Companies with consolidated accounts only (2.242 holdings)
2. C2/U2: Companies with both types of accounts (5.904 holdings)
3. U1: Companies with unconsolidated accounts only (25.956 holdings)
4. NRF/LF/NRLF/NF: companies with no (recent)/limited financial information (5.346 holdings)

1. *C1 companies*

   Do not require any special attention as these firms do report consolidated accounts across the years.

   → To control for variation caused by the self-made consolidation method in the next consolidation type of companies we added a dummy variable which takes the value 1 if the consolidated account was not available for that particular year, and the consolidation was performed by the researchers. For *C1 companies* the consolidation account was available for all observations, so takes value 0.

2. *C2/U2 Companies*

   Do require attention as these firms have reported consolidated accounts in some particular years. However, the consolidated accounts are not available for all reporting years. The following procedure is followed to consolidate all fiscal years:

- ○ Search and download strategy:
  - → Load Holding BvDIDNumbers in Orbis (36.740 holdings)
  - → Type of accounts: C2/U2 (Companies with both types of accounts) (5.904 holdings)
  - → Searches settings file 1 (Holdings financials - consolidated) with settings: *a preferred account* and then on *consolidated accounts*
  - → Searches settings file 2 (Holdings financials - unconsolidated) with settings: *a preferred account* and then on *unconsolidated accounts*
  - → Import the majority owned subsidiaries in Orbis (25.703 subsidiaries)

→ Search settings file 3 (Subsidiary financials - unconsolidated) with settings: *a preferred account* and then on *unconsolidated accounts* (due to download limitations on Orbis downloaded 3 separate files)

- 545 holdings, no subsidiary
- 219 holdings dropped because one of the subsidiaries has cons. code C1.
  - In case of match survey respondents with these data, we have to manually check the right numbers.
- 2 holdings dropped, because the subsidiary is the holding itself

○ Consolidation procedure:

→ Summed all unconsolidated subsidiary information per holding (21.500 subsidiaries, 5.138 holdings)

→ Merge unconsolidated holding information, self-consolidated subsidiary information and consolidated holding information and transpose and order data to panel data set

→ Determine the right consolidation method:

- For patents: the sum of unconsolidated holding information and self-consolidated subsidiary information
- For Employees: the consolidated holding information or if missing the unconsolidated holding information + self-consolidated subsidiary information. If the consolidated holding employee amount is lower than 10 we also take the unconsolidated holding information + self-consolidated subsidiary information
- For the financial variables two type of consolidated accounts are available:
  - With elimination entries (=>2 years consolidated information from holding): The consolidated account taken into account is equal to the given holding consolidated account or if the holding consolidated account is missing, then we calculated the mean elimination entry over the number of years consolidated holding information was available and subtracted this mean from the self-consolidated financials (unconsolidated holding information + consolidated subsidiary information).
  - Without elimination entries (<2 years consolidated information from holding): The consolidated account taken into account is equal to the given holding consolidated account or if the holding consolidated account is missing we take into account the self-consolidated financials

(unconsolidated holding information + consolidated subsidiary information).

→ To control for variation caused by the consolidation method we added a dummy variable which takes the value 1 if the consolidated account was not available for that particular year, so we had to do the consolidation ourselves.

→ 5.138 holdings consolidated

○ Monitoring and improving the consolidation method:

→ The holding is sometimes a subsidiary of the holding itself. To avoid double counting we drop the subsidiary observation (91 Subsidiaries)

→ Double counting of financials; it appears that in the holdings the financials are already partly consolidated. To avoid this problem, we used an elimination entry estimation procedure.

3. *U1 Companies*

Do require attention as these firms have reported no consolidated accounts in all years. The following procedure is followed to consolidate all fiscal years:

○ Search and download strategy:

→ Load Holding BvD ID Numbers in Orbis (36.740 holdings)

→ Type of accounts: U1 (Companies with unconsolidated accounts only) (25.957 holdings)

→ Searches settings file 1 (Holdings Financials): the most recent accounts available (due to download limitations on Orbis downloaded as 3 separate files)

→ Import the majority owned subsidiaries in Orbis (45.300 subsidiaries)

→ Searches settings file 2 (Subsidiary Financials): the most recent accounts available (due to download limitations on Orbis downloaded as 5 separate files)

■ 8.411 holdings, no subsidiary

■ 158 holdings dropped because one of the subsidiaries has cons. code C1.

■ 4 holdings dropped, because the subsidiary is the holding itself

○ Consolidation procedure:

→ Summed all unconsolidated subsidiary information per holding (41.456 subsidiaries, 17.384 holdings)

→ Merge unconsolidated holding information with self-consolidated subsidiary information and transpose and order data to panel data set

→ Summed all self-consolidated subsidiary information to the unconsolidated holding information.

→ 601 holdings (3.982 subsidiaries) have subsidiaries with C2/U2 code (companies with consolidated and unconsolidated accounts. This is an issue, since some yearly accounts could be/are consolidated while others are not consolidated. This would result in double-counting if we are consolidating manually ourselves. Solve this by running the procedure again for these companies.

→ Searches settings file 3 (Subsidiary Financials) with settings: *a preferred account* and then on u*nconsolidated accounts*

○ Consolidation procedure:

→ Summed all unconsolidated subsidiary information per holding (3.982 subsidiaries, 601 holdings)

→ Merge the self-consolidation data of holdings with C2/U2 subsidiaries with the main file

→ We checked whether the new consolidated information is better than the initial consolidation value. In many cases (2.678 accounting years) the initial value corresponds to the new consolidation value. However, in some cases the values largely deviate (new value <20% of initial) from each other (621 accounting years). We take certain consolidation values in the following cases:

  ○ The number of employees of "*the most recent account available (step 3)*" consolidation is more than or equal to 5 times the number of employees of "*the preferred account and then unconsolidated accounts (step 7)*" consolidation -> take the number of employees of "*the most recent account available (step 3)*" consolidation value

  ○ The number of employees of "*the most recent account available (step 3)*" consolidation is less than 5 times the number of employees of "*the preferred account and then unconsolidated accounts (step 7)*" consolidation

    -> take the the number of employees of "*the preferred account and then unconsolidated accounts (step 7)*" consolidation value

  ○ To maintain consistency within the dataset we adjust the other financial variables in accordance with the employee adjustment procedure.

→ To control for variation caused by the self-consolidation method in the next consolidation type of companies we added a dummy variable which takes the value 1 if the consolidated account was not available for that particular year, so we had to do the consolidation ourselves. For *U1 companies* the consolidation account was not available for all observations, so the dummy variable always takes value 1.

→ To control for variation caused by the elimination entries we added a dummy variable which takes the value 1 if the employee ratio (see elimination entries adjustment) is below 0.2.

→ Most problems seem to be solved with this solution, however it is still required to manually check the financial variables in the final dataset.

4. *NRF/LF/NRLF/NF Companies*

Do require attention as these firms have not reported consolidated or unconsolidated accounts or have reported limited financial information for all the years. The following procedure is followed to consolidate all fiscal years:

○ Search and download strategy:

→ Load Holding BvD ID Numbers in Orbis (36.740 holdings)

→ Type of accounts: NRF/LF/NRLF/NF (Companies with unconsolidated accounts only) (5.346 holdings - with 22.520 subsidiary companies)

→ Searches settings file 1 (Holdings financials - consolidated): *a preferred account* and then on *consolidated accounts*

→ Searches settings file 2 (Holdings financials - unconsolidated): *a preferred account* and then on *unconsolidated accounts*

→ Searches settings file 3 (Subsidiary Financials) with settings: *a preferred account* and then on u*nconsolidated accounts* (due to download limitations on Orbis downloaded 3 separate files)

- 1.298 holdings, no subsidiary
- 10 subsidiaries do not belong to the holding anymore
- 258 holdings dropped because one of the subsidiaries has cons. code C1.
- 2 holdings & 49 subsidiaries dropped, because the subsidiary is the holding itself

○ Consolidation procedure:

→ Summed all unconsolidated subsidiary information per holding (16.475 subsidiaries, 3.788 holdings)

→ Merge unconsolidated holding information, self-consolidated subsidiary information and consolidated holding information and transpose and order data to panel data set

→ Determine the right consolidation method:

- For patents: the sum of unconsolidated holding information and self-consolidated subsidiary information

- ■ For Employees: the consolidated holding information or if missing the unconsolidated holding information + self-consolidated subsidiary information. If the consolidated holding employee amount is lower than 10 we also take the unconsolidated holding information + self-consolidated subsidiary information
- ■ For the financial variables two type of consolidated accounts are available:
  - With elimination entries (=>2 years consolidated information from holding): The consolidated account taken into account is equal to the given holding consolidated account or if the holding consolidated account is missing, then we calculated the mean elimination entry over the number of years consolidated holding information was available and subtracted this mean from the self-consolidated financials (unconsolidated holding information + consolidated subsidiary information).
  - Without elimination entries (<2 years consolidated information from holding): The consolidated account taken into account is equal to the given holding consolidated account or if the holding consolidated account is missing we take into account the self-consolidated financials (unconsolidated holding information + consolidated subsidiary information).
- → To control for variation caused by the consolidation method we added a dummy variable which takes the value 1 if the consolidated account was not available for that particular year, so we had to do the consolidation ourselves.

5. *Wrap - up: Finalized companies vs. non-finalized companies*
   - ○ Finalized companies (Total = 28.551 holdings):
- → C1 companies: 2.242 holdings
- → C2/U2 companies: 5.138 holdings
- → U1 companies: 17.383 holdings
- → NRF companies: 3.788 holdings
   - ○ Non-finalized companies (Total = 10.897 holdings):
- → C1 companies: no holdings
- → C2/U2 companies: 764 holdings
   - ○ 545 holdings no subsidiary
     - ■ Take either the consolidated or the unconsolidated financial value

- - - To control for variation caused by the consolidation method we added a dummy variable which takes the value 1 if the consolidated account was not available for that particular year.
    - 219 holding one of the subsidiaries has cons. code C1.
  - → U1 companies: 8573 holdings
    - 8.411 holdings no subsidiary
      - Take the unconsolidated financial value
      - To control for variation caused by the consolidation method we added a dummy variable which takes the value 1 if the consolidated account was not available for that particular year (i.e. all cases)
    - 154 holdings one of the subsidiaries has cons. code C1.
  - → NRF companies: 1558 holdings
    - 1.298 holdings, no subsidiary
      - Take the unconsolidated financial value
      - To control for variation caused by the consolidation method we added a dummy variable which takes the value 1 if the consolidated account was not available for that particular year (i.e. all cases)
    - 258 holdings dropped because one of the subsidiaries has cons. code C1.

    - Final dataset (Total = 36.498)
  - → C1 companies: 15,694 lines 2.242 holdings
  - → C2/U2 companies: 35,966 lines 5.138 holdings
  - → U1 companies: 121.688 lines 17.384 holdings
  - → NRF companies: 1.480 holdings unique
    - 182 lines not matched: 26 holdings
    - 26.334 lines matched: 3.762 holdings
      - Double counts C1: 133 lines, 19 holdings
      - Double counts C2: 15.393 lines, 2.199 holdings
      - Double counts U1: 448 lines, 64 holdings
  - → Non-finalized companies - no subsidiary: 71.778 lines 10.254 holdings

- Consolidation Assumptions:
  1. Differences in financials due to partial consolidation, Orbis database reliability
  2. 100% consolidation of majority owned subsidiaries

3. "Vastgoed beheer" not taken into account. Mostly the real estate is placed in a company that is part of the holding. However, sometimes it is spited from the holding and then the assets are not taken into account.

6. For efficiency purposes we determined the age, diversification and industry of the corporate group in a different file.
   - ○ Search and download strategy:
     - → Load Holding BvD ID Numbers in Orbis (36.740 holdings)
     - → Load Subsidiary BvD ID Numbers in Orbis (109.425 Subsidiaries)
     - → Download the following variables:
       - ■ Date of incorporation
       - ■ US SIC Core code (3-digit)
       - ■ US SIC primary code (option: first line only)
       - ■ Total Assets EUR Last avail.
     - → Subsidiaries: Transformed the data to determine the incorporation year, main industry and diversification levels.
       - ■ Corporate incorporation year: The oldest incorporation year of the subsidiary companies per corporate group. This measure is used to determine the company age.
       - ■ Main industry per corporate group: SIC code related to the industry in which the company has devoted the highest amount of total assets.
       - ■ Diversification 1: The number of different primary SIC codes per holding
       - ■ Diversification 2: A dummy variable which takes the value 1 if the number of different primary SIC codes per holding is larger than 2, and 0 otherwise.
       - ■ Diversification 3: Weighted average diversification measure (Montgomery, 1982)[14]
     - → Holding: Merge Subsidiary file into Holding file.
       - ■ Checked whether subsidiary incorporation year is the earliest incorporation year of the corporate group.
       - ■ Main industry

---

[14] Montgomery, C. A., (1992). The Measurement of Firm Diversification: Some New Empirical Evidence. *Academy of Management Journal.* June 1, 1982 vol. 25 no. 2 299-307

- If SIC code is 6712/6719 we do not take the amount of assets into account. SIC 6000-6999 is a financial firm.
- To define the corporate group main industry code, we compare the amount of assets of the holding industry and the main subsidiary industry. The industry with the highest amount of assets is taken as the corporate group main industry code.

Appendix D: Survey set-up

Strategy

Method: Cross-Sectional Survey, collection of data at a single point in time from a sample drawn from a specified population.
Sampling:
- Proposed strategy: target all family companies within the dataset.
- Survey respondent:
  - CEO: Believed to be one of the most informed person within the company, but difficult to directly target.
  - Other employees: mostly family members and senior management.

Proposed strategy: All the questions are suited to be filled in by family members or senior management. This way, we think the response rate will be higher than when we solely focus on CEOs.
- Targeting strategy:
  - Letter: send a letter to the holding address (in many cases this will be the home address of the family owners) with the request to fill in the survey online. As we often know the manager name, we could address the letter to a specific person (t.a.v. CEO/de heer/mevrouw) and the response rate is believed to be higher than in other methods.
    - Link to Tilburg University website (create weblog with main results with the goal that companies are willing to fill in survey, because they get something in return)

  - General e-mail addresses (info@ addresses): Send to known e-mail addresses an email with request to fill-in the survey. This is a cheap, easy and very scalable method.
    1. After online check recognised that most companies use an info@website.nl address. Thus, we could try to send the email to this address of each of the companies.

2. Is it possible to use Tilburg University logo in mail to create trust and therefore the response rate.

3. Unique log-in / identification code to be able to match survey data to current database

4. Reminder emails

5. T.a.v. mail -> less effective than in 'closed' letter

○ FBNed/other websites (Tilburg University, LinkedIn, Facebook): Create posts on websites and social media with request to fill in survey. This is also a very easy, cheap and scalable method.

1. It needs to be clear who the audience group is in the online post.

2. In the survey that is automatically send, you can make an identifier to link the answers to the dataset. The answer from online replies do not have this identifier.

Proposed strategy: We would like to target respondents via a mailing and a letter, since we belief that this will be the most efficient method. It would be interesting to link the survey to a Tilburg University website to increase the response rate and awareness of the existence of the institution. One major problem, as we would like to target a large group of family companies, is linking the survey results directly to our database. We hope to solve this by generating a unique code linked to each questionnaire and company.

- Questionnaire design: Online questionnaire, which can be directly linked to the main database with company information (e.g. financial and management).
- Question type: Both open and closed-ended questions (see also research set-up document)
- Question order/survey length, will be determined after the pretesting period.
  ○ Possibly rotate answers in closed-ended questions for response order effects
  ○ Start with simple questions, then more difficult and grouped per topic.
  ○ Change question order if possible
  ○ Asking to believes from the past is often incorrect, because people think they have always believed what they believe now

- Pretesting method - face-to-face survey
  ○ Who? -> Pesonal network . In total we target to try out the survey at 15-25 companies to make sure the questions are formulated clearly and the length is right.
  ○ Data → Take the data that we have to some face-to-face companies to check if the data derived from the Orbis database is right. For example: are the people from the board similar and is the number of employees correct.
- Survey period: 5-6 weeks and send reminder after 2 weeks.

    ○   Self-administered surveys take a quiet long collection period.

We would like to target all the identified family firms in the database (52%[15] of 36.470 = 19.000 FFs) via a combined email and a letter mailing. As we belief that the response rate is relatively low for these methods (i.e. between 5-20%), we are going to identify several methods to increase the response rate. Next to this, we would like to leverage several family business networks, to increase the awareness and the willingness to participate in the survey.

Time-line

14 September: Send-out survey

14 September: Publish posts on Twitter/LinkedIn/ (Facebook?)

14 September: Publish article PwC / TiU

18 September: Publish article FBNed

21 September: Publish article Fambizz

29 October: Close survey

Channels[16] (prior launch)

Total family population = 19.000 (52%)

Prediction accuracy = 93% -> 17.670 FFs

- ● Email Mailing

  Exp. target rate= Appr. 65% of the dataset-> exp. respondents = 1.150/575 (Rr=10%/5%)

  - ○ Apr. 11 % of the email addresses could be derived from Orbis.
    - ■ 1.000/36.000 holdings =3%)
    - ■ 20.000/134.000 subsidiaries (complete list of subsidiaries) = 3.000/36.000 holdings (=8%)
  - ○ Apr. 60% of the email addresses could be derived by transforming the website-addresses → As done in name matching sample
  - ○ Identification: send unique respondent URL for Qualtrics (survey 1)

- ● Letter

---

[15] Vanaf empl> 10 https://www.cbs.nl/-/.../16/2017ep23%20familiebedrijven%20in%20nederland.pdf

[16] Calculations are based on assumptions

Expected target rate = Apr. 90% -> exp. respondents = 3.180/1.590 (Rr = 20%/10%)

- ○ It is required for companies to fill in the company address for the registration at the chamber of commerce, so we have many postal addresses to send the letter to.
- ○ Identification: send unique log-in code for Qualtrics (survey 1)

- Social Networks

Expected target rate = Apr. 10% -> exp. Respondents = 1.400/1.200 (Rr = 80%/70%)

- ○ Leverage the networks, FBNed, LinkedIn and Tilburg University
- ○ Identification: provide link to general survey and ask several identification questions in Qualtrics (survey 2)
- ○ Twitter https://twitter.com/succesopvolging @succesopvolging tweeted about the research
- ○ NAEVO (news and knowledge platform for entrepreneurship) wrote an article about the research: http://naevo.nl/nieuw-onderzoek-naar-prestaties-familiebedrijven/
- ○ Rendement.nl wrote an article about the research: https://www.rendement.nl/nieuws/id20176-familiebedrijven-kunnen-deelnemen-aan-onderzoek.html

Methods to increase response rate (prior launch)

- Tilburg University - Tilburg Institute for Family Business
  - ○ We belief that we could leverage the name of the university to increase the response rate.
  - ○ Link survey to website of institute - make a subpage on the TiFB page at ww.tiu.nl → https://www.tilburguniversity.edu/research/institutes-and-research-groups/tifb/
  - ○ Use name and colours of Tilburg University

- Newspapers & Other media channels -> FamBizz & FBNed
- Family Business Environment
  - ○ Many family businesses believe that being a family business brings many advantages and family firms are vital for the economy, especially after the financial crisis.
  - ○ -> Higher willingness to participate in events/surveys who support and acknowledge the importance of family businesses.
  - ○ -> Family business cases (e.g. *alles is familie* television programme) show the similarities between family businesses and reveal factors which are very recognizable (e.g. grandfather still working in company) -> many family business owners like this programme
  - ○ -> Conclusion: create a "family business environment" around the survey.

Channels to increase response rate (Communication 2.0 - After launch)

- Reminder E-mail
- How to reach large companies, sales >75 million?
    - Personal Network
    - Add survey link to signature
    - Direct mailings to close contacts
- How to reach other medium to large companies?
    - Gelders familiebedrijven gilde
    - Brabants familiebedrijven genootschap
    - KvK
- Twitter:

#TilburgUniversity doet uniek onderzoek naar #familiebedrijven. #Familiebedrijf: deel uw ervaring via http://bit.ly/TiFBonderzoek

    - @VNONCW (21K)
    - @tilburgresearch (2488)
    - @succesopvolging
    - @FBNed (515)
    - @fambizznl (959)
    - @mkbnl (45K)
    - @OP_Nederland (90K)
    - @Dormaar (0.7K) - redacteur Rtl Z (familiebedrijven artikel)
    - @BV_Rendement (0.9K) - published article about research

Method: Cross-Sectional Survey, collection of data at a single point in time from a sample drawn from a specified population.

Sampling:

- Proposed strategy: target all family companies within the dataset.
- Survey respondent:
    - CEO: Believed to be one of the most informed person within the company, but difficult to directly target.
    - Other employees: mostly family members and senior management.

Proposed strategy: All the questions are suited to be filled in by family members or senior management. This way, we think the response rate will be higher than when we solely focus on CEOs.

- Targeting strategy:

- ○ Letter: send a letter to the holding address (in many cases this will be the home address of the family owners) with the request to fill in the survey online. As we often know the manager name, we could address the letter to a specific person (t.a.v. CEO/de heer/mevrouw) and the response rate is believed to be higher than in other methods.
    - ■ Link to Tilburg University website (create weblog with main results with the goal that companies are willing to fill in survey, because they get something in return)

- ○ General e-mail addresses (info@ addresses): Send to known e-mail addresses an email with request to fill-in the survey. This is a cheap, easy and very scalable method.
    6. After online check recognised that most companies use an info@website.nl address. Thus, we could try to send the email to this address of each of the companies.
    7. Is it possible to use Tilburg University logo in mail to create trust and therefore the response rate.
    8. Unique log-in / identification code to be able to match survey data to current database
    9. Reminder emails
    10. T.a.v. mail -> less effective than in 'closed' letter

- ○ FBNed/other websites (Tilburg University, LinkedIn, Facebook): Create posts on websites and social media with request to fill in survey. This is also a very easy, cheap and scalable method.
    3. It needs to be clear who the audience group is in the online post.
    4. In the survey that is automatically send, you can make an identifier to link the answers to the dataset. The answer from online replies do not have this identifier.

Proposed strategy: We would like to target respondents via a mailing and a letter, since we belief that this will be the most efficient method. It would be interesting to link the survey to a Tilburg University website to increase the response rate and awareness of the existence of the institution. One major problem, as we would like to target a large group of family companies, is linking the survey results directly to our database. We hope to solve this by generating a unique code linked to each questionnaire and company.

- Questionnaire design: Online questionnaire, which can be directly linked to the main database with company information (e.g. financial and management).
- Question type: Both open and closed-ended questions (see also research set-up document)
- Question order/survey length, will be determined after the pretesting period.
    - ○ Possibly rotate answers in closed-ended questions for response order effects

- ○ Start with simple questions, then more difficult and grouped per topic.
- ○ Change question order if possible
- ○ Asking to believes from the past is often incorrect, because people think they have always believed what they believe now

- Pretesting method - face-to-face survey
  - ○ Who? -> Pesonal network . In total we target to try out the survey at 15-25 companies to make sure the questions are formulated clearly and the length is right.
  - ○ Data → Take the data that we have to some face-to-face companies to check if the data derived from the Orbis database is right. For example: are the people from the board similar and is the number of employees correct.
- Survey period: 5-6 weeks and send reminder after 2 weeks.
  - ○ Self-administered surveys take a quiet long collection period.

We would like to target all the identified family firms in the database (52%[17] of 36.470 = 19.000 FFs) via a combined email and a letter mailing. As we belief that the response rate is relatively low for these methods (i.e. between 5-20%), we are going to identify several methods to increase the response rate. Next to this, we would like to leverage several family business networks, to increase the awareness and the willingness to participate in the survey.

Time-line

14 September: Send-out survey

14 September: Publish posts on Twitter/LinkedIn/ (Facebook?)

14 September: Publish article PwC / TiU

18 September: Publish article FBNed

21 September: Publish article Fambizz

29 October: Close survey

Channels[18] (prior launch)

Total family population = 19.000 (52%)

---

[17] Vanaf empl> 10 https://www.cbs.nl/-/.../16/2017ep23%20familiebedrijven%20in%20nederland.pdf

[18] Calculations are based on assumptions

Prediction accuracy = 93% -> 17.670 FFs

- Email Mailing

  Exp. target rate= Appr. 65% of the dataset-> exp. respondents = 1.150/575 (Rr=10%/5%)

  - Apr. 11 % of the email addresses could be derived from Orbis.
    - 1.000/36.000 holdings =3%)
    - 20.000/134.000 subsidiaries (complete list of subsidiaries) = 3.000/36.000 holdings (=8%)
  - Apr. 60% of the email addresses could be derived by transforming the website-addresses → As done in name matching sample
  - Identification: send unique respondent URL for Qualtrics (survey 1)

- Letter

  Expected target rate = Apr. 90% -> exp. respondents = 3.180/1.590 (Rr = 20%/10%)

  - It is required for companies to fill in the company address for the registration at the chamber of commerce, so we have many postal addresses to send the letter to.
  - Identification: send unique log-in code for Qualtrics (survey 1)

- Social Networks

  Expected target rate = Apr. 10% -> exp. Respondents = 1.400/1.200 (Rr = 80%/70%)

  - Leverage the networks, FBNed, LinkedIn and Tilburg University
  - Identification: provide link to general survey and ask several identification questions in Qualtrics (survey 2)
  - Twitter **Liane en Carola** https://twitter.com/succesopvolging @succesopvolging tweeted about the research
  - NAEVO (news and knowledge platform for entrepreneurship) wrote an article about the research: http://naevo.nl/nieuw-onderzoek-naar-prestaties-familiebedrijven/
  - Rendement.nl wrote an article about the research: https://www.rendement.nl/nieuws/id20176-familiebedrijven-kunnen-deelnemen-aan-onderzoek.html

Methods to increase response rate (prior launch)

- Tilburg University - Tilburg Institute for Family Business
  - We belief that we could leverage the name of the university to increase the response rate.
  - Link survey to website of institute - make a subpage on the TiFB page at ww.tiu.nl → https://www.tilburguniversity.edu/research/institutes-and-research-groups/tifb/
  - Use name and colours of Tilburg University

- Newspapers & Other media channels -> FamBizz & FBNed
- Family Business Environment
  - Many family businesses believe that being a family business brings many advantages and family firms are vital for the economy, especially after the financial crisis.
  - -> Higher willingness to participate in events/surveys who support and acknowledge the importance of family businesses.
  - -> Family business cases (e.g. *alles is familie* television programme) show the similarities between family businesses and reveal factors which are very recognizable (e.g. grandfather still working in company) -> many family business owners like this programme
  - -> Conclusion: create a "family business environment" around the survey.

Channels to increase response rate (Communication 2.0 - After launch)

- Reminder E-mail
- How to reach large companies, sales >75 million?
  - Personal Network
  - Add survey link to signature
  - Direct mailings to close contacts
- How to reach other medium to large companies?
  - Gelders familiebedrijven gilde
  - Brabants familiebedrijven genootschap
  - KvK
- Twitter:

#TilburgUniversity doet uniek onderzoek naar #familiebedrijven. #Familiebedrijf: deel uw ervaring via http://bit.ly/TiFBonderzoek
  - @VNONCW (21K)
  - @tilburgresearch (2488)
  - @succesopvolging
  - @FBNed (515)
  - @fambizznl (959)
  - @mkbnl (45K)
  - @OP_Nederland (90K)
  - @Dormaar (0.7K) - redacteur Rtl Z (familiebedrijven artikel)
  - @BV_Rendement (0.9K) - published article about research