Evaluating Selection Bias in Early-Stage Investment Returns

Katja Kisseleva Aksel Mjøs David T. Robinson*

Abstract

This paper investigates sample selection bias in early-stage investment. We use comprehensive administrative data on the universe of new firm starts in Norway, allowing us to compare venture-backed firms with ex ante similar firms that do not receive venture funding. The valuation premium for venture backing is sizeable at firm birth and doubles over the first five years, implying a substantial upward bias in VC returns relative to comparable firms. In contrast, the premium for firms receiving multiple rounds of outside equity emerges only after the first year and remains significantly smaller than the VC premium throughout the firm lifecycle.

*Katja Kisseleva, k.kisseleva@fs.de, Frankfurt School of Finance & Management. Aksel Mjøs, aksel.mjos@nhh.no, Norwegian School of Economics. David T. Robinson, davidr@duke.edu, Duke University Fuqua School of Business and NBER. We thank the anonymous reviewer, Manuel Adelino and Manju Puri for helpful comments. Participants at the 2018 Entrepreneurial Financial Management conference, 2018 Bergen Entrepreneurship and Finance Conference, 2019 Annual Private Capital Conference, 2019 KWC Conference on Entrepreneurial Finance, 2020 American Finance Association Annual Meeting, and 2020 SoCal Private Equity conference provided many helpful comments on an earlier version of this project. We thank the Norwegian tax authority for granting us data access. All errors are our own.

I. Introduction

Understanding the risk and return to investing in early-stage, high-potential, innovative firms is one of the central questions in entrepreneurial finance. Nonetheless, nearly a century after Georges Doriot's pioneering investment in Digital Equipment Corporation effectively invented the modern venture capital (VC) industry, we still lack a firm grasp of VC risk and return dynamics. This problem fundamentally stems from a lack of data.

As an illustration, take Tesla Motors. Investors in Tesla's public equity earned an average annualized return of 51 percent per year in the fourteen years since its 2010 IPO, amounting to a cumulative return of over 21,000 percent. This is, however, a poor measure of investing in the electric car industry, as attempts to create commercially viable electric cars go back as far as the 19th century (US Dept of Energy). In general, we see the innovative investments that succeed, but we are much less likely to see investments that fail, and we almost never observe firms that seek capital for innovation but cannot obtain it. Only a tiny fraction of new firm starts ultimately receive venture funding (Puri and Zarutskie (2012)). While many of the firms not matched to VC investors are likely "lifestyle" businesses (Hurst and Pugsley (2011)), many others are potentially early-stage, innovative firms that do not go on to receive venture funding presumably because investors look unfavorably on the particularly uncertain risk–return trade-off associated with investing in such firms.

The bias that arises from only observing firms that successfully receive venture funding is well-documented and has been addressed in a number of important papers

(prominent examples include Cochrane (2005), Korteweg and Sorensen (2010)). There are two key biases at play. One arises from the fact that firms receiving *any* venture investment likely grow faster than ex ante observationally similar firms that do not. This makes VC returns a biased measure of the returns to investment in innovative startups in general. The second bias comes from the fact that many VC investors make multiple investments in successful firms; firms receiving multiple rounds of investment are presumably on different growth trajectories than firms receiving only a single round of investment.[1] Prior work has focused on statistical corrections that offer indirect mechanisms for addressing these selection issues. In this paper, we approach these issues directly by using data that allow us to observe previously unobserved investment returns. Specifically, we focus not only on the returns earned by VC investors but more broadly on the returns to investing in all new potentially innovative firm "at risk" of receiving venture funding or other types of growth capital. This brings new evidence to light on the extent of sample selection bias in the reported returns of early-stage innovative firms.

To do so, we exploit a large, detailed data set provided by the Norwegian tax authority that contains tax declarations of transaction values (gross of transaction costs and any fund fees), dates and year-end holdings for the entire population of domestic and foreign investors in the population of Norwegian limited liability firms. Limited liability firms in Norway are analogous to C-corporations in the US; however, unlike their counterparts in the US, where many new firms initially organize as sole proprietorships, Norwegian business registration practices are such that our data cover essentially all new

[1]This is consistent with the firm in question not receiving multiple rounds of investment potentially because outside investors after receiving updated information no longer saw promise in the firm's future.

firm starts. These data allow us to go beyond measuring performance limited to VC investments and measure the returns that an investor wishing to invest in innovative startups may expect, in full knowledge that she must make her investments before the success or failure of a firm is known.

We begin by building an ex ante empirical measure of future innovation potential. This measure allows us to identify a set of firms observably similar to those that will later receive venture funding: firms that are essentially at risk of receiving venture funding, whether they succeed or not. To do this, we rely on the core idea in Guzman and Stern (2015, 2020) that entrepreneurs make choices at the time of firm establishment based on their ambitions and expectations that then predict actual entrepreneurial quality. We adapt this logic to our empirical setting by identifying four indicators for high innovation potential (HIP) based on firm characteristics observable at incorporation. The indicators are based on the firms' having an English-language name, being located near the largest university cities, operating in a potentially innovative industry and having at least one geographically distant board member. HIP firms possess any two of these characteristics and account for only 14% of the universe of new establishments but draw 88% of all the VC funding that flows into firms in Norway in our sample period. Our flags are also strong predictors of firm-level revenue growth, patent applications and positive later-stage firm exits through, for example, mergers and acquisitions (M&As) and IPOs. Thus, this sampling procedure allows us to study the set of firms that might appear at the "top of the funnel" for venture investors (see Gompers, Gornall, Kaplan, and Strebulaev (2020)). Using our equity transaction data, we (i) identify whether each of these firms raises equity financing at all, and if so, whether it does so in one or multiple financing rounds, (ii)

identify whether each of the firms receives venture funding, and (iii) calculate a panel of firm valuations for each of the firms based on all share purchases by any investor in either a financing round or a secondary trade.

Next, we report the distribution of firm returns across both the venture-invested sample as well as the broader sample of high innovation potential firms. This highlights the importance of unobserved market transactions for the findings of earlier studies on the returns to investments in early-stage innovative firms (see Cochrane (2005), Korteweg and Sorensen (2010), Harris, Jenkinson, and Kaplan (2014), Robinson and Sensoy (2013), Robinson and Sensoy (2016) among others). We construct a series of firm return benchmarks for different subsamples of firms and estimate three alternative measures of firm returns based on different assumptions about firms that cease to report earnings data and whether to include market valuations implied by secondary trades.

Across all our measures, investment returns demonstrate pronounced right skewness. While the median HIP firm barely returns invested capital, the average equally weighted firm ´total value to paid-in capital´ (TVPI) is around 1.6. Under strong assumptions about the failure of firms with missing data, investors in HIP firms lose on average 30% of the invested capital. Average returns for VC-backed firms are significantly higher with the average TVPI observed in financing rounds of 2.6 and slightly lower of 2.0, when accounting for failed firms. The skewness is pronounced within both samples: the 99th percentile TVPI is between three and nine times larger than the 90th percentile TVPI depending on the empirical specification.

Cochrane (2005), Korteweg and Sorensen (2010) also stress the importance of skewness in their analysis of VC risk and return. Note, however, that Cochrane (2005)

focuses primarily on obtaining estimates of the asset pricing parameters associated with the VC investment class: The selection bias in his data stems from the fact that not all initial venture investments receive follow-on funding, and the skewness that he observes is limited to that in the returns from venture investments. Cochrane (2005) does not speak to the skewness in the total underlying population of firms from which the venture investments are drawn. In contrast, the skewness that we measure also arises from the fact that many firms with high innovation potential receive outside equity funding of types other than VC. The skewness in our data is several orders of magnitude greater than what has been previously documented from venture investment data.

Relatedly, Bessembinder (2018) shows skewness in long-horizon public stock returns, showing that only a handful of publicly traded stocks account for the vast majority of the overall performance of the US stock market since 1926, and Farago and Hjalmarsson (2023) show that strong positive skewness implies that the mean compound return will serve as a poor guide for typical long-horizon investment outcomes. Our results on skewness in firm returns connect our work to a large literature on firm dynamics (e.g., Hurst and Pugsley (2011), Haltiwanger, Jarmin, and Miranda (2013), Adelino, Ma, and Robinson (2017)) and, in turn, connect this literature to a large literature in finance measuring investment performance.

In the final step of our analysis, we adapt techniques commonly applied to repeat-sales indexes for housing, art and other illiquid assets to measure the valuation premia associated with different types of firms as a function of firm age. We estimate the valuation premium of VC-backed and firms receiving multiple equity injections (multiround firms) relative to single round and non-VC-backed firms. Comparing the evolution of the

indexes over time measured as firm age allows us to gauge the magnitude of sample selection bias in commonly observed data. Our results here indicate that there is a pronounced initial difference in valuations between venture and non-venture backed firms, and that this difference grows through the first four–five years of a firm's life. While there is also a premium for firms that receive multiple rounds of investment, this premium arises not until the early years and is generally smaller in magnitude than the VC premium. The venture premium is much more volatile as firms age, indicating that older still-private firms are a mix of firms that are stagnant and struggling to generate an exit event for investors and firms that are continuing to grow.

The remainder of the paper is structured as follows. Section 2 provides institutional background and data description. Section 3 describes our sample. Section 4 sets the benchmark investment performance. Section 5 constructs startup valuation premium indexes. Section 6 concludes.

II. Institutional Background and Data

A. Entrepreneurial Environment

Although Norway is a small country in terms of population (with just over five million inhabitants), its economic structure and the relative importance of private capital markets to its economy are similar to those of other advanced Western economies. Figure 1 reports the average size of the VC sector from 2007 to 2018 scaled by population,[2] ranking

[2]Because the oil and gas sector constitutes a large share of Norwegian GDP, this scaling facilitates comparisons with other European countries.

Norway third in Europe (behind Sweden and Switzerland) and sixth globally (after the United States, Israel, and Canada). Norwegian private capital markets attract significant domestic and international investment. Between 2007 and 2023, 29% of VC and seed investments by private equity (PE) investors in Norwegian firms originated from foreign investors, while foreign investors hold roughly one-third of the share value on the Oslo Stock Exchange.[3]

[Insert Figure 1 approximately here]

Norway also performs strongly in innovation, ranking seventh in Europe on the European Innovation Scoreboard (Reid and Markianidou (2024)) and second in "crowdfunding market readiness" among thirteen European countries (Shneor, Wenzlaff, Boyko, Baah-Peprah, Odorovic, and Okhrimenko (2023)). In 2021, over 80% of 1.8 billion Norwegian kroner (NOK) in VC and seed investments was directed to IT, life sciences, and clean technology sectors. Further background on the Norwegian innovation system can be found in Hvide and Jones's (2018) Appendix V. Thus, it is reasonable to expect that lessons learned from the Norwegian innovation-related private capital markets can be generalized to other economic settings.

B. Data Sources

Norwegian administrative data are known for their high level of detail and quality and have been used prominently in research in labor economics, finance and innovation (for recent examples, see Fagereng, Mogstad, and Rønning (2021), Hvide and Meling (2022),

[3]Source: Menon Economics on behalf of the Norwegian Venture Capital Association, a special extract from the data underlying their general market report; see `https://nvca.no`.

Ring (2023), Hvide, Meling, Mogstad, and Vestad (2024)). We draw data from two sources. First, we retrieve the annual corporate information and annual financial statements from the national corporate registry, the Brønnøysund Register Centre.[4] All Norwegian limited liability firms, similarly to U.S. C-corporations must digitally register their incorporation and provide information on founders, received equity, articles of association, industry, address, CEO, board of directors, and auditor. Upon this registration, they are allocated a unique and permanent national firm identifier (*organisasjonsnummer*), which is consistently used in all firm registries and allows the data to be merged to other databases. This identification number is also allocated to foreign institutional investors. All firms are required to submit their financial statements for each calendar year by the end of July of the following year (Mjøs and Selle (2022)), as well as ad hoc notification of any corporate changes (bankruptcy filings, change in parent company, etc.).

Second, we receive corporate and individual tax records on all equity transactions covering the entire population of Norwegian firms and their investors (including foreign) from the Norwegian tax authority.[5] This information is based on annual reporting to tax authorities about share purchases, their realizations and shareholdings provided by each individual or institutional investor for wealth tax purposes. Firms submit a separate corporate annual tax report on their equity and investors. These reports have been digitally collected and stored in a data warehouse since 2004, and we obtain data through to the end of calendar year 2018. The data identify firms' investors and their shareholdings and all equity purchase, sale and liquidation transactions. All investors (in fact, all

[4]English-language website: `https://www.brreg.no/en/`

[5]English-language website: `https://www.skatteetaten.no/en/business-and-organisation/`.

individuals), including foreign individual investors, have a unique and permanent individual identifier (*løpenummer*). Finally, we merge tax records to the financial statements data, corporate registry data and firm incorporation documents. [6]

For each transaction, we have the date and amount of the purchase or realization (through either share sales or liquidation), number of shares transacted and whether a purchase is primary or secondary. Primary transactions are purchases of newly issued shares in a firm financing round, while secondary transactions are purchases of already issued shares from existing investors. Our inclusion here of secondary transactions partially addresses the stale price problem. Because secondary purchases are not driven by the firm's demand for new capital, they are free from the endogeneity problems that typically surround equity capital fundraising events. However, they may be subject to other (e.g., investor-related) endogeneity challenges. Figure 2 illustrates the data structure with a timeline of all equity transactions of two hypothetical firms, the first of which is still independently operating at the end of our sample period (12/31/2018) and the second acquired in full on 03/25/2016. We distinguish issuance of new shares at firm inception (blue) and in financing rounds (orange) and secondary trades, where shares are transferred from existing investors to buyers (pink). Shares can be realized through sale (to either a third party or the firm itself) or liquidation at different points in time, or they are held untraded at the end of our sample period. At the incorporation of the first firm, the founder injects NOK 28,000 (200,000 shares at NOK 0.14 each), while Investor 1 injects NOK 69,300 (495,000 shares at NOK 0.14 each); subsequently, we see equity capital

[6]Internet Appendix provides detailed information on the structure of the raw data, and describes how we process these raw data.

injections in primary transactions by both the founder (on 10/31/2008) and Investor 1 (on 10/12/2004, 05/12/2005 and 10/31/2008), each time with different numbers of shares and share prices. On 06/30/2009, Investor 1 sells all her shares to Investor 2 (602,000 shares at NOK 1.82 each). Both the founder and Investor 2 continue to invest in the firm through primary transactions on 02/15/2011 and 06/20/2013. The last observed transaction by the founder, 66,000 shares at NOK 0.50 each, represents the share price used for the end-of-sample valuation of the firm as a whole. The transactions in the second firm follow a similar pattern, except that we also introduce Investor 3, who makes her first investment in a primary transaction of 150,000 shares at NOK 1.03 each on 02/15/2011. In this firm, all shares are acquired in a takeover at NOK 0.75 per share on 03/25/2016. This represents an exit, and all investments are thus realized. The purchase price also provides the final firm valuation, and the firm leaves our dataset.

[Insert Figure 2 approximately here]

C. Calculating Firm Valuations

Equity valuations for publicly listed firms are easy to observe from stock prices in public capital markets. These markets feature a level of liquidity that is generally not found for shares of private, unlisted firms. Private firms transact far less frequently than listed firms, and when transactions do occur, they are connected either to an equity fundraising event or a secondary trade (a trade between two investors rather than a sale of shares by the firm to an investor). To calculate valuations of private firms, we need to rely on actual valuations observed in these arm´s-length *market* transactions, not on estimates

of net asset values as reported by PE investors. Thus, we can estimate market valuations only for firms with at least one primary or secondary market transaction.

Firms are required by law to provide a minimum equity injection at their founding (NOK 100,000 before 2012, NOK 30,000 from 2012 onward). Because the legal act of firm incorporation does not coincide with an outside fundraising event, we exclude the initial injection if it does not exceed the legally minimum required equity amount. We also exclude the founder's initial equity at the time of firm incorporation.[7] Furthermore, we drop penny equity issues with a purchase price of less than one NOK, total purchase amounts of less than NOK 10,000 (approximately USD 1,300), purchases of only one share or ownership transitions of less than 1% or more than 99%.[8] None of these purchases are likely to represent arm's-length market transactions relevant for valuing a firm's equity.

We aggregate primary and secondary purchases for each firm to the monthly level (i.e., we collapse all purchases in a firm within one calendar month to a common valuation point). Primary financing rounds are defined as calendar months when new shares are issued. Because secondary trades typically occur at different points in time from firm fundraising events, they provide additional time-series identification of the evolution of firm value. Thus, we aggregate the number of shares issued, number of shares purchased in secondary trades, amount of equity raised by the firm, and purchase amount of shares in secondary trades for each firm in each month in which they occur. This approach allows us

[7]We still consider these initially injected equity amounts as equity raised for other purposes.

[8]The USD:NOK exchange rate is roughly 1:6.6 on average.

to calculate either an average share price for that month based on primary purchases only:

$$Price_{it} = \frac{Equity\ Raised_{it}}{Shares\ Issued_{it}}, \tag{1}$$

or an average share price for that month based on all—both primary and secondary—purchases:

$$Price_{it}^{'} = \frac{Equity\ Raised_{it} + \sum Secondary\ Amount_{it}}{Shares\ Issued_{it} + \sum Shares\ Traded_{it}}. \tag{2}$$

Accordingly, for each firm and for each calendar month, we compute two alternative postmoney valuations as either:

$$Postmoney_{it} = Price_{it} \times Number\ Shares\ Outstanding_{it}, \tag{3}$$

which corresponds to the conventional calculation of dividing the total equity raised by the ownership purchased in the financing round, or

$$Postmoney_{it}^{'} = Price_{it}^{'} \times Number\ Shares\ Outstanding_{it}, \tag{4}$$

. This procedure forms an unbalanced panel of two versions of firms' postmoney valuations.

III. Sample Construction and Description

A. Population of Newly Established Firms

To construct our sample of interest, we begin by identifying all newly established limited liability firms (analogous to C-corporations in the US) incorporated between 2004 and 2018 in the firm registry data. We use industry codes to remove financial services and real estate firms. We define a parent company as a firm that owns 90% or more of another firm both at incorporation and on average across the included firm-years, and exclude newly formed subsidiaries of established firms. In case of a holding firm directly owning an operating subsidiary, we aggregate these two firms into one entity. Furthermore, we exclude firms with missing equity transaction data. Finally, we drop firms that go public in their first year of existence, which would imply that these are not newly established firms. Table 1 Panel A shows that of the remaining 124,348 newly formed firms, a total of 928 (0.75%) are backed by a professional VC investor. The percentage of VC-backed companies resembles the number in Puri and Zarutskie (2012), which further confirms the external validity of the findings in the Norwegian context. Among the firm population, 37% of the firms receive at least one equity financing round, and 20% of the firms have shares that are traded in at least one secondary trade. Note, however, that more than the half of the population of firms have no equity transaction event beyond the legally required minimum equity provision at the time of incorporation.

[Insert Table 1 approximately here]

B. Identifying Firms with High Innovation Potential

Of course, it is unlikely that most of the newly established 124,348 firms have growth aspirations or intend to develop large-scale commercial innovation. As Hurst and Pugsley (2011) shows, most small business owners (in the U.S.) have no desire to grow, operating their businesses primarily for lifestyle purposes. To identify firms with high innovation potential, we draw on the entrepreneurial quality index developed by Guzman and Stern (2015). Guzman and Stern (2015, 2020) start by recognizing that a practical first step for any growth-oriented entrepreneur in the US is to register her business in the state in which she operates: This facilitates paying payroll taxes, unemployment insurance, etc. Incorporated businesses are significantly more likely to grow than non-incorporated ones. To adapt these insights to the Norwegian business context, we develop a series of flags or indicators that signal the likely intention on the part of the founder(s) at the time of incorporation to grow the firm. These flags are observed when the firm first appears in our data. Although any one flag may have false negatives and false positives associated with it, by developing a series of flags to be applied together, we aim to create a robust measure of high innovation potential based on ex ante observable characteristics. The population counts of firms satisfying the criteria for these flags are reported in Table 1 Panels B and C.

The first flag denotes whether the firm has an English-language firm name. A total of 35,200 firms, or 28.3% of the starting sample, satisfy this criterion. The idea behind this flag is that because Norway is a small economy, an English-language firm name helps the firm be recognizable to a broader, international audience and therefore could signal that the entrepreneur intends to grow her firm internationally. While giving the firm an

English-language name would not necessarily confer a natural advantage if the firm's objectives were to serve only the local market, doing so would be a natural choice for entrepreneurs with more ambitious objectives, especially in Northern Europe, where English is commonly spoken as a second language. We apply several approaches to detect English firm names. These include detection by artificial intelligence with the support of ENIN, see `https://www.enin.ai/`, assessment of the language by Google Translate and manual checks throughout the entire pool of firms to flag discrepancies and to double-check and exclude firms with names that include only generic English-language words (e.g., *holding*, *shipping*, *group*, *consulting*). We also classify neologisms and artificial English-sounding names as satisfying this flag. Examples of firm with this flag include: *AKA Telecom*, *Theta Development*, *Anzym*, *Girlcode*, and *Pixab*, while firms with Norwegian names that do not satisfy the flag criterion include *Juristbyrået*, *Felger og Dekk*, *Haugen Utleie*, *Trygve Sivertsen*, and *Sjøhav*.

The second flag denotes whether the firm is located in a regional innovation hub in Norway. These are Oslo, Bergen, Stavanger and Trondheim—the country's four largest cities, each of which is host to a major research university with an associated technology cluster (Hvide and Jones (2018)). The idea here is to construct a geographical flag of startup location choice that would correspond to a firm's starting in Silicon Valley, Route 128, Austin (Texas), or the Research Triangle Park area in the United States. A total of 34,217 firms, or 27.5% of the starting sample, are started in one of these innovation hubs during our sample period.

Of course, Silicon Valley contains pizza parlors and barber shops too. Similarly, many firms founded in these innovation clusters are not likely to be innovative. The third

flag tracks whether the firm operates in an innovative industry. We obtain the firm's 5-digit industry code listed at the time of its incorporation from Statistics Norway for the entire population of firms. Using negative selection, we rule out industries with assumed low innovation potential— those that are heavily regulated (e.g., banks, insurers, and oil and gas producers), have high levels of public-sector involvement or ownership (e.g., hospitals, theaters and schools), are heavily supported via taxes and/or subsidies (e.g., agriculture and shipping), or are highly unlikely to engage in value-creating innovative growth projects (e.g., hairdressers, plumbers and trading agents). For such industries, we expect nonfinancial objectives such as government policies or a founder's wish to become an entrepreneur to be especially important, at the expense of significant value-creating innovation and growth. We manually double-check Statistics Norway's list of all possible industry classifications based on both its 2002 and 2007 classifications [9] and apply negative selection to rule out firms in such industries found in either of the classifications.[10] After these exclusions, 63.7% (over 79,000) of the firms in the starting sample remain and receive this industry flag.

The final flag tracks whether at least one of the firm's nonexecutive board members lives far from the city in which the firm resides in the year of firm incorporation. For this,

[9] See `https://www.ssb.no/en/klass/klassifikasjoner/6` for details.

[10] The specific excluded industries are the following: *agents/traders*, *agriculture*, *banks*, *brokers*, *cultural event producers*, *direct health services*, *education*, *fisheries*, *food production*, *gym/sports facilities*, *hotels*, *insurers*, *investment management*, *kindergartens*, *garages*, *mail order*, *mining*, *museums*, *oil and gas production*, *physical shops*, *public services*, *publishing*, *real estate*, *restaurants*, *shipping firms*, *wholesale traders*, and *direct services* (e.g., *hairdressers*, *for tourists*, *car rental*, *lawyers*, *maintenance*, *accountants*, *auditors*, *builders*, *plumbers*, *electricians*, *undertakers*, *taxis*).

we use a ZIP-code concordance and define "far" as a ZIP-code difference of 1,500 digits between the firm's address at initial registration and the board member's address at the same time. This implies an average beeline distance of more than 300 kilometers. Far fewer firms (20,011 or 16.1% of all firms) satisfy the criterion for this flag. The idea here is that the choice of a distantly located board member *in the year of establishment* is a potential indication that the founder(s) (or investor(s)) have recruited a board member with specific technical or market expertise not readily available in their immediate local network.

Comparing across the four flags, we find that VC investors almost solely invest in potentially innovative industries (92.8% of all VC-backed firms), and that less than half of these firms have board members who live far away. Furthermore, more than 40% of the VC-backed firms neither have an English name nor are located in the innovation hubs, indicating that VC investors select firms based on a wider information set. These lower shares also show that there may well be other high innovation potential firms that are overlooked by VC investors.

In some cases, these flags may overlap, while in others, the presence of one flag could make the presence of another unlikely. For example, a firm founded in an innovation hub may not need to recruit a geographically distant board member for technical expertise. To remain agnostic about which of the flags is more or less salient in a particular setting, we define a firm as potentially innovative if we apply any two flags to it. This yields a HIP sample of 51,243 firms or 41.2% of the population of newly established operating firms.

Our sample of HIP firms contains 88% of all VC-backed firms in our data, despite the fact that selection by a VC investor is not a criterion for inclusion in our sample. Note, however, that the fraction of firms receiving any equity capital in the HIP sample is nearly

identical to that of the broader sample–roughly 40%. Thus, around 60% of all high innovation potential firms never raise an equity financing round beyond the legal amount minimum required to start the business. While some of these businesses may obtain debt financing, or may generate operating cash flows sufficient to fund the business, the fact remains that many high innovation potential businesses never succeed at raising any outside funding.

To demonstrate the power of our flags to predict later-stage outcomes, Table 2 relates a series of firm outcomes to the presence of these flags, both individually and collectively. Panel A focuses on future financing events. In particular, this panel shows that each of these flags, either alone or when combined, is highly predictive of a firm's receiving VC investment or an innovation-related governmental grant.

[Insert Table 2 approximately here]

Panel B focuses on future milestones related to firm growth and innovation. The first part of Panel B focuses on patents.[11] In particular, not only firms with English names and firms operating in potentially innovative industries but also firms with a geographically distant board member are much likelier to apply for a patent at some point than firms not designated with any of these flags. When all three flags are combined, the location-related flag is no longer significant, but the R^2 increases to 9.0%. All the flags are highly predictive of the firm's achieving an exit through an IPO, merger or acquisition, as can be seen in the middle portion of Panel B. Last, the far-right portion of Panel B shows that these flags predict four-year revenue growth. The latter outcome also implicitly measures firm survival.

Another way to gauge the salience of these innovation flags is to look at capital

[11] We are grateful to Jorge Guzman for suggesting this outcome.

flows into and out of firms flagged as HIP firms under our procedure and compare them to those of the overall (relevant) firm population, as defined in Table 1, Panel A. This angle is especially important if we want to derive market valuations of these firms. Table 3 shows the amounts of equity capital invested by investors in financing rounds or secondary trades of the shares of all sample firms before their exit events, as well as the amounts received by investors from the sale of their (earlier purchased) shares to other investors or amounts received from the firm due to share liquidations.[12] This offers a marketwide, macro-level overview of the capital that innovative firms garner relative to that drawn by other firms.

[Insert Table 3 approximately here]

Our HIP sample receives 87% of the total equity capital invested in all newly established businesses in Norway in our sample period. The 813 VC-backed firms with at least two ex ante innovation flags comprise only 0.7% of the population but garner 15% of the equity capital raised. The latter firms represent an even larger share of the volume of secondary purchase transactions. The vast majority of the total capital paid out through share sales or share liquidations corresponds to the firms in our HIP sample. These statistics provide further evidence that our selection on ex ante flags captures firms with a high chance of raising significant funding to support their investments to grow.

[12]Liquidation means that a firm is closed down; usually due to the firm's no longer going to operate, or if the shareholders want to end their cooperation. In this case all assets are sold, debts are paid, and the remaining values are distributed to the shareholders. In a partial liquidation only part of the firm's assets are sold and distributed to the shareholders, while the rest of the firm continues to exist and operate. Share liquidations exclude bankruptcies which almost never result in any payback to shareholders. See Internet Appendix on how alternative forms of share realizations are recorded in our dataset.

C. Defining Exit Events

To define firm outcomes, we rely on the information in the business registry, which covers all firms regardless their current status or equity transactions. Our exit categories include *independently operating*, *merger*, *(full) acquisition*, *IPO*, *bankruptcy filing*, *(partial) liquidation*, *disappearance* and *zombie*.

A firm is still independently operating if it has not officially filed for bankruptcy, has neither (partially) liquidated nor merged nor been (fully) acquired, has not gone public, still submits its financial accounts to the business registry as of 2018 and has been generating revenues in the most recent five years (2014—2018). A firm is classified as a zombie if it meets all of those criteria except that it has generated no revenues in the most recent five years (2014–2018). A firm is instead classified as disappeared if it satisfies those criteria but has stopped submitting its financial accounts to the business registry before the end of our sample period. IPO, merger and bankruptcy filings are identified as such in the business registry. We identify full acquisition through the existence of a parent company with an ownership share of >90% after its incorporation in the business registry. We supplement this categorization by identifying full or partial liquidations of shares in the tax records on equity transactions. We define a firm as (partially) liquidated if we observe at least one share liquidation transaction as described above. In the case of a partial liquidation, the firm continues to exist, while the assets have been partially liquidated.

Table 4 shows the distribution of exit events for firms for which we observe at least one market-based firm equity valuation, including an implicit valuation of zero for failed

firms.[13] Specifically, to account for failures in our valuation indices we create an additional (implicit) financing round and assign a value of zero to this implicitly "observed" valuation at the time a firm files for bankruptcy, is (partially) liquidated or disappears, and on 12/31/2018 for zombie firms.[14]

[Insert Table 4 approximately here]

More than a quarter of the VC-backed firms in our HIP sample experience an exit event (IPO, merger or full acquisition), compared to 12% of all HIP firms. Only 8.5% of the firms with less than two flags (non-HIP firms) have such positive exits, indicating that they are of a different type from those selected into our HIP sample. These non-HIP firms do not go public, while the HIP firms do, with the subsample of VC-backed firms showing the highest fraction of firms that eventually have an IPO (2.6%). The VC-backed firms are also more likely to have other positive exit events such as an acquisition (17.2%) or merger (7.4%), than all HIP firms (8.7% of which are acquired and 3.5% of them merged). This discrepancy is indicative of one source of selection bias, as pointed out in the previous literature, in that VCs invest with a planned positive market exit strategy. We observe the opposite pattern with regard to negative firm outcomes. In total, only a quarter of VC-backed firms go bankrupt, are liquidated, disappear, or become zombies, while more than half of all HIP and non-HIP firms do (53.7% of all HIP and 53.5% of non-HIP firms).

[13]The remainder of our analyses of valuations necessarily only include firms for which we observe at least one market valuation, thus, we only describe exit events for the same subset of firms.

[14]We do not count this implicit financing round into the classification as a multiround firm.

D. Description of Financing Rounds and Secondary Trades

Table 5 reports descriptive statistics of equity raising events and secondary trades for both HIP and non-HIP firms, covering all sample firms with at least one capital market equity transaction. Among HIP firms, 38% raise equity, averaging 1.6 times (median 1.0), slightly higher than the 36% of non-HIP firms, which raise equity an average of 1.3 times (median 1.0). VC-backed firms in our HIP sample raise equity far more frequently (94%) with an average of four times (median 3.0). Although generally the first equity raising occurs when the firm is, on average, 1.3 (VC-backed 1.5) years old (median 1.0), the equity-raising activity varies by firm age across different types of firms. Nearly all HIP firms (84%), and an even larger proportion of non-HIP firms (88%), raise equity before reaching two years of age, with the frequency of financing events decreasing notably as firms mature (or disappear). In contrast, almost all VC-backed firms raise equity when they are between two and three years old, though both younger and older firms in this subsample also frequently engage in equity raising.

The non-founder (outside) ownership dilution decreases over time in financing events; however, it remains fairly stable across our subsamples within each firm age cohort. The main difference between the types of firms is the amount of non-founder (outside) equity raised. VC-backed firms raise the largest amounts of non-founder equity across all age cohorts. For example, VC-backed firms aged two to three years raise an average of NOK 24.3 million (median NOK 4.3 million), whereas the full sample of HIP firms of the same age raises a significantly lower average of NOK 10.6 million (median NOK 0.8 million). By comparison, non-HIP firms of the same age raise the smallest amounts from

outside investors, averaging NOK 2.9 million (median NOK 0.2 million). These differences result in significant divergence in postmoney valuations across firm types. VC-backed firms are valued at an average of NOK 46.4 million (median NOK 2.7 million) when younger than two years, rising to NOK 87.7 million (median NOK 39.7 million) when older than six years. In contrast, valuations for the full sample of HIP firms range from NOK 5.9 million (median NOK 0.1 million) to NOK 50.4 million (median NOK 8.7 million), while non-HIP firm valuations range from NOK 1.1 million (median NOK 0.1 million) to NOK 12.8 million (median NOK 1.4 million). These patterns illustrate the selection bias favoring larger equity amounts and higher valuations in VC-backed firms. It also illustrates potential survivorship bias, as about six times the share of the VC-backed firms raise equity when they are over six years old compared to for other firm types. Consistent with prior literature on the U.S. market, the decision to repeatedly raise equity—which is endogenous to firm development and financing demand or supply (see Cochrane (2005), Korteweg and Sorensen (2010))—occurs more frequently in VC-backed firms than in other innovative firms.

[Insert Table 5 approximately here]

Table 5 also provides data on secondary trades. While already issued shares are traded among existing (and potentially new) investors in 58% of VC-backed firms, this figure drops to 23% for all HIP firms and 18% for non-HIP firms. The average pooled firm valuations in secondary trades follow a similar pattern across the firm subsamples to postmoney valuations in equity-raising events. The average aggregated amounts traded, in the months with secondary trading, are NOK 7.2 million in VC-backed firms (median NOK 1.0 million) compared to NOK 1.8 million (median NOK 0.1 million) for all HIP firms and NOK 0.7 million (median NOK 0.1 million) for non-HIP firms. These market transactions

reveal a notable selection bias, with VC-backed firms achieving more successful exits, raising larger amounts of equity, engaging in more secondary trades, and attaining higher average valuations in all market transactions compared to both the full sample of HIP firms, and in particular to the non-HIP firms.

IV. Benchmarking Investment Performance

To illustrate the importance of the lack of data on unobserved market transactions in earlier studies (e.g., Cochrane (2005), Korteweg and Sorensen (2010), Harris et al. (2014), Robinson and Sensoy (2013), Robinson and Sensoy (2016) among others) for our understanding of the returns to investing in early-stage innovative firms, we begin by presenting the distribution of investment returns using the 'total value to paid-in capital' (TVPI) measure, which is commonly used in private markets research. We measure this at the firm level by taking the last observed firm valuation and dividing by the sum of all invested equity in the firm up to that point. Formally, the TVPI_{it} for firm i at time t is defined as

$$TVPI_i = \frac{Postmoney_{it}}{\sum_{t=0}^{t} Equity\ Raised_{it}}, \tag{5}$$

where POSTMONEY_{it} is the last observable market valuation of firm i in our sample period at time t and $\sum_{t=0}^{t} \text{EQUITY\_ RAISED}_{it}$ is the total equity injected in firm i from

inception at $t = 0$ until the latest observable firm valuation at time t by any investor.[15] Table 6 reports this measure computed for different subsamples and in a variety of ways based on different assumptions about unobserved firm deaths. The main measure, "Observed TVPI", is simply based on the observed postmoney valuation in the latest financing round at time t. To account for "silent liquidations", we add in firms that filed for bankruptcy, were liquidated or disappeared, or were zombie firms, assigning them a terminal TVPI of 0. This is referred to as "Add Implied Failures." Finally, we use the $\text{POSTMONEY}'_{it}$ valuation, which also takes into account the firm market valuation implicit in the last observed secondary trade of the firm. This is labeled "Add Secondary Trades." These are computed both on an equally weighted basis and weighted by the total invested capital at the time of the last valuation.

[Insert Table 6 approximately here]

Although the number of included VC-backed firms increases only slightly when we report broader TVPI measures, the sample of HIP firms is dramatically affected by accounting for failures and zombie firms, as well as by including valuations in secondary trades. The sample of HIP firms increases by 55% when we include failed firms and by another 15% when we also allow for TVPIs based on valuations implicit in secondary trades. This comparison offers an additional perspective on the sample selection bias associated with estimation of the returns to early-stage investments from publicly observable data, which are typically returns to VC investments that fund managers have chosen to publicize. This

[15] It also includes the founder's initial equity and any other legally required initial capital, which we exclude from the calculation of postmoney valuations as described in Section 2.3.

suggests that the broader population of potentially innovative firms experiences a much higher failure rate than VC-backed firms.

The average equally weighted (capital-weighted)[16] observed TVPI is 2.62 (1.60) in the VC-backed firms in our HIP sample.[17] The distribution of returns in VC-backed firms is heavily right skewed, and after we winsorize the full TVPI distribution at the 99.5th percentile, the average decreases to a TVPI of 2.37 (1.58), driven by replacing the extreme equally weighted returns in the right tail, i.e. higher than the reported 99th percentile of 23.91, equally weighted. The median firm valuation exceeds the total equity raised by 31% on the equally weighted basis or is just as much as the total equity raised on the capital-weighted basis.

Once we add implied failures, the equally weighted (capital-weighted) average TVPI of 1.96 (0.77) is significantly lower than the firm TVPIs based only on the valuations in financing rounds. The winsorized average TVPI is 1.72 (0.74). Failing firms are usually unobservable and are typically excluded from estimations of return distributions.

The most comprehensive return measure adds valuations implicit in secondary

[16]We present the capital-weighted distribution, which uses the denominator in equation (5), $\sum_{t=0}^{t}$ EQUITY_RAISED$_{it}$ as a weight on TVPIs.

[17]There are two important differences between the returns in our sample and those reported in Harris, Jenkinson, Kaplan, and Stucke (2023), Robinson and Sensoy (2016), Harris et al. (2014). One is that we measure total returns over the observed life of each firm, while VC fund returns reflect the aggregated investment returns in various firms only during part of the firms' life. Our invested capital is also the total equity raised by the firm from all investors, rather than the amounts the VC funds actually have paid for their shares. Second, the returns reported in previous literature are net of manager, transaction and other fees borne by the investor, while we report gross returns at firm level.

trades, and the equally weighted (capital-weighted) average TVPI is 4.72 (0.80). The winsorized average TVPI is 2.99 (0.80), significantly lower on an equally weighted basis, reflecting that the distribution of TVPI based on secondary trades exhibits even higher levels of skewness. The right tails of these TVPI distributions weighted by total capital invested are significantly lower than those of the equally weighted distributions, reflecting that lower-performing firms receive a relatively larger share of equity invested in VC-backed firms.[18] The returns to investing in the entire sample of HIP firms are significantly lower (equally weighted) compared to the returns to VC-backed firms up to the 95th percentile, equally weighted. In the very right tail of the distribution, the full sample of HIP firms includes some significantly higher return observations compared to those in the VC-backed subsample. These findings are indicative of a potential for overlooked opportunities outside traditional VC pipelines. Comparing the capital-weighted return distributions indicates that VC investors invest relatively less capital in the lower performing firms.

Multiround firms—firms that raise more than one financing round, regardless of whether it is from VC- or other investors–generate higher average equally weighted TVPIs compared to the overall sample of HIP firms, even after including implied failures. However, when including secondary trades, the full sample of HIP firms achieves lower average TVPI, even if the skewness causes significantly higher returns from the 95th percentile onwards, equally weighted. The capital-weighted average TVPIs are lower for multiround firms than for the full sample of HIP firms after accounting for implied failures, indicating that larger equity investments are associated with worse-performing firms.

[18]Including VC-backed firms that are not captured by our HIP flags has no impact on the returns distributions.

TVPIs are also less skewed for multiround firms, irrespective of any weighting. Within the subsample of multiround firms, VC-backed firms deliver higher returns, except in the right tail of the equally weighted returns including valuations from secondary trades. This illustrates that the likelihood of some higher performing firms that are not selected by VC investors are also present among multiround firms, as shown above for the full sample.

Overall, this descriptive table provides first evidence on the extent of upward misrepresentation of the returns to investing in early-stage innovative firms when only VC-backed firms and/or firms that raise multiple rounds of financing are studied. The returns are even more highly skewed than recognized in previous research, and there is potential for overlooked opportunities outside traditional VC pipelines. We also find evidence that relatively more equity is invested in worse-performing firms, a tendency that further reduces the combined returns from the whole innovation economy considered as an overall portfolio of HIP firms.

V. Startup Valuation Premium Index

A. Quantifying the Selection Bias

Table 6 offers a cross-sectional picture of total firm return multiples, measured as of the final observed valuation. In this section, we approach the venture premium from a time-series perspective by constructing a valuation index that varies over time as a function of the firm's age and other factors.

We borrow tools and insights from the literature that constructs a modern, hybrid

version of the repeat-sales technique for durable assets such as art or housing, starting with Bailey, Muth, and Nourse (1963), Case and Shiller (1989), Adelino, McCartney, and Schoar (2020). In the standard housing market setting, a repeat-sales index is constructed on the basis of multiple valuations from sales of the same residential property. These valuations are regressed on monthly time dummy variables. The coefficients on these time dummy variables then constitute a calendar time series (index) of property values.

While the standard repeat-sale index is primarily focused on the evolution of an asset value over calendar time, our aim is different. We are interested in understanding how asset values evolve over a firm's life, from birth, through subsequent funding rounds, to some potential exit event. Thus, our use of the index is different than the standard approach.

Specifically, we estimate the following regression equation:

$$
\begin{aligned}
X_{it} = \alpha + \sum \beta_{1t} Age_t + \beta_2 VC_{it} + \sum \beta_{3t} Age_t \times VC_{it} \\
+ \beta_4 Multiround_{it} + \sum \beta_{5t} Age_t \times Multiround_{it} + \epsilon_{it}.
\end{aligned} \tag{6}
$$

X is the firm i's actual postmoney valuation (in NOK million) at age t (measured in calendar months), winsorized at the 99.5th percentile.[19] The postmoney valuations are derived from valuations in primary financing rounds, not including secondary transactions. (Including them has no effect on the findings; see Appendix figures.) We do not consider the initial equity injection at firm establishment to be a financing round if it does not

[19]Winsorization of the top 0.5% observations in the full valuation distribution is sufficient to robustly mitigate the influence of a few outliers.

exceed the legally required equity amount. To account for firm failures in our valuation premium indexes we create an additional (implicit) final financing round and assign a value of zero to this implicitly "observed" valuation at the time a firm files for bankruptcy, is (partially) liquidated or disappears, and on 12/31/2018 for zombie firms. (This implicit financing round does not count towards being classified as a multiround firm.)

To disentangle the VC effect from the multiround effect, we first define VC-backed and multiround firm categories as mutually exclusive. VC_{it} takes the value one when a firm first receives VC-backing, and remains one thereafter, including when it potentially receives follow-on investing of any type. The variable MULTIROUND_{it} takes the value one for non-VC-backed firms when they raise a second financing round. Thus, we do not involve any look-ahead bias. The VC category takes precedence over multiround status. So if a firm is both VC-backed and multiround in a given year, it is coded as VC-backed.

The set of point estimates β_{3t} from the interaction effect between firm age and the VC-backed dummy variable plus β_2 on the latter constitutes the startup valuation premium due to firm selection by venture investors. The set of point estimates β_{5t} from the interaction effect between firm age and the multiround dummy variable plus β_4 on the latter constitutes the startup valuation premium due to the ability to raise several rounds of non-VC-equity financing.

The standard application of repeat-sales methodology would be to recover the set of β_{1t} and use them to generate a measure of how the average postmoney valuation evolves over time. Instead, we make use of the fact that $\beta_2 + \beta_{3t}$ (or $\beta_4 + \beta_{5t}$) is a measure of the difference in the average valuation of a firm at age t of one type versus the other. Plotting

these sums yields a measure of how the type-specific premium evolves as a function of firm age. All estimations are based on the sample of HIP firms.

The standard identifying assumption in the traditional setting is that unobserved asset characteristics do not change between two sales of that property and that price changes over time are driven by a common, market-wide factor, not changes in the fundamental nature of the asset itself. Our approach does not require this assumption because we are not attempting to establish the "market value" of a unit of early-stage innovation or track its change over time; instead, we are interested in understanding how the quantity of innovation, in market value terms, evolves over time in different samples.[20]

Figure 3 provides a visual description of the sample selection bias and presents estimates of the VC valuation premium and multiround valuation premium over non-VC-backed single-round firms, using equation (6). The average VC premium over non-VC-backed, single-round firms is approximately NOK 74 million, while the average multiround premium is approximately NOK 30 million. This results in a VC selection bias of 2.5 times the multiround bias. The upward slope of both lines through age four is consistent with the average firm becoming relatively more valuable over time if it is

[20]Hwang, Quigley, and Woodward (2005), Peng (2001) attempt to overcome the selection bias and to construct an index of VC valuations. Our analyses differ from theirs on three dimensions. First, we account for not only the selection bias arising from the observability of financing rounds or the lack thereof in the firms selected into VC financing but also the bias due to firm selection into VC financing itself. Second, similar to the real estate literature, they create an index of value for VC investments to measure movements in the value of VC-backed firms over time. Third, we create an index by firm age, not calendar time. For an extensive discussion of the applicability of the statistical properties of this method to (early-stage) firms, please refer to Hwang et al. (2005).

VC-backed or otherwise receives multiple rounds of funding relative to a non-VC-backed firm with only a single round of investment. This in turn is consistent with the idea that these firms are growing relative to excluded firms, at least through age four–five.[21]

[Insert Figure 3 approximately here]

Appendix Figure A1 replicates Figure 3 but the postmoney valuations are based on all observable trades in both financing rounds and secondary trades. The results remain robust with the average VC premium over non-VC-backed, single-round firms of approximately NOK 77 million, and the average multiround premium of approximately NOK 31 million (thus, the VC selection bias 2.5 larger than the multiround bias).

B. Understanding the Venture Premium

Figure 4 explores the venture premium. To do so, we estimate equations of the following form:

$$(7) \quad X_{it} = \alpha + \sum \beta_{1t} Age_t + \beta_2 \mathrm{VC_i} + \sum \beta_{3t} Age_t \times \mathrm{VC_i} + \epsilon_{it}$$

$$(8) \quad X_{ijt} = \alpha + \gamma_j + \sum \beta_{1t} Age_t + \beta_2 \mathrm{VC_i} + \sum \beta_{3t} Age_t \times \mathrm{VC_i} + \epsilon_{ijt}$$

$$(9) \quad X_{ijst} = \alpha + \gamma_j + \kappa_s + \sum \beta_{1t} Age_t + \beta_2 \mathrm{VC_i} + \sum \beta_{3t} Age_t \times \mathrm{VC_i} + \epsilon_{ijst}$$

$$(10) \quad X_{ijst} = \alpha + \gamma_j + \kappa_s + \sum \beta_{1t} Age_t + \beta_2 \mathrm{VC_i} + \sum \beta_{3t} Age_t \times \mathrm{VC_i} + \beta_4 \mathrm{TIC_i} + \epsilon_{ijst}.$$

[21] The winsorization of the top 1% of the full valuation distribution (untabulated) consistently shows the VC premium bias of 2.5 times the multiround premium, with the average VC premium over non-VC-backed, single-round firms of approximately NOK 64 million, and the average multiround premium of approximately NOK 25 million.

where X is the firm's i postmoney valuation at age t, γ_j are industry fixed effects, κ_s are calendar-month fixed effects, VC_i is an indicator for whether a firm receives venture funding at some point in time (and regardless of whether it is going to become a multiround firm or not), and TIC is the sum total of invested capital raised up to the current date. Both X and TIC are winsorized at the winsorized at the 99.5th percentile.

Figure 4 plots the VC valuation premium as a function of the firm's age. In contrast to Section 5.1, where our objective is to sharply disentangle VC premium from multiround premium, the VC dummy estimated here is a static, time-invariant, flag defined at the firm level, meaning that it equals one for all points in time for a firm that at some point receives VC financing. It is not a flag that indicates the point at which a firm receives VC financing. This means that Figure 4 is not depicting the valuation change that occurs when a firm gets financing in year t; it is instead the year t valuation difference between firms that at some point will receive VC financing relative to those that never will. For example, even if a firm receives a first VC investment at age three, its postmoney valuation (if observed) at ages one and two is used to estimate the β_{3t} parameter for those time points. This procedure allows us to assess whether firms that eventually receive VC funding rounds are distinct from others from the outset.

[Insert Figure 4 approximately here]

Figure 4 illustrates that at firm birth, a firm that will at some point be venture-backed is valued at approximately NOK 40 million higher compared to what a non-VC-backed firm is initially worth. To be included in the sample at that point in time it needs to have an observed postmoney valuation, so this effectively measures the difference in initial value between firms that receive VC funding at birth and those that do not. This valuation

premium varies as a function of firm age, but in general averages around NOK 60 million. The increased volatility as a function of firm age presumably reflects the fact that older-aged firms with observed valuations are a mix of firms continuing to grow and firms that are stagnant but nonetheless receive capital from outside investors.

As we shift from plotting estimates obtained from equation (7) to equations (8)–(10), the resulting coefficients are purged of the confounding effects of the included variables. Incorporating industry fixed effects has virtually no effect on the estimated venture premium over time. Similarly, incorporating calendar-month fixed effects (controlling for the fact that some age t valuation events occurred in January of 2016 while others occurred in February of 2009, for example) has only a modest downward effect, particularly in the early years, on the observed valuation premia.

Because more equity is being invested into a firm at each financing event, we include total amount of equity raised up to the time of the respective valuation of the firm in equation (10). When this is omitted, we confound the evolution of the underlying firm values with the injection of additional capital flowing into the firms as they develop over time. The plot of coefficients from equation (10) illustrates that the venture premium is for the most part a premium associated with raising large amounts of outside capital. The premium associated with being venture-backed substantially decreases when we control for the total capital raised.[22]

Interpreting this version of venture premium requires caution. Table 5 shows that VC-backed firms raise significantly larger amounts of equity capital compared to other

[22]See Appendix Figure A3 for valuation premium when the postmoney valuations are derived both from valuations in financing rounds and in secondary trades, $\text{POSTMONEY}'_{it}$. The results are unchanged.

HIP-firms, particularly in the first financing rounds. Both median and average amounts of funding raised for venture-backed firms aged 0–1 is ten times that of similarly aged full sample of HIP firms. Effectively, venture backing is selecting on firms with high demand for outside capital, which is consistent with higher overall valuations, reflecting the firms' higher growth potential.

C. Understanding the Multiround Premium

Figure 5 estimates the multiround valuation premium using equations (7)–(10), replacing VC_i with a (static, time-invariant) dummy variable MR_i taking value one if a firm i is a multiround firm. A firm is flagged as multiround if it ever receives more than one round of financing, regardless of its source (not including a final implied value of zero for zombie firms). Multiround firms and VC-backed firms are overlapping firm categories in Figures 4 and 5, so that multiround firms in Figure 5 may include VC-backed firms and vice versa.

[Insert Figure 5 approximately here]

We observe a different pattern for the multiround valuation premium. The initial multiround premium at firm incorporation or, potentially, even before the official registration hovers around zero, indicating that firms that go on to raise multiple financing rounds at some point do not appear to be valued differently at the time of incorporation than firms that receive financing only once. However, the multiround premium begins to emerge after age one and grows steadily with firm age. The estimates also exhibit less volatility than those for VC-backed firms, particularly at older firm ages. Once we control for the total equity raised, the valuation premium of multiround firms is largely attenuated

until the firm age of six. Beyond age six, the total equity–adjusted valuation premium becomes negative compared to single-round firms, consistent with investors continuing to contribute capital also to support firms' survival and not only to achieve superior returns.[23] As illustrated in Table 7, multiround firms raise proportionally larger *total* amounts of equity capital compared to single-round firms over our sample period, particularly, if they are not VC-backed. Our results are consistent with multiround firms being valued at a premium compared to single-round firms, but this premium being explained by capital raised.

Appendix Figure A4 replicates Figures 4 (Panel A) and 5 (Panel B) for a subsample of firms that have survived for at least ten years. The valuation premium is less pronounced than when using the whole sample, with the possible exception that we see a stable positive premium for the oldest multiround firm cohorts.

Appendix Figure A5 examines the robustness of our sample selection of HIP firms and replicates Figures 4 (Panel A) and 5 (Panel B) for varying definitions of HIP firms. All graphs show the respective valuation premium indexes, estimated using postmoney valuations in financing rounds, POSTMONEY_{it}, and the fully specified model in equation (10). The blue graph represents our sample of HIP firms, which satisfy the criteria for any two ex ante innovation flags at the time of firm incorporation. The red graph represents firms satisfying the criterion for at least one ex ante flag, while the green graph represent firms operating in an innovative industry and having at least one further ex ante flag. The yellow graph represents firms having any three ex ante innovation flags. The resulting

[23]See Appendix Figure A4 for valuation premium when the postmoney valuations are derived both from valuations in financing rounds and in secondary trades, $\text{POSTMONEY}'_{it}$. The results are unchanged.

valuation premium indexes are very similar to those using our standard definition of HIP firms, except that they become increasingly volatile for the more restrictive definitions of HIP firms, necessarily leading to fewer firms. These similarities are sustained over the observed lifetime of our sample firms, thus, our ex ante flag selection approach is robust to the use of different numbers of flags.

D. Digging Deeper Into Fundraising

Because our estimation of the venture and multiround premia, on average and over time, is heavily influenced by whether we control for the total amount of prior funding raised, a natural final step in our analysis is to examine the sources and structure of equity financing for VC and non-VC backed firms. This allows us to hone the interpretation of the indexes by studying the sources of capital being supplied to VC- and non-VC-backed firms.

Table 7 documents the average equity ownership structure (the "capitalization table" or "cap table") and average equity investment amount, conditional on the specific investor type making an investment. In addition to VC investors, we differentiate between individual investors—founders, their family members, more sophisticated (professional) repeat angels and other individuals—and other institutional investors, such as financial institutions and other nonfinancial corporations.

Multiround firms—particularly those backed by VCs—raise significantly more capital in total, on average, than their single-round counterparts. Founders in VC-backed firms contribute more capital per investment in single-round firms than in multiround firms, but retain a similar level of ownership. In contrast, founders in non-VC-backed firms

retain much larger ownership shares—46.3% in single-round and 23.9% in multiround cases—despite investing similarly modest amounts. These patterns are consistent with the idea that VC-backed founders face more consistent dilution, but benefit from larger absolute capital inflows and potentially higher valuations. It further follows that the effects of the selection and treatment by VC investors on firm valuations dominate any potential negative incentive effects of the reduction in ownership for founders (Sorensen (2007), Hsu (2004)). Other individual investors also invest more when the firm is VC-backed, consistent with the idea that VC involvement serves a certification function that encourages greater participation. Ownership shares for these investor types are typically lower in VC-backed firms, reflecting larger round sizes and more competition for equity. In VC investors' absence, investors of other types purchase higher ownership stakes on average in any financing event of non-VC-backed firms, as the round sizes are also significantly smaller, and vice versa in firms also matched to VC investors.

VC investors specialize in investing in portfolios of early-stage firms and are likely to understand these firms better before selecting them and to be more able to support their growth. Matching with VC investors provides "certification" which represents a positive externality for other investors leading to larger investments. This larger investor interest in investing in VC-backed firms is also consistent with such firms having higher valuations, as shown earlier. Top VC investors contribute critical expertise in capital markets and growth, as well as provide professional networks, enabling their portfolio firms to seize opportunities for rapid scale-up and expansion. As such, their capabilities and investment-related investigations into the firm compensate for any initial firm-specific information asymmetry.

[Insert Table 7 approximately here]

Overall, this table illustrates how ownership structure, equity provision and implicit valuations depend on whether VC investors have selected a firm, and that these patterns are consistent across the financing events. The premium valuations of VC-backed firms can be explained by venture investors' superior selection and treatment capabilities, founders in these firms having invested more equity and thus possessing more "skin in the game" and more competition from interested investors, and investors of other types effectively placing a higher value on VC-backed firms.

VI. Conclusion

Most attempts at innovation are unobservable because they occur in privately held firms that never go on to become the success stories that capture the public's attention. This fact leaves us unable to draw a clear picture of the risk and return to investing in early-stage innovative firms. Our study sheds new light on the risk and return dynamics of early-stage investments by leveraging comprehensive Norwegian administrative data from 2004 to 2018 to address the longstanding challenge of sample selection bias.

In contrast to prior research that has primarily examined venture-backed firms, we leverage data that allow us to go beyond measuring the performance of VC investments. Instead, we measure the returns that an investor wishing to invest in any innovative startup might experience, in full knowledge that she must make the investment before the success or failure of the firm is known. Not only is this question important in its own right; it helps us establish a reference point for observed venture performance by allowing us to

compare observed VC returns to the returns that investors could have obtained had they invested differently but still in firms with ex ante indicators of high innovation potential.

Our comparison of the valuation performance of venture-backed firms and firms receiving multiple rounds of equity with that of the broader set of innovative firms directly illustrates the sample selection biases that have been the focus of prior work. We find that these biases are indeed pronounced; however, they contain two components: Venture bias, driven by the selection process of VCs, has a larger impact on valuation differentials than the bias associated with firms' ability to raise multiple rounds of financing. This insight underscores the pivotal role of matching into venture financing in shaping startup outcomes, but also the superior investment returns achieved by venture capital investors relative to the universe of comparable investments. At the same time, our findings on the pronounced skewness in returns emphasize the importance of diversification and the potential for overlooked opportunities outside traditional VC pipelines.

Politicians and policymakers seeking to boost entrepreneurship by "helping to identify the next Google" must confront the fact that, among the hundreds of thousands of new firm starts, only a tiny, and hard-to-spot, fraction deliver growth and thrive. Most either fail or remain stagnant. Our findings help quantify the extent of this phenomenon and highlight the need to design funding mechanisms and innovation policies that reduce selection bias and foster equitable access to capital, especially for firms operating in regions or industries that may lack immediate venture appeal.

References

Adelino, M.; S. Ma; and D. Robinson. "Firm Age, Investment Opportunities, and Job Creation." *The Journal of Finance*, 72 (2017), 999–1038.

Adelino, M.; W. B. McCartney; and A. Schoar. "The Role of Government and Private Institutions in Credit Cycles in the U.S. Mortgage Market." (2020). NBER Working paper.

Bailey, M. J.; R. F. Muth; and H. O. Nourse. "A Regression Method for Real Estate Price Index Construction." *Journal of the American Statistical Association*, 58 (1963), 933–942.

Bessembinder, H. "Do Stocks Outperform Treasury Bills?" *Journal of Financial Economics*, 129 (2018), 440 – 457.

Case, K. E., and R. J. Shiller. "The Efficiency of the Market for Single-Family Homes." *The American Economic Review*, 79 (1989), 125–137.

Cochrane, J. H. "The Risk and Return of Venture Capital." *Journal of Financial Economics*, 75 (2005), 3–52.

Fagereng, A.; M. Mogstad; and M. Rønning. "Why Do Wealthy Parents Have Wealthy Children?" *Journal of Political Economy*, 129 (2021), 703–756.

Farago, A., and E. Hjalmarsson. "Long-Horizon Stock Returns Are Positively Skewed." *Review of Finance*, 27 (2023), 495–538.

Gompers, P. A.; W. Gornall; S. N. Kaplan; and I. A. Strebulaev. "How Do Venture Capitalists Make Decisions?" *Journal of Financial Economics*, 135 (2020), 169–190.

Guzman, J., and S. Stern. "Where is Silicon Valley?" *Science*, 347 (2015), 606–609.

Guzman, J., and S. Stern. "The State of American Entrepreneurship: New Estimates of the Quantity and Quality of Entrepreneurship for 32 US States, 1988–2014." *American Economic Journal: Economic Policy*, 12 (2020), 212–43.

Haltiwanger, J.; R. S. Jarmin; and J. Miranda. "Who Creates Jobs? Small versus Large versus Young." *The Review of Economics and Statistics*, 95 (2013), 347–361.

Harris, R. S.; T. Jenkinson; and S. N. Kaplan. "Private Equity Performance: What Do We Know?" *The Journal of Finance*, 69 (2014), 1851–1882.

Harris, R. S.; T. Jenkinson; S. N. Kaplan; and R. Stucke. "Has Persistence Persisted in Private Equity? Evidence from Buyout and Venture Capital Funds." *Journal of Corporate Finance*, 81.

Hsu, D. H. "What Do Entrepreneurs Pay for Venture Capital Affiliation?" *The Journal of Finance*, 59 (2004), 1805–1844.

Hurst, E., and B. W. Pugsley. "What Do Small Businesses Do?" *Brookings Papers on Economic Activity*, 42 (2011), 73–142.

Hvide, H. K., and B. F. Jones. "University Innovation and the Professor's Privilege." *American Economic Review*, 108 (2018), 1860–98.

Hvide, H. K., and T. G. Meling. "Do Temporary Demand Shocks Have Long-Term Effects for Startups?" *The Review of Financial Studies*, 36 (2022), 317–350.

Hvide, H. K.; T. G. Meling; M. Mogstad; and O. L. Vestad. "Broadband Internet and the Stock Market Investments of Individual Investors." *The Journal of Finance*, 79 (2024), 2163–2194.

Hwang, M.; J. M. Quigley; and S. E. Woodward. "An Index For Venture Capital, 1987-2003." *Contributions in Economic Analysis & Policy*, 4.

Korteweg, A., and M. Sorensen. "Risk and Return Characteristics of Venture Capital-Backed Entrepreneurial Companies." *The Review of Financial Studies*, 23 (2010), 3738–3772.

Mjøs, A., and S. F. Selle. "Regnskapsdatabasen - Norwegian Corporate Accounts." SNF Report no. 04/22 (2022).

Peng, L. "Building a Venture Capital Index." (2001). `https://ssrn.com/abstract=281804`.

Puri, M., and R. Zarutskie. "On the Life Cycle Dynamics of Venture-Capital- and Non-Venture-Capital-Financed Firms." *The Journal of Finance*, 67 (2012), 2247–2293.

Reid, A., and P. Markianidou. "European Innovation Scoreboard." (2024). Report by the EU Commission's Directorate-General for Research and Innovation Directorate-General for Internal Market, Industry, Entrepreneurship and SMEs.

Ring, M. A. K. "Entrepreneurial Wealth and Employment: Tracing Out the Effects of a Stock Market Crash." *The Journal of Finance*, 78 (2023), 3343–3386.

Robinson, D. T., and B. A. Sensoy. "Do Private Equity Fund Managers Earn Their Fees? Compensation, Ownership, and Cash Flow Performance." *The Review of Financial Studies*, 26 (2013), 2760–2797.

Robinson, D. T., and B. A. Sensoy. "Cyclicality, Performance Measurement, and Cash Flow Liquidity in Private Equity." *Journal of Financial Economics*, 122 (2016), 521–543.

Shneor, R.; K. Wenzlaff; K. Boyko; P. Baah-Peprah; A. Odorovic; and O. Okhrimenko. "The European Crowdfunding Market Report." (2023). The findings of the European Crowdfunding Market Survey 2023.

Sorensen, M. "How Smart Is Smart Money? A Two-Sided Matching Model of Venture Capital." *The Journal of Finance*, 62 (2007), 2725–2762.

FIGURE 1: **VC Investments Across Countries**

Figure 1 shows the cross-country comparison of total venture capital (VC) investments in USD per capita over 2007–2018 (source: OECD.Stat).

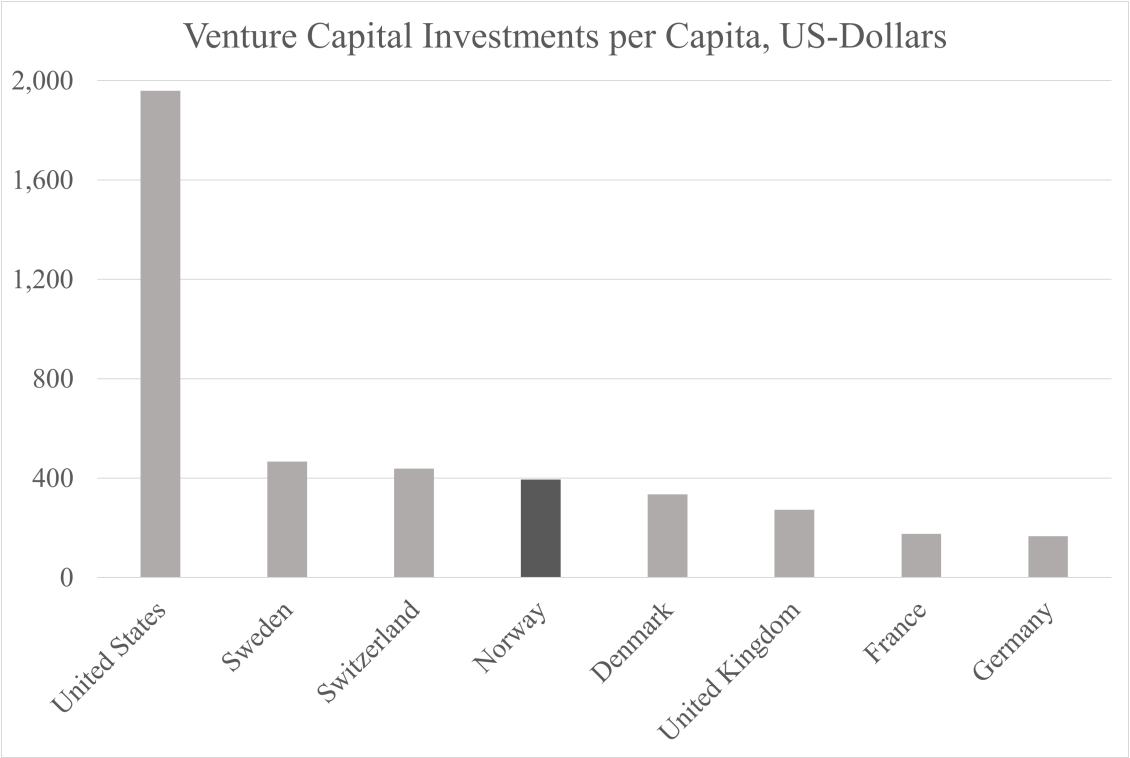


FIGURE 2: **Data Structure**

Figure 2 shows the data structure by representing the timeline of exemplary equity transactions in two fictitious firms. The top firm is assumed to still be independently operating, while the bottom firm is assumed to have been acquired. We distinguish primary transactions, in which new shares are issued at firm inception (blue) or in financing rounds (orange), and secondary trades, where shares are transferred from existing investors to buyers (pink). Shares can be realized through a sale (either to a third party or the firm itself) or liquidation at different points in time, or they are held untraded by the end of our sample period (12/31/2018). Our data provide information on the identities of investors, purchase and realization types, purchase and realization dates, numbers of shares transacted and purchase and realization amounts. Based on the investor identities, we can differentiate across various investor types.

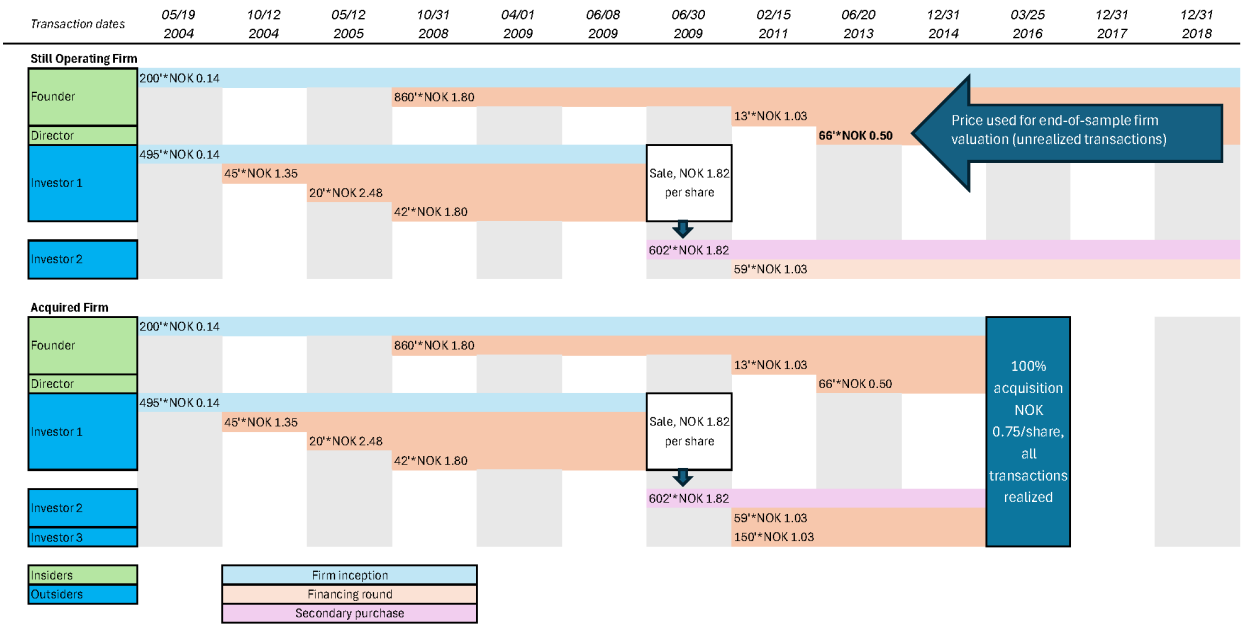


FIGURE 3: **Quantifying the Selection Bias**

Figure 3 provides a visual description of the sample selection bias and presents estimates of the VC valuation premium and multiround valuation premium over non-VC-backed single-round firms, using equation (6). In particular, the VC valuation premium is the plot of the $\beta_2 + \beta_{3t}$ coefficients, and the multiround premium is the plot of the $\beta_4 + \beta_{5t}$ coefficients. All estimations are based on the sample of HIP firms. Firms are categorized as VC-backed firms only from the time of the first equity capital injection from a VC investor and as multiround firms only from their second financing round. If the firm is both VC-backed and has multiple rounds at the same time, we classify it as VC-backed. We do not consider the initial equity injection at firm establishment to be a financing round if it does not exceed the legally required equity amount. We add an implicit last financing round with the implicit postmoney valuation of zero for failed (bankrupt, liquidated, disappeared or zombie) firms at the time of their exit.

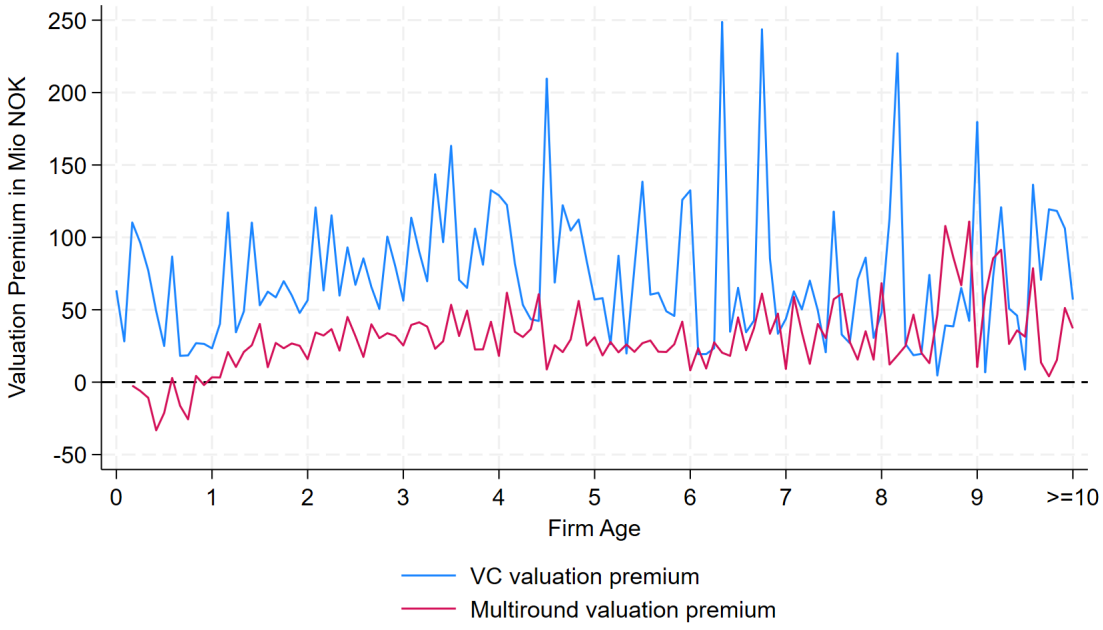


FIGURE 4: **VC Valuation Premium**

Figure 4 plots the VC valuation premium as a function of the firm's age and presents estimates of the VC valuation premium, using equations (7)–(10) with VC_i estimated as a static, time-invariant, flag defined at the firm level, meaning that it equals one for all points in time for a firm that at some point receives VC financing. Each graph plots the time-series formed by summing $\beta_2 + \beta_{3t}$ for each of the estimated equations. The blue graph represents the total VC valuation premium, the red graph excludes variation in valuations due to industry matching, the green graph in addition excludes the variation in valuations due to investment timing (in calendar months). The yellow graph in addition controls for the non-VC equity raised, while the purple one does so for the total equity raised by the respective firm. All estimations are based on the sample of HIP firms. We do not consider the initial equity injection at firm establishment to be a financing round if it does not exceed the legally required equity amount. We add an implicit last financing round with the implicit postmoney valuation of zero for failed (bankrupt, liquidated, disappeared or zombie) firms at the time of their exit.

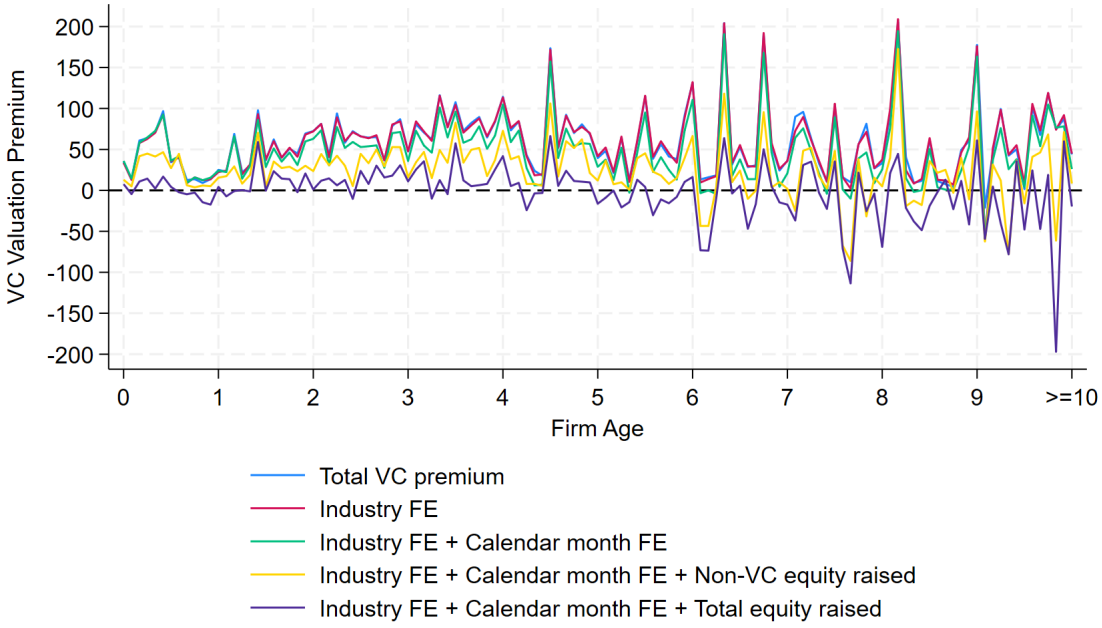


FIGURE 5: **Multiround Valuation Premium**

Figure 5 plots the multiround valuation premium as a function of the firm's age and presents estimates of the multiround valuation premium, using equations (7)–(10), but replacing VC_i with a (static, time-invariant) dummy variable MR_i taking value one if a firm i is a multiround firm. Each graph plots the time-series formed by summing $\beta_2 + \beta_{3t}$ for each of the estimated equations. The blue graph represents the total multiround valuation premium, the red graph excludes variation in valuations due to industry matching, the green graph in addition excludes the variation in valuations due to investment timing (in calendar months) and the yellow graph in addition controls for the total equity previously raised by the respective firm. All estimations are based on the sample of HIP firms. We do not consider the initial equity injection at firm establishment to be a financing round if it does not exceed the legally required equity amount. We add an implicit last financing round with the implicit postmoney valuation of zero for failed (bankrupt, liquidated, disappeared or zombie) firms at the time of their exit.

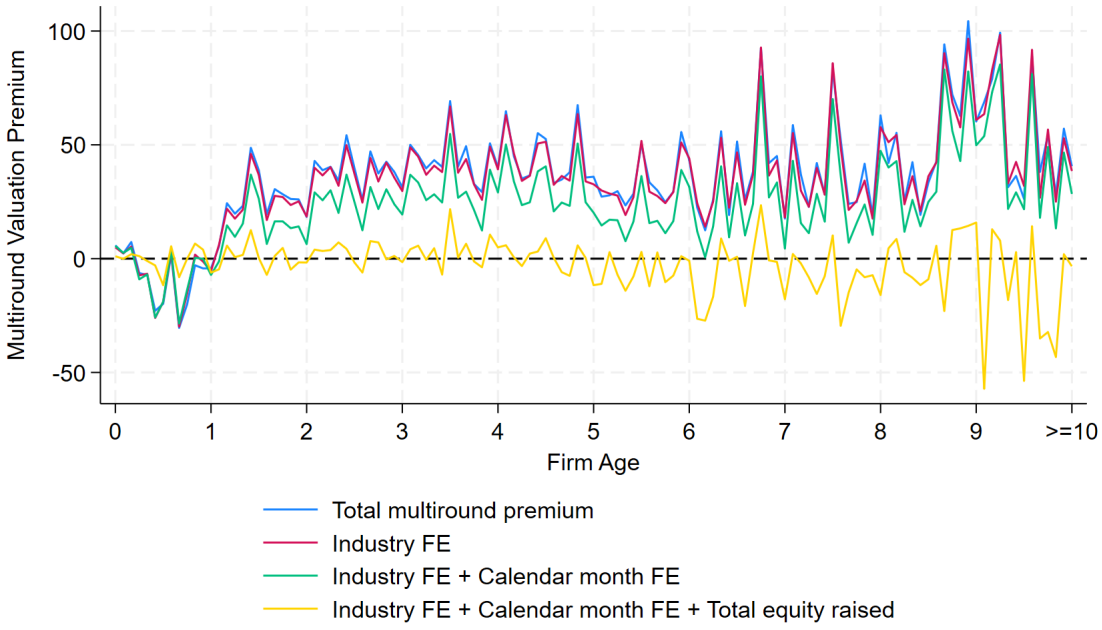


TABLE 1: **Sample Construction**

Table 1 describes our sample construction process. Panel A begins with all new establishments in Norway between 2004 and 2018 as found in the business registry. Panel B describes our process for identifying the subsample of firms with a high propensity to engage in innovation. Panel C shows the resulting sample of HIP firms identified by at least two flags (the sample used in our analysis) and its relative share of the starting sample and of firms that received at least one VC investment. Venture backing includes traditional, corporate or government-affiliated VCs, early-stage investment funds associated with traditional private equity (PE) groups, and incubators.

Panel A. Full Sample

	Firms	% of (A)
New establishments in 2004–2018 (in the registry)	358,745	
- Not a limited liability firm (=C-corp)	-36,669	
- Financial services and real estate firms	-143,496	
- Subsidiaries of established firms	-17,917	
- Holding structures	-10,059	
- Transaction data not matched	-26,256	
Population of newly established operating firms: (A)	**124,348**	100.00%
of which at least one equity financing round:	45,725	36.77%
of which multiple equity financing rounds:	9,698	7.80%
of which at least one secondary trade:	24,814	19.96%
of which no equity transactions (above legal requirement):	69,664	56.02%
***of which VC-backed:* (B)**	928	0.75%

Panel B. Ex Ante Innovation Flags

	Firms	% of (A)
English name	35,200	28.31%
of which VC-backed:	555	59.81% of (B)
Located in an innovation hub (Oslo, Bergen, Stavanger, Trondheim)	34,217	27.52%
of which VC-backed:	522	56.25% of (B)
Operating in a potentially innovative industry	79,196	63.69%
of which VC-backed:	861	92.78% of (B)
At least one board member who lives far from the firm	20,011	16.09%
of which VC-backed:	413	44.50% of (B)

Panel C. Sample Firms with High Innovation Potential

	Firms	% of Baseline
Any two ex ante innovation flags	51,243	41.21% of (A)
of which VC-backed:	813	87.61% of (B)
And at least one equity financing round:	19,721	38.49%
of which VC-backed:	761	93.60%

TABLE 2: **Predicting Later-Stage Firm Outcomes with Ex Ante Innovation Flags**

Table 2 reports the results of a regression of later-stage firm outcomes on the four flags used to define our HIP sample in Table 1. In Panel A, the dependent variables are indicator variables for receipt of any VC financing or governmental innovation-related grant (logit estimations). In Panel B, in the first two sets of regressions, the dependent variables are indicator variables for whether the firm has applied for a patent and has experienced a successful exit, defined as a merger, acquisition or IPO (logit estimations). In the final set of regressions, the dependent variable is the growth in revenues between the end of the first year and the end of the fourth year of the firm's life (ordinary least squares (OLS) estimation). All regressions include a year-of-founding fixed effect. A constant term is estimated but suppressed for brevity. Robust standard errors are reported in parentheses. One, two and three asterisks denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A. Predicting Future Financing

	VC Investment (1/0)					Governmental Innovation Grant (1/0)				
English name (1/0)	1.377***				1.005***	1.382***				1.117***
	[0.068]				[0.070]	[0.067]				[0.069]
Innovation hub (1/0)		1.249***			0.931***		0.477***			0.174*
		[0.066]			[0.069]		[0.067]			[0.069]
Innovative industry (1/0)			2.036***		1.717***			2.199***		1.955***
			[0.127]		[0.129]			[0.132]		[0.133]
Distant board member (1/0)				1.384***	1.074***				0.843***	0.590***
				[0.070]	[0.073]				[0.070]	[0.071]
Observations	124,348	124,348	124,348	124,348	124,348	124,348	124,348	124,348	124,348	124,348
Pseudo R-squared	5.1%	4.5%	5.4%	4.6%	12.4%	7.4%	3.9%	8.1%	4.7%	11.6%

Panel B. Predicting Future Firm Outcomes

	Patent Application (1/0)					Value-Creating Firm Exit (1/0)					4-Year Revenue Growth				
English name (1/0)	1.434***				1.226***	0.281***				0.191***	1.970***				1.612***
	[0.067]				[0.068]	[0.023]				[0.024]	[0.256]				[0.261]
Innovation hub (1/0)		0.169*			-0.132		0.242***			0.143***		1.379***			0.971***
		[0.070]			[0.071]		[0.023]			[0.024]		[0.243]			[0.245]
Innovative industry (1/0)			2.029***		1.798***			0.191***		0.103***			0.944***		0.531**
			[0.124]		[0.126]			[0.022]		[0.023]			[0.196]		[0.199]
Distant board member (1/0)				0.747***	0.527***				0.682***	0.635***				1.971***	1.626***
				[0.072]	[0.073]				[0.024]	[0.024]				[0.290]	[0.292]
Observations	124,348	124,348	124,348	124,348	124,348	124,348	124,348	124,348	124,348	124,348	47,846	47,846	47,846	47,846	47,846
(Pseudo) R-squared	5.4%	1.3%	5.3%	2.1%	9.0%	8.4%	8.3%	8.3%	9.3%	9.5%	0.3%	0.2%	0.1%	0.2%	0.4%

TABLE 3: **Total Capital in Private Capital Market**

Table 3 shows the aggregated distribution of total capital invested in and paid out from our sample of newly established operating firms, denoted category (A) in Table 1, and our subsample of HIP firms. Amounts are reported in million NOK (with 6.6 NOK approximately equal on average to one USD). Percentages are expressed in terms of the overall population indicated in each specific row.

	Overall Population	HIP Firms			
				and VC-Backed	
Number of firms	124,348	51,243		*813*	
		41.2% of Total		*0.7% of Total*	
Total amount:					
Invested in financing rounds	706,053	614,621	*87.1%*	105,735	*15.0%*
Invested in secondary trades	184,716	155,513	*84.2%*	32,056	*17.4%*
Paid out through share sales	183,770	152,991	*83.3%*	37,742	*20.5%*
Paid out through liquidation of shares	29,584	27,029	*91.4%*	4,341	*14.7%*

TABLE 4: **Firm Outcomes**

Table 4 shows the distribution of exit events for firms for which we observe at least one capital market-based firm equity valuation, including an implicit valuation of zero for failed firms. A firm is still independently operating if it has not officially filed for bankruptcy, has neither (partially) liquidated nor merged nor (fully) acquired, has not gone public, still submits its financial accounts to the business registry as of 2018 and has been generating revenues in the most recent five years (2014—2018). A firm is classified as a zombie if it meets all of those criteria except that it has generated no revenues in the most recent five years (2014–2018). A firm is instead classified as disappeared if it satisfies those criteria but has stopped submitting its financial accounts to the business registry before the end of our sample period. IPO, merger and bankruptcy filings are identified as such in the business registry. We identify full acquisition through the existence of a parent company with an ownership share of >90% after its incorporation in the business registry. We supplement this categorization by identifying (partial) liquidations of shares in the tax records on equity transactions.

	HIP firms		Non-HIP firms
		of which VC-backed	
No of firms	33,503	800	44,608
Independently operating	34.1%	47.2%	38.0%
IPO	0.1%	2.6%	0.0%
Full acquisition (>90%)	8.7%	17.2%	6.3%
Merger	3.5%	7.4%	2.2%
Bankruptcy filing	17.0%	6.0%	20.9%
(Partial) Liquidation	17.6%	10.1%	17.3%
Disappearance	11.0%	4.5%	9.6%
Zombie	8.1%	4.9%	5.7%

TABLE 5: **Summary Statistics — Equity-Raising Events and Secondary Trades**

Table 5 reports descriptive statistics of equity-raising events and secondary trades in our sample of HIP firms and in non-HIP firms. An equity-raising event is defined at the firm–firm age level in case new shares are issued. We do not consider the initial equity injection at firm establishment to be a financing round if it does not exceed the legally required equity amount. Postmoney valuation is calculated as shown in equation (3). Amounts are reported in million NOK (with 6.6 NOK approximately equal on average to one USD). The amount of equity raised and firm valuations are winsorized at the 99.5th percentile.

	HIP firms						Non-HIP firms		
				of which VC-backed					
	N	Mean	p50	N	Mean	p50	N	Mean	p50
Equity raised									
% of firms	38%			94%			36%		
N equity raised per firm	19,721	1.6	1.0	761	4.1	3.0	25,985	1.3	1.0
Firm age first equity raised	19,721	1.3	1.0	761	1.5	1.0	25,985	1.3	1.0
Firm age 0–1									
Non-founder ownership	16,626	64.3%	61.7%	532	70.8%	80.0%	22,860	59.9%	50.0%
Non-founder equity raised (MNOK)	16,626	3.76	0.10	532	31.04	1.22	22,860	0.70	0.08
Postmoney valuation (MNOK)	16,626	5.94	0.12	532	46.39	2.72	22,860	1.11	0.12
Firm age 2–3									
Non-founder ownership	5,739	37.7%	31.2%	737	39.2%	31.0%	4,501	35.6%	33.3%
Non-founder equity raised (MNOK)	5,739	10.59	0.77	737	24.28	4.26	4,501	2.89	0.20
Postmoney valuation (MNOK)	5,739	28.54	2.80	737	66.66	15.75	4,501	7.16	0.75
Firm age 4–6									
Non-founder ownership	3,360	28.9%	21.1%	575	27.4%	20.6%	2,869	29.6%	24.8%
Non-founder equity raised (MNOK)	3,360	9.66	1.00	575	19.18	5.60	2,869	2.45	0.23
Postmoney valuation (MNOK)	3,360	38.15	5.47	575	82.93	28.24	2,869	7.69	1.00
Firm age >6									
Non-founder ownership	1,548	26.0%	17.4%	350	24.6%	16.1%	1,450	26.6%	20.0%
Non-founder equity raised (MNOK)	1,548	11.29	1.45	350	16.11	5.98	1,450	3.19	0.30
Postmoney valuation (MNOK)	1,548	50.40	8.74	350	87.68	39.68	1,450	12.80	1.37
Secondary trades									
% of firms	23%			58%			18%		
Firm age	22,274	4.0	3.0	1,321	4.3	3.0	21,175	4.2	3.0
Amount traded (MNOK)	22,274	1.79	0.10	1,321	7.17	1.00	21,175	0.66	0.09
Firm valuation (MNOK)	22,274	13.52	0.58	1,321	62.75	13.08	21,175	3.05	0.30

TABLE 6: **Benchmark Returns**

Table 6 reports equally weighted (EW) and capital-weighted ($CapW$) distributions of firm-level TVPI, defined as the latest observable firm valuation divided by the total equity raised by the firm from its incorporation date until that valuation (equation (5)). The capital-weighted distribution uses the denominator in equation (5) as weights. We show both unwinsorized means and means winsorized at the 99.5th percentile.

		Mean				Percentiles					
	N	Unwins	Wins	SD	Skew.	p10	p50	p75	p90	p95	p99
All VC-backed Firms in HIP Sample											
Observed TVPI (EW)	761	2.62	2.37	4.43	6.50	0.52	1.31	2.78	5.14	7.65	23.91
Add Implied Failures	786	1.96	1.72	3.57	5.53	0.00	1.01	2.17	4.80	6.50	19.31
Add Secondary Trades	800	4.72	2.99	36.90	16.32	0.00	1.00	2.21	4.95	7.83	37.20
Observed TVPI (CapW)	761	1.60	1.58	2.43	4.10	0.08	1.00	1.47	3.24	8.50	8.50
Add Implied Failures	786	0.77	0.74	1.46	9.04	0.00	0.41	1.02	1.67	2.41	4.95
Add Secondary Trades	800	0.80	0.80	1.98	70.37	0.00	0.17	1.02	1.66	2.48	6.28
All HIP Firms											
Observed TVPI (EW)	19,721	1.59	1.46	3.42	10.78	0.61	0.98	1.01	2.53	4.63	13.48
Add Implied Failures	30,663	0.71	0.62	2.39	13.17	0.00	0.00	0.92	1.11	2.44	8.74
Add Secondary Trades	33,503	139.80	2.78	23,974	182.90	0.00	0.00	1.00	2.89	8.35	78.93
Observed TVPI (CapW)	19,721	1.39	1.33	2.42	14.68	0.51	1.00	1.26	2.28	3.01	8.50
Add Implied Failures	30,663	0.88	0.85	1.32	13.65	0.00	0.75	1.02	2.00	2.09	4.72
Add Secondary Trades	33,503	1.77	0.87	1,972	2,225	0.00	0.75	1.00	1.47	2.16	5.56
HIP, VC-backed, multi-round											
Observed TVPI (EW)	561	2.53	2.36	3.67	5.62	0.40	1.51	3.04	5.00	7.46	17.78
Add Implied Failures	561	2.08	1.84	3.61	5.92	0.00	1.14	2.56	4.80	6.41	16.91
Add Secondary Trades	561	2.31	2.31	4.81	6.79	0.00	1.09	2.54	4.86	7.14	23.91
Observed TVPI (CapW)	561	1.61	1.60	2.44	3.29	0.08	1.00	1.50	4.08	8.50	8.50
Add Implied Failures	561	0.72	0.71	1.19	8.81	0.00	0.19	1.02	1.67	2.43	4.76
Add Secondary Trades	561	0.76	0.76	1.59	9.62	0.00	0.14	1.02	1.70	2.62	5.74
HIP, non-VC-backed, multi-round											
Observed TVPI (EW)	4,516	2.31	2.19	3.51	5.87	0.65	1.05	2.35	4.99	7.68	17.52
Add Implied Failures	4,516	1.72	1.53	3.30	6.41	0.00	0.96	1.72	4.20	6.85	14.31
Add Secondary Trades	4,516	2.33	2.27	8.14	16.51	0.00	0.95	1.87	4.75	8.56	27.26
Observed TVPI (CapW)	4,516	1.25	1.24	1.42	10.05	0.51	0.85	1.47	2.79	2.79	5.28
Add Implied Failures	4,516	0.84	0.82	1.25	11.28	0.00	0.51	1.02	2.09	2.09	4.66
Add Secondary Trades	4,516	0.78	0.78	1.65	31.96	0.00	0.51	0.86	1.47	2.14	5.56

TABLE 7: **Average Equity Structure in Equity-Raising Events**

Table 7 reports the capitalization table of equity-raising events in our sample of HIP firms with the average amount purchased and purchased ownership stake by investor type, conditional on investment. Initial equity injection at firm establishment is not considered if it does not exceed the legally required equity amount. Total equity raised by firm and investment equity amounts for each investor type are winsorized at the 99.5th percentile. Units are reported in millions of Norwegian Kroner (approximately 6.6 to the dollar).

	Single-round Firms				Multiround Firms			
	VC-backed		Non-VC-backed		VC-backed		Non-VC-backed	
	N	mean	N	mean	N	mean	N	mean
Total equity raised by firm	203	78.39	14,467	5.37	561	90.58	4,505	22.91
Equity Investments by Investor Type:								
VC								
Equity amount	149	39.61			1,475	13.72		
Ownership	149	40.8%			1,475	20.6%		
Founder								
Equity amount	26	4.93	1,313	1.69	356	1.79	3,107	0.85
Ownership	26	11.8%	1,313	46.3%	356	10.7%	3,107	23.9%
Family								
Equity amount	4	0.99	2,449	0.56	36	4.46	899	0.98
Ownership	4	6.2%	2,449	42.4%	36	14.1%	899	26.3%
Repeat angels								
Equity amount	52	7.71	2,949	1.25	787	2.56	3,858	1.36
Ownership	52	16.7%	2,949	35.3%	787	11.2%	3,858	19.0%
Other individuals								
Equity amount	49	7.69	6,699	0.42	694	1.82	5,093	0.82
Ownership	49	20.4%	6,699	42.7%	694	12.6%	5,093	23.5%
Financial institutions								
Equity amount	4	150.50	2	28.30	47	25.92	49	36.91
Ownership	4	29.6%	2	13.7%	47	9.6%	49	11.1%
Other corporations								
Equity amount	123	20.41	5,584	4.70	1,707	9.27	7,760	6.25
Ownership	123	43.7%	5,584	58.2%	1,707	19.5%	7,760	33.6%
Unknown								
Equity amount	26	75.40	930	13.72	396	16.01	1,365	19.66
Ownership	26	29.3%	930	48.6%	396	12.8%	1,365	26.1%

Appendices

Appendix A. Additional Figures

FIGURE A1: **Quantifying the Selection Bias—All Trades**

Figure A1 replicates Figure 3 but derives the postmoney valuations from valuations in both financing rounds and secondary trades, $\text{POSTMONEY}'_{it}$. All estimations are based on the sample of HIP firms. Firms are categorized as VC-backed firms only from the time of the first equity capital injection from a VC investor and as multiround firms only from their second financing round. If the firm is both VC-backed and has multiple rounds at the same time, we classify it as VC-backed. We do not consider the initial equity injection at firm establishment to be a financing round if it does not exceed the legally required equity amount. We add an implicit last financing round with the implicit postmoney valuation of zero for failed (bankrupt, liquidated, disappeared or zombie) firms at the time of their exit.

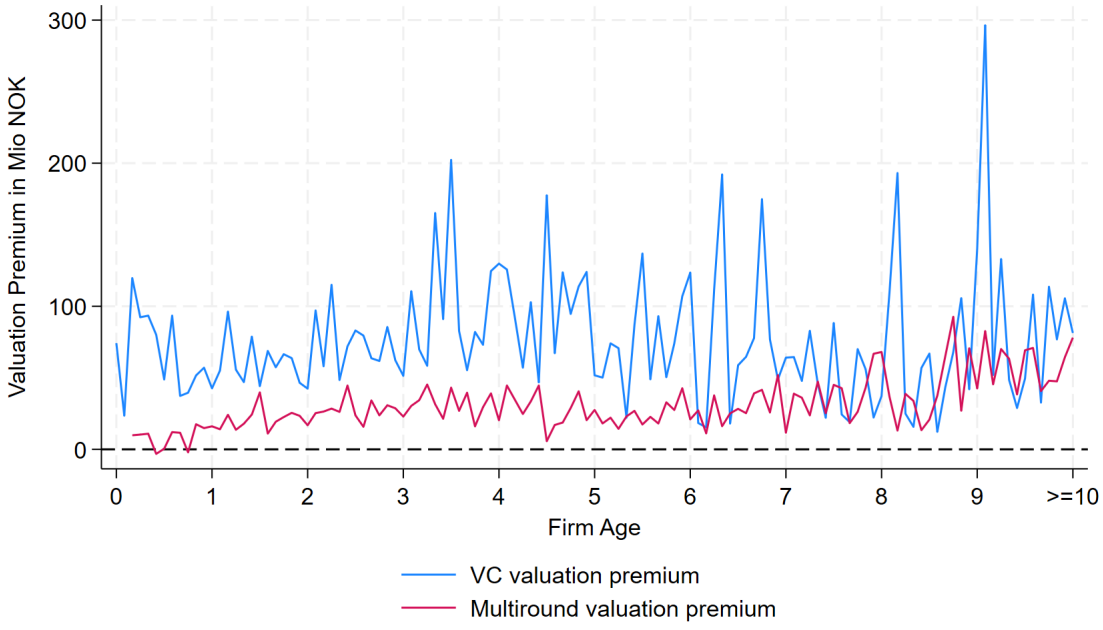


FIGURE A2: **VC Valuation Premium—All Trades**

Figure A2 replicates Figure 4 but derives the postmoney valuations from valuations in both financing rounds and secondary trades, $\text{POSTMONEY}'_{it}$. The blue graph represents the total VC valuation premium, the red graph excludes variation in valuations due to industry matching, the green graph in addition excludes the variation in valuations due to investment timing (in calendar months). The yellow graph in addition controls for the non-VC equity raised, while the purple one does so for the total equity raised by the respective firm. All estimations are based on the sample of HIP firms. We do not consider the initial equity injection at firm establishment to be a financing round if it does not exceed the legally required equity amount. We add an implicit last financing round with the implicit postmoney valuation of zero for failed (bankrupt, liquidated, disappeared or zombie) firms at the time of their exit.

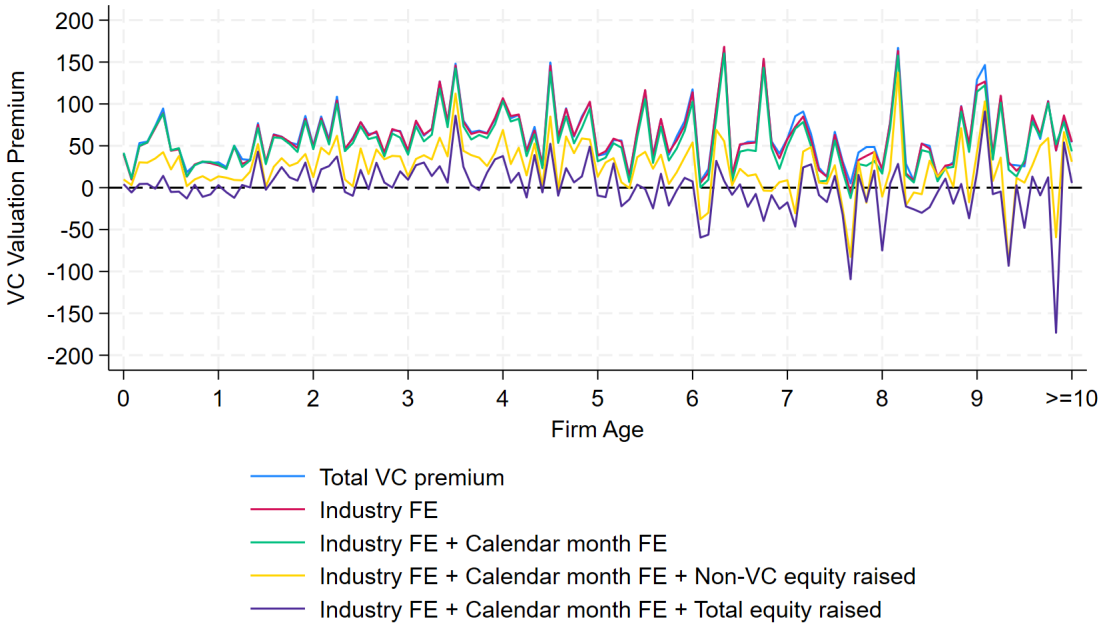


FIGURE A3: **Multiround Valuation Premium—All Trades**

Figure A3 replicates Figure 5 but derives the postmoney valuations from valuations in both financing rounds and secondary trades, $\text{POSTMONEY}'_{it}$. The blue graph represents the total multiround valuation premium, the red graph excludes variation in valuations due to industry matching, the green graph in addition excludes the variation in valuations due to investment timing (in calendar months) and the yellow graph in addition controls for the total equity previously raised by the respective firm. All estimations are based on the sample of HIP firms. We do not consider the initial equity injection at firm establishment to be a financing round if it does not exceed the legally required equity amount. We add an implicit last financing round with the implicit postmoney valuation of zero for failed (bankrupt, liquidated, disappeared or zombie) firms at the time of their exit.

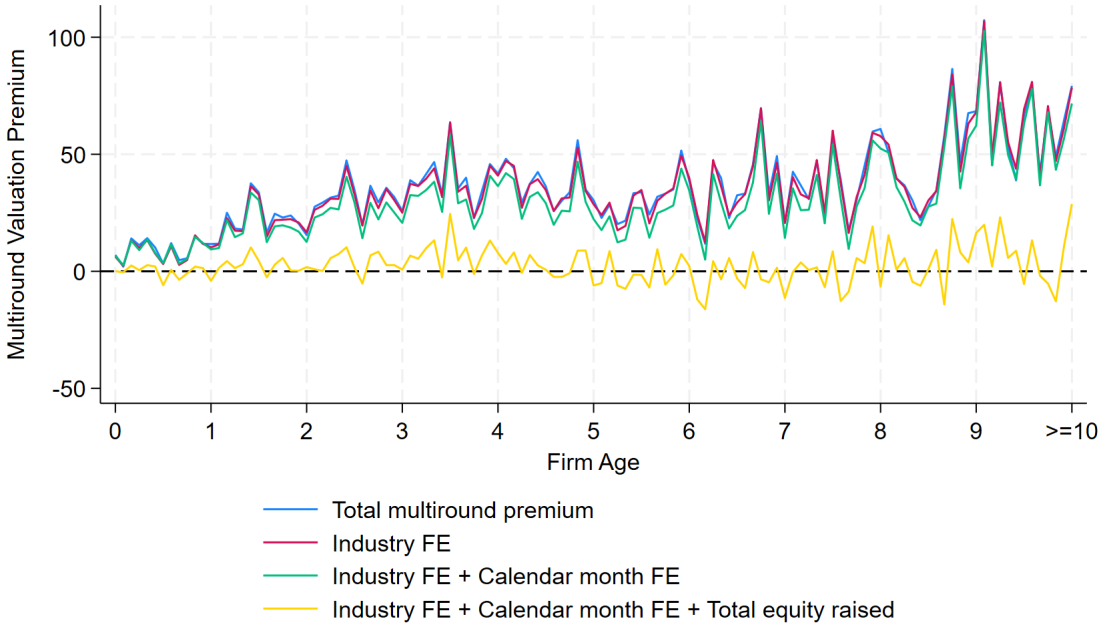
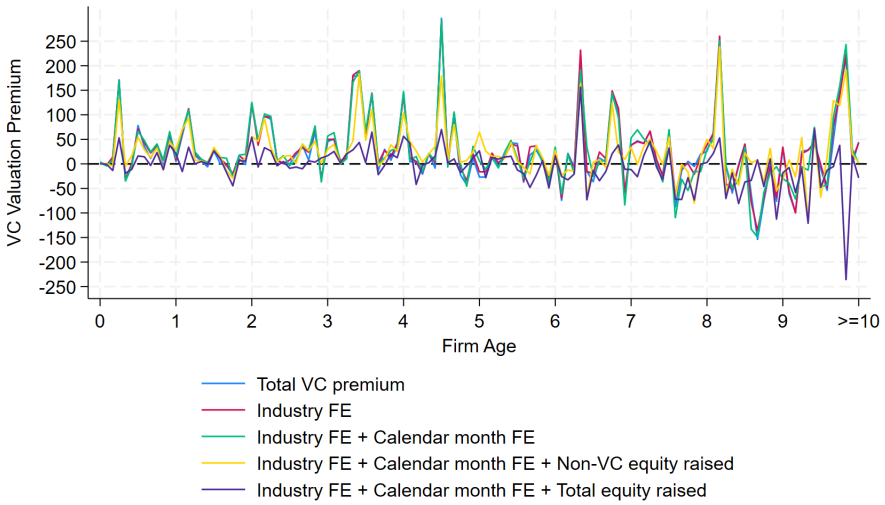


FIGURE A4: **Firm Attrition**

Figure A4 replicates Figures 4 and 5 for the subsample of firms that have survived for at least ten years.

Panel A. VC Valuation Premium



Panel B. Multiround Valuation Premium

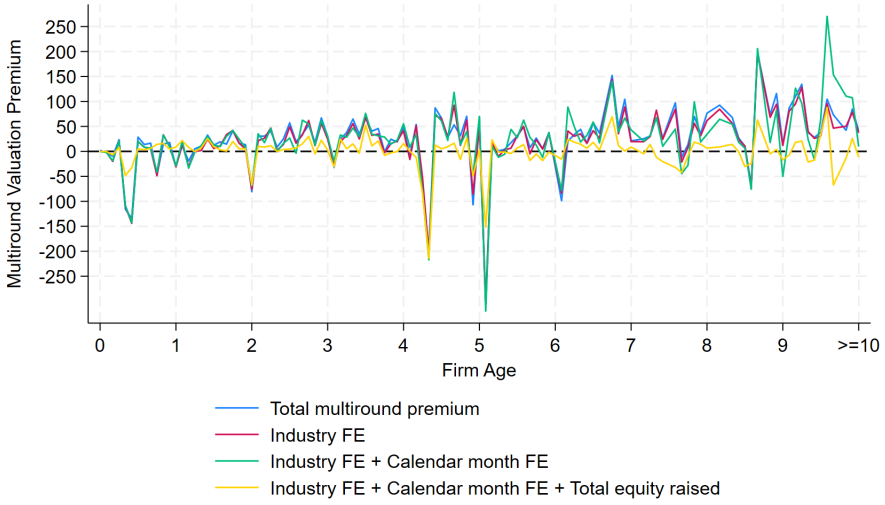
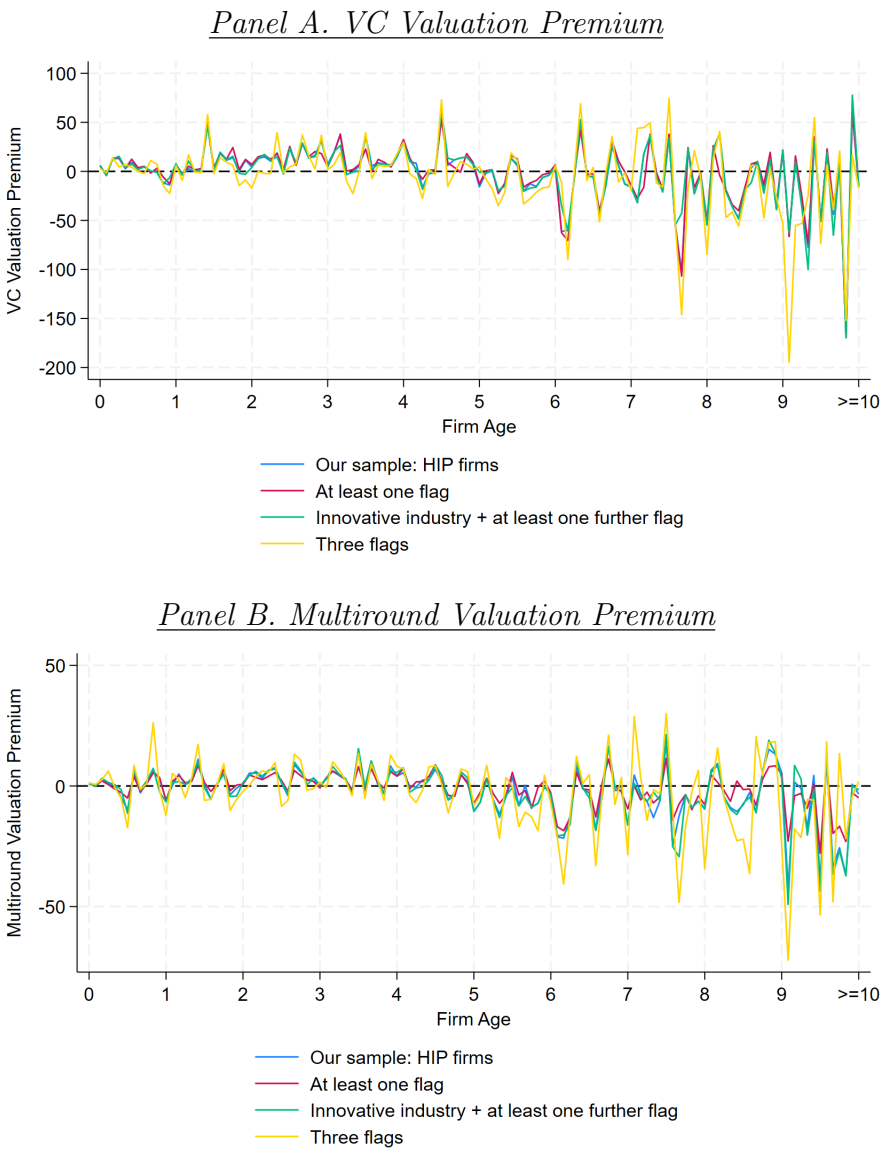


FIGURE A5: **Robustness of Selection of HIP Firms**

Figure A5 examines the robustness of our sample selection of HIP firms and replicates the full specification in equation (10), as presented by the purple line in Figure 4 and yellow line in Figure 5, for varying definitions of a HIP firms. The blue graph represents our sample of HIP firms, which satisfy the criteria for any two ex ante innovation flags at the time of firm incorporation. The red graph represents firms satisfying the criterion for at least one ex ante flag, while the green graph represent firms operating in an innovative industry and having at least one further ex ante flag. The yellow graph represents firms having any three ex ante innovation flags.



Appendix B. Raw Equity Transaction Data

This internet appendix provides a detailed description of how the Norwegian tax authority data that we use are collected from firms and investors and provided by the tax authority to us for our research.

Reporting to Tax Authorities

All limited liability firms (*aksjeselskap*, or AS, analogous to C-corporations in the U.S.) report annually to the tax authorities their total number of shares outstanding and associated equity raised using form *RF-1086*, and separately report the shareholdings of each of their investors using form *RF-1086U*.

Figure B1 shows an illustrative snapshot of the corporate firm-level reporting of a small family-owned firm received from a family member of one of this paper's coauthors. This form must be submitted digitally by January 31 of the year following the reporting year. The firm-level form includes:

- Firm identification: The national firm identifier (*organisasjonsnummer*), firm name and postal address, share class, and ISIN identifier if applicable.

- Firm information: Total equity raised, total equity raised in the reported share class (one form submitted per share class), nominal value per share, number of shares in the reported share class, and paid-in share premium. This information is provided as of 01/01 and 12/31 of the reporting year.

- Dividends: Dividends paid out (total and per share), and date(s) of any dividend payouts.

- Contact person: Name, role in the firm, phone number and email address.

- Equity-related events (new share issuance at incorporation or in financing events, mergers, transfers, share buybacks, etc.): In this example, this includes an issuance of new shares in exchange for cash. The information included is the number of new shares issued, total number of shares after the issuance, date of the issuance, nominal value per share, share premium paid-in per share, and number of shares held by the firm itself that are transferred to investors as the part of the issuance. The information on newly issued shares in this table results from aggregation of the information on the forms prepared for each investor, as shown below in Figure B2.

The corporate reporting for each investor uses form *RF-1086U*, which is shown in Figure B2. This investor-level form includes:

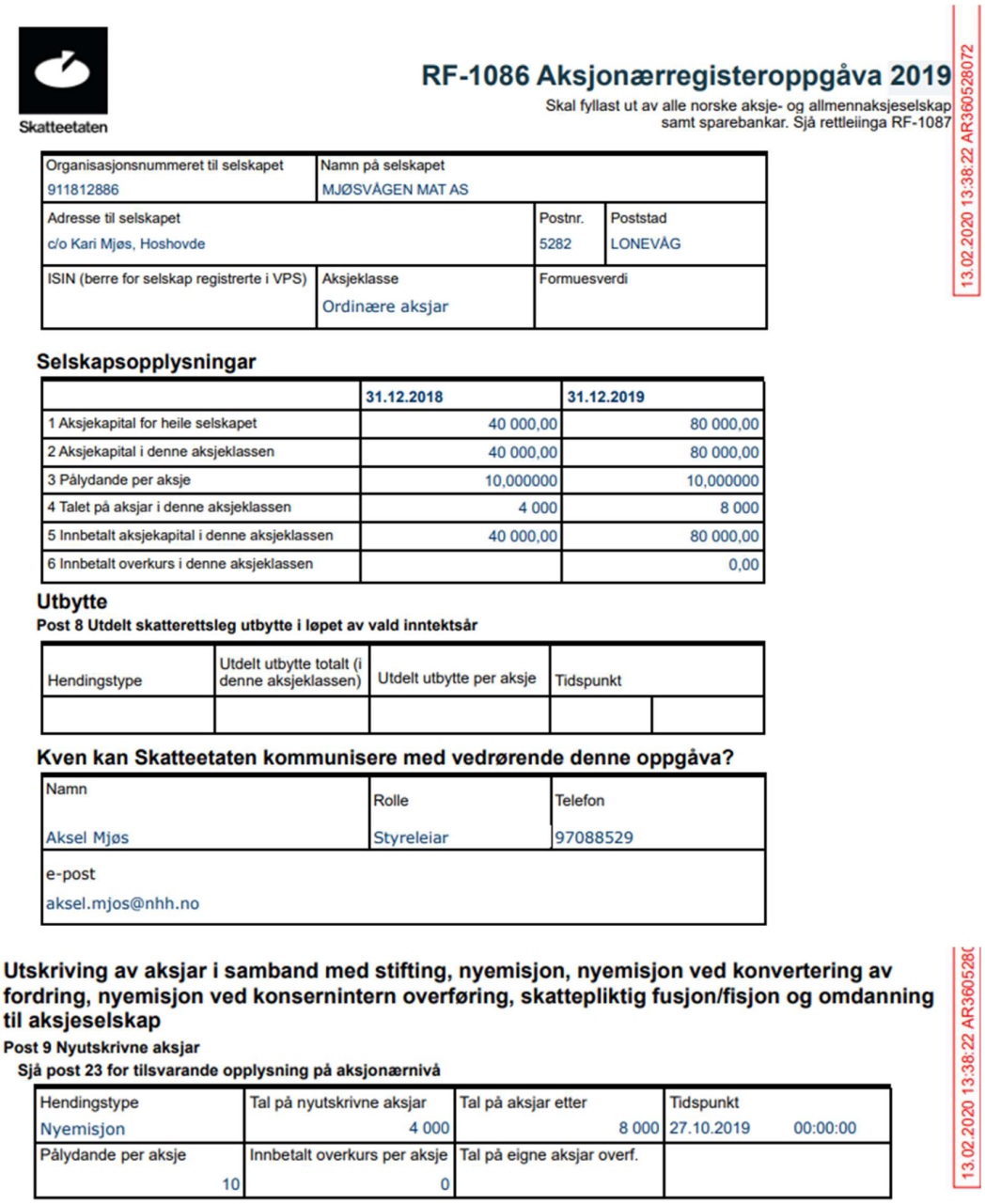
- Firm identification: The national firm identifier (*organisasjonsnummer*), share class, and ISIN identifier, if applicable.

- Investor information: Individual identifier (*løpenummer*) for Norwegian and foreign individual investors, national firm identifier (*organisasjonsnummer*) for institutional investors, firm (individual) name, country, postal address, and number of shares held as of 01/01 and 12/31 of the reporting year.

- Investor equity transactions (per transaction in this firm): Transaction type (primary purchase [at firm incorporation or other new equity issuance], purchase in a

FIGURE B1: **Corporate Annual Tax Reporting (*RF 1086*)**

Figure B1 presents a snapshot of the corporate annual reporting of total number of shares outstanding and associated equity raised from form *RF-1086*, which must be digitally submitted by each firm before January 31 of the year following the reporting year.



Skatteetaten

RF-1086 Aksjonærregisteroppgåva 2019

Skal fyllast ut av alle norske aksje- og allmennaksjeselskap samt sparebankar. Sjå rettleiinga RF-1087

13.02.2020 13:38:22 AR360528072

Organisasjonsnummeret til selskapet	Namn på selskapet		
911812886	MJØSVÅGEN MAT AS		
Adresse til selskapet		Postnr.	Poststad
c/o Kari Mjøs, Hoshovde		5282	LONEVÅG
ISIN (berre for selskap registrerte i VPS)	Aksjeklasse	Formuesverdi	
	Ordinære aksjar		

Selskapsopplysningar

	31.12.2018	31.12.2019
1 Aksjekapital for heile selskapet	40 000,00	80 000,00
2 Aksjekapital i denne aksjeklassen	40 000,00	80 000,00
3 Pålydande per aksje	10,000000	10,000000
4 Talet på aksjar i denne aksjeklassen	4 000	8 000
5 Innbetalt aksjekapital i denne aksjeklassen	40 000,00	80 000,00
6 Innbetalt overkurs i denne aksjeklassen		0,00

Utbytte

Post 8 Utdelt skatterettsleg utbytte i løpet av vald inntektsår

Hendingstype	Utdelt utbytte totalt (i denne aksjeklassen)	Utdelt utbytte per aksje	Tidspunkt	

Kven kan Skatteetaten kommunisere med vedrørende denne oppgåva?

Namn	Rolle	Telefon
Aksel Mjøs	Styreleiar	97088529
e-post		
aksel.mjos@nhh.no		

Utskriving av aksjar i samband med stifting, nyemisjon, nyemisjon ved konvertering av fordring, nyemisjon ved konsernintern overføring, skattepliktig fusjon/fisjon og omdanning til aksjeselskap

Post 9 Nyutskrivne aksjar

Sjå post 23 for tilsvarande opplysning på aksjonærnivå

13.02.2020 13:38:22 AR360528(

Hendingstype	Tal på nyutskrivne aksjar	Tal på aksjar etter	Tidspunkt	
Nyemisjon	4 000	8 000	27.10.2019	00:00:00
Pålydande per aksje	Innbetalt overkurs per aksje	Tal på eigne aksjar overf.		
10	0			

secondary trade, realization of shares through share sale or share liquidation, inheritance, or merger), number of shares transacted, transaction date, and total transaction amount. In this example, the investor acquired 2,000 shares through a secondary trade on 09/01/2019 and 4,000 shares through new issuance of shares from the firm on 10/27/2019. In both cases, the price paid is NOK 10 per share.

The information in this investor-level reporting is aggregated to the firm level in the reporting form shown in Figure B1. The national tax authority aggregates, on behalf of each taxpayer, these corporate reportings from all limited liability firms and produces an overview of all shareholdings and equity transactions in all firms for each investor. This overview is digitally distributed to each investor separately in reporting form *RF-1088* (see Figures B3 and B4 below). Investors add the provided information on their shareholdings and equity transactions to their income tax returns. Because foreign investors are also taxed in Norway on any income from their shareholdings (both dividends and gains from sales), these forms are also prepared for them.[24]

To supplement corporate reporting, financial intermediaries are required to report all equity transactions they carry out. Finally, the taxpayer is herself responsible for making sure that the data in her tax return are complete and correct. Secondary trades are, thus, self-reported to a higher extent than primary equity transactions by both buyers and sellers of the shares. Given that the number of shares is aggregated in the corporate reporting, we find support for a high level of compliance from investors. As shown in Table 5, we observe 12,828 financing round transactions and 22,274 secondary trades in

[24]Within the European Economic Area, a limited liability firm's income from investing in other limited liability firms is not taxed. However, it still needs to report its shareholding as described here.

FIGURE B2: **Investor-Level Corporate Annual Tax Reporting (*RF-1086U*)**

Figure B2 presents a snapshot of the corporate annual tax reporting for each investor's holdings and transactions – *RF-1086U*. The form must be prepared for each investor and digitally submitted by the firm before January 31 of the the year following the reporting year.

Selskapsidentifikasjon

Org. nr.	Aksjeklasse	ISIN
911812886	Ordinære aksjar	

Aksjonæropplysningar - (namn, adresse, tal på aksjar og utbytte)

Post 19 Namnet på aksjonæren, adresse, fødselsnummer, organisasjonsnummer, utanlandsk aksjonær-ID og land

Aksjonæridentifikasjon (fødsels-/D-nummer, organisasjonsnummer, utanlandsk aksjonær-ID)			
25126140854			
Namn			Land
Kari Laurhammer Mjøs			NOREG
Adresse		Postnummer	Poststad
Hoshovde 28		5282	Lonevåg
Post 20 Tal på aksjar per aksjonær	Tal på 31.12.2018 2 000	Tal på 31.12.2019 8 000	

Transaksjonar (aksjar i tilgang/avgang) for denne aksjonæren

Til dømes kjøp, stifting, skattepliktig fusjon/fisjon (for fullstendig oversikt over transaksjonstypar, sjå rettleiinga)

Post 23 Aksjar i tilgang (anskaffing)

Sjå post 9 for opplysningar på selskapsnivå

Transaksjonstype	Tal på aksjar	Tidspunkt	
Kjøp	2 000	01.09.2019	00:00:00
Total anskaffingsverdi	Givers/arvelaters f.nr.	Givers org.nr.	
2 000			

Transaksjonstype	Tal på aksjar	Tidspunkt	
Nyemisjon	4 000	27.10.2019	00:00:00
Total anskaffingsverdi	Givers/arvelaters f.nr.	Givers org.nr.	
40 000			

13.02.2020 13:38:22 AR360528072

early-stage innovative firms, volumes that make us confident that the data cover a very high proportion of secondary trades.

Figures B3 and B4 are snapshots of an annual investor reporting form *RF-1088* received by one of the coauthors of this paper. For illustration purposes, these snapshots include only the shares held in one firm. The pages of the form include[25]:

- Page 2 represents the shareholding overview (Figure B3) and includes one line per invested firm and share class, if multiple.
  - Firm name and national firm identifier (*organisasjonsnummer*)
  - Share class
  - Number of shares held as of 01/01 and 12/31 of the reporting year
  - Taxable dividends received during the reporting year
  - Taxable gain/loss from the realization of shares
  - Taxable value of shareholding calculated for wealth tax purposes. Left empty in the example, as it was not finally calculated by the firm at the time of reporting. The investor in this example had to add in this taxable value himself later before filing his tax return. We do not use this information in our analyses.
- Page 3 represents the equity transactions overview (Figure B4) and includes one table per invested firm and share class, if multiple. This example includes only primary share purchases but is representative for the reporting on any type of transaction.

[25]The first page of the form includes only the taxpayer's identity and contact details, as well as some guidance on filling in the form. The last page duplicates the values from the previous page of the form. Both are excluded here for brevity.

FIGURE B3: **Investor-Level Shareholdings Overview (*RF-1088* Page 2)**

Figure B3 presents the second (first informative) page of the annual investor-level overview of all shareholdings and equity transactions prepared by the national tax authority for each taxpayer. This page is dedicated to all the investor's shareholdings. This form is sent to each investor for subsequent inclusion in her tax return.

Side 2 av 4 03.03.22 25036446105 +§I

Oversikt over aksjar du eigde i 2020

Her finn du ei oversikt over selskapa du har hatt aksjar i i løpet av 2021. Skattepliktig utbytte, skattepliktig gevinst, tap med fradragsrett og formuesverdi er viste i denne tabellen summert per selskap og per aksjeklasse. Dersom du har aksjar i andre selskap, skal du opplyse om desse. Er du ein av dei som har motteke ein ny temabasert elektronisk skattemelding, kan du endre direkte i den nye skattemeldinga, hvis ikkje bruker du skjema RF-1159.

200 Namn på selskap	201 Organisasjons-nummer	202 Aksjeklasse	203 Talet på aksjar ved inn-gangen av året	204 Talet på aksjar ved ut-gangen av året	210 Skattepliktig utbytte	211 Skattepliktig gevinst	212 Tap med fradragsrett	209 Formuesverdi
HOLBERG EEG AS	995 109 670	Ordinære aksjer	1 110	1 110				Ukjent

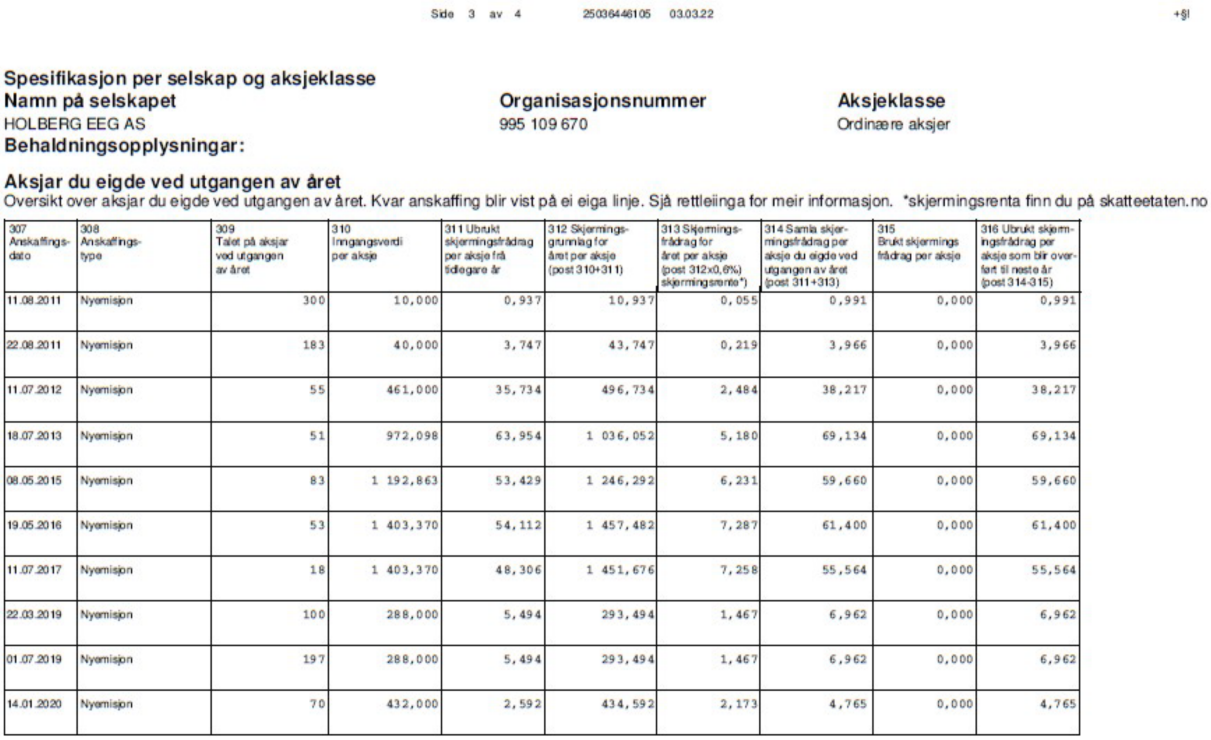
- Transaction date
- Type of transaction – in this example, new issuance of shares (after incorporation)
- Number of shares transacted
- Price paid per share
- (Last six columns) Tax-shielding deduction mechanism to reduce the actual tax payment in case of sales gains. We do not use this information in our analyses.

Accessed Data

The Norwegian tax authority provided us with the annual firm-and investor-level corporate reporting and investor-level shareholding and equity transaction overviews for

FIGURE B4: **Investor-Level Shareholdings Overview (*RF-1088* Page 3)**

Figure B4 presents the third (second informative) page of the annual investor-level overview of all shareholdings and equity transactions prepared by the national tax authority for each taxpayer. This page is dedicated to all the investor's equity transactions. Each line represents a transaction that contributes to the investor's shareholding in the firm as of 31/12 of the reporting year. This form is prepared by the tax authorities and sent to each investor for subsequent inclusion in her tax return declaration.



Side 3 av 4 25036446105 03.03.22 +§l

Spesifikasjon per selskap og aksjeklasse

Namn på selskapet	**Organisasjonsnummer**	**Aksjeklasse**
HOLBERG EEG AS	995 109 670	Ordinære aksjer

Behaldningsopplysningar:

Aksjar du eigde ved utgangen av året

Oversikt over aksjar du eigde ved utgangen av året. Kvar anskaffing blir vist på ei eiga linje. Sjå rettleiinga for meir informasjon. *skjermingsrenta finn du på skatteetaten.no

307 Anskaffings-dato	308 Anskaffings-type	309 Talet på aksjar ved utgangen av året	310 Inngangsverdi per aksje	311 Ubrukt skjermingsfrådrag per aksje frå tidlegare år	312 Skjermings-grunnlag for året per aksje (post 310+311)	313 Skjermings-frådrag for året per aksje (post 312x0,6%) skjermingsrente*)	314 Samla skjer-mingsfrådrag per aksje du eigde ved utgangen av året (post 311+313)	315 Brukt skjermings frådrag per aksje	316 Ubrukt skjerm-ingsfrådrag per aksje som blir over-ført til neste år (post 314-315)
11.08.2011	Nyemisjon	300	10,000	0,937	10,937	0,055	0,991	0,000	0,991
22.08.2011	Nyemisjon	183	40,000	3,747	43,747	0,219	3,966	0,000	3,966
11.07.2012	Nyemisjon	55	461,000	35,734	496,734	2,484	38,217	0,000	38,217
18.07.2013	Nyemisjon	51	972,098	63,954	1 036,052	5,180	69,134	0,000	69,134
08.05.2015	Nyemisjon	83	1 192,863	53,429	1 246,292	6,231	59,660	0,000	59,660
19.05.2016	Nyemisjon	53	1 403,370	54,112	1 457,482	7,287	61,400	0,000	61,400
11.07.2017	Nyemisjon	18	1 403,370	48,306	1 451,676	7,258	55,564	0,000	55,564
22.03.2019	Nyemisjon	100	288,000	5,494	293,494	1,467	6,962	0,000	6,962
01.07.2019	Nyemisjon	197	288,000	5,494	293,494	1,467	6,962	0,000	6,962
14.01.2020	Nyemisjon	70	432,000	2,592	434,592	2,173	4,765	0,000	4,765

reporting years 2004–2018. We received the investor-level shareholding and equity transaction overviews split into three files: shareholdings, share purchases and share realizations. We use only the latter two and the firm-level corporate reporting in our analyses.

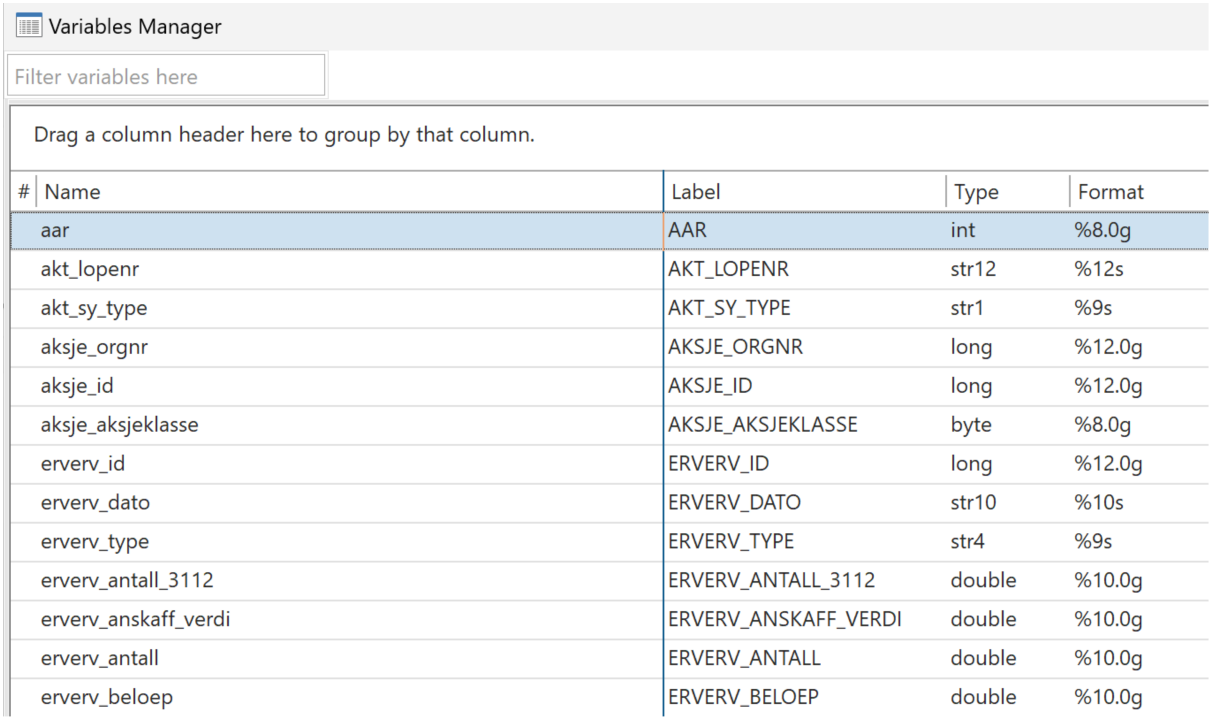
Purchases File *(RF-1088)*

Figure B5 presents a snapshot of the variables we use from the annual data files received on shareholdings and equity transactions at the investor level (*RF-1088*). The file includes one line per share purchase transaction. We use the following variables from this purchase file for our analyses:

- *aar*: Reporting year
- *akt_lopenr*: Unique investor identifier
- *akt_sy_type*: Investor type as categorized by the tax authority
- *aksje_orgnr*: National firm identifier (*organisasjonsnummer*) of the invested firm
- *aksje_id*: Unique share identifier allocated by the tax authority
- *aksje_aksjeklasse*: Share class
- *erverv_id*: Unique transaction (share purchase) identifier allocated by the tax authority
- *erverv_dato*: Transaction (share purchase) date

FIGURE B5: **Relevant Variables in the Purchase File (*RF-1088*)**

Figure B5 presents the snapshot of variables that we use in the annual data files received on shareholdings and equity transactions at investor level (*RF-1088*). The file includes one line per share purchase transaction.


Variables Manager

Filter variables here

Drag a column header here to group by that column.

#	Name	Label	Type	Format
	aar	AAR	int	%8.0g
	akt_lopenr	AKT_LOPENR	str12	%12s
	akt_sy_type	AKT_SY_TYPE	str1	%9s
	aksje_orgnr	AKSJE_ORGNR	long	%12.0g
	aksje_id	AKSJE_ID	long	%12.0g
	aksje_aksjeklasse	AKSJE_AKSJEKLASSE	byte	%8.0g
	erverv_id	ERVERV_ID	long	%12.0g
	erverv_dato	ERVERV_DATO	str10	%10s
	erverv_type	ERVERV_TYPE	str4	%9s
	erverv_antall_3112	ERVERV_ANTALL_3112	double	%10.0g
	erverv_anskaff_verdi	ERVERV_ANSKAFF_VERDI	double	%10.0g
	erverv_antall	ERVERV_ANTALL	double	%10.0g
	erverv_beloep	ERVERV_BELOEP	double	%10.0g

- *erverv_type*: Purchase type (primary purchase at firm incorporation or in the financing round, secondary trade)
- *erverv_antall_3112*: Total number of shares held as of 12/31 of the reporting year
- *erverv_anskaff_verdi*: Price paid per share
- *erverv_antall*: Number of shares transacted (purchased)
- *erverv_beløp*: Purchase amount of shares

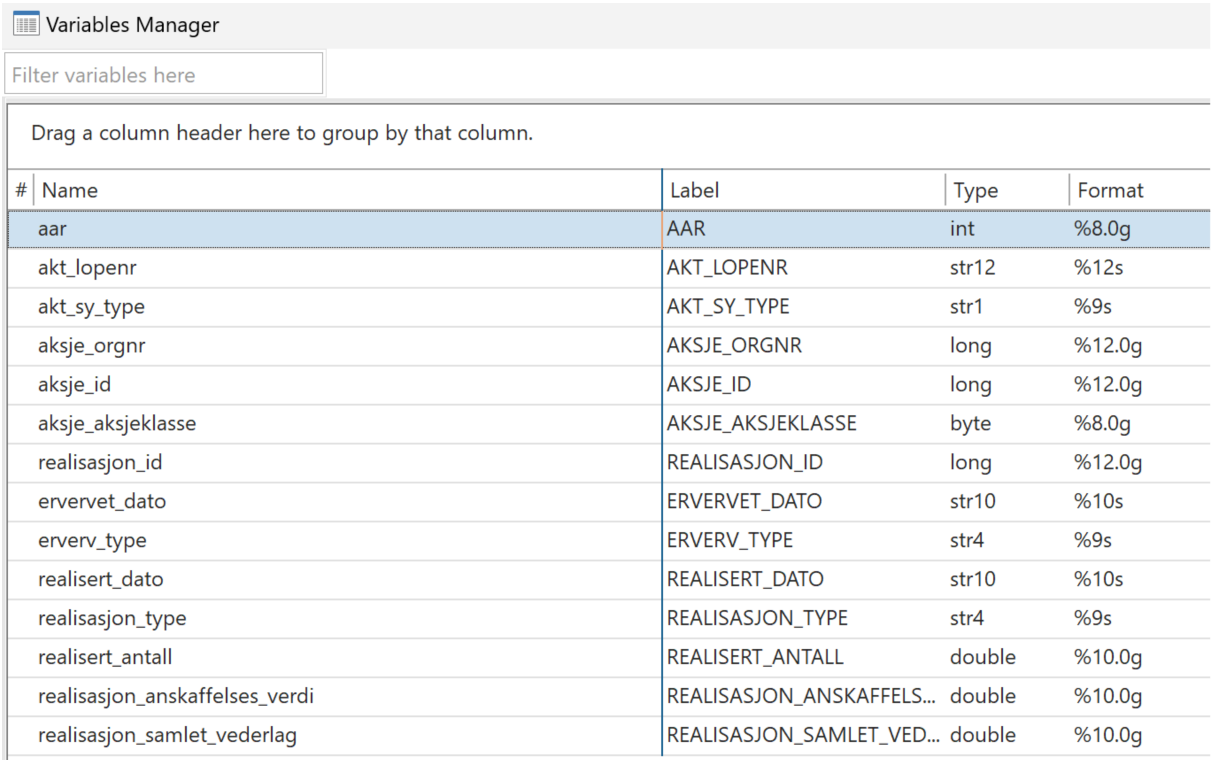
Realization File *(RF-1088)*

Figure B6 presents a snapshot of the variables we use from the annual data files received on shareholdings and equity transaction at investor level (*RF-1088*). The file includes one line per share realization transaction. We use the following variables from this file for our analyses:

- *aar*: Reporting year
- *akt_lopenr*: Unique investor identifier
- *akt_sy_type*: Investor type as categorized by the tax authority
- *aksje_orgnr*: National firm identifier (*organisasjonsnummer*) of the invested firm
- *aksje_id*: Unique share identifier allocated by the tax authority
- *aksje_aksjeklasse*: Share class
- *realisasjon_id*: Unique transaction (share realization) identifier allocated by the tax authority
- *ervervet_dato*: Purchase date of realized shares
- *erverv_type*: Purchase type (primary purchase at firm incorporation or in a financing round, secondary trade) of realized shares
- *realisert_dato*: Transaction (share realization) date
- *realisasjons_type*: Realization type (sale, liquidation)
- *realisert_antall*: Number of shares transacted (realized)

FIGURE B6: **Relevant Variables in the Realization File (*RF-1088*)**

Figure B6 presents a snapshot of the variables we use from the annual data files received on shareholdings and equity transactions at investor level (*RF-1088*). The file includes one line per share realization transaction.


Variables Manager

Filter variables here

Drag a column header here to group by that column.

#	Name	Label	Type	Format
	aar	AAR	int	%8.0g
	akt_lopenr	AKT_LOPENR	str12	%12s
	akt_sy_type	AKT_SY_TYPE	str1	%9s
	aksje_orgnr	AKSJE_ORGNR	long	%12.0g
	aksje_id	AKSJE_ID	long	%12.0g
	aksje_aksjeklasse	AKSJE_AKSJEKLASSE	byte	%8.0g
	realisasjon_id	REALISASJON_ID	long	%12.0g
	erververt_dato	ERVERVET_DATO	str10	%10s
	erverv_type	ERVERV_TYPE	str4	%9s
	realisert_dato	REALISERT_DATO	str10	%10s
	realisasjon_type	REALISASJON_TYPE	str4	%9s
	realisert_antall	REALISERT_ANTALL	double	%10.0g
	realisasjon_anskaffelses_verdi	REALISASJON_ANSKAFFELS...	double	%10.0g
	realisasjon_samlet_vederlag	REALISASJON_SAMLET_VED...	double	%10.0g

- *realisasjon_anskaffelses_verdi*: Purchase amount of realized shares
- *realisasjon_samlet_vederlag*: Realization amount of shares

Firm-Level Corporate Reporting (*RF-1086*)

Figure B7 presents a snapshot of the variables we use from the annual firm-level corporate reporting (*RF-1086*). The file includes one line per firm–year. We use the following variables from this file for our analyses:

- *aar*: Reporting year
- *orgnr*: National firm identifier (*organisasjonsnummer*)
- *oppgave_id*: Submitted corporate reporting identifier
- *oppgave_status*: Submission status of the corporate report (submitted, entered the system, correction). In case of corrections, a firm can submit several reports for the same reporting year.
- *aksje_id*: Unique share identifier allocated by the tax authority
- *aksje_aksjeklasse*: Share class
- *aksje_tot_ant_0101*: Total shares outstanding as of 01/01 of the reporting year
- *aksje_tot_ant_3112*: Total shares outstanding as of 12/31 of the reporting year

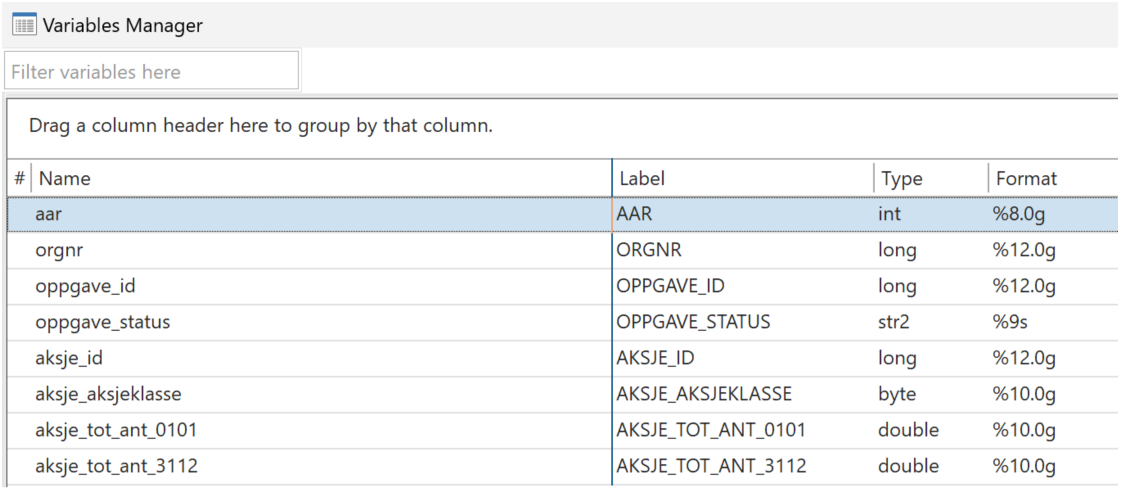
Data Processing

While the corporate information and annual financial statements from the business registry are straightforward and include unique firm–year-level observations, the raw tax records are received in multiple files and require extensive preparation for our analyses.

In the first step, we uncover holding firms in all files. We identify an operating firm's ultimate owner if we observe that the reporting investor is a nonoperating firm and 100% of the shares in that firm belong to only one individual investor. This implies, that we combine the operating characteristics of the lowest holding layer with the ownership of the highest holding layer in one observation. Next, we hand-collect the names and firm

FIGURE B7: **Relevant Variables in the Firm-Level Corporate Reporting (*RF-1086*)**

Figure B7 presents a snapshot of the variables we use from the annual firm-level corporate reporting (*RF-1086*). The file includes one line per firm–year.


Variables Manager

Filter variables here

Drag a column header here to group by that column.

#	Name	Label	Type	Format
	aar	AAR	int	%8.0g
	orgnr	ORGNR	long	%12.0g
	oppgave_id	OPPGAVE_ID	long	%12.0g
	oppgave_status	OPPGAVE_STATUS	str2	%9s
	aksje_id	AKSJE_ID	long	%12.0g
	aksje_aksjeklasse	AKSJE_AKSJEKLASSE	byte	%10.0g
	aksje_tot_ant_0101	AKSJE_TOT_ANT_0101	double	%10.0g
	aksje_tot_ant_3112	AKSJE_TOT_ANT_3112	double	%10.0g

identifiers of all Norwegian and foreign VCs registered in Norway and identify them among the investors in all files. For our purposes, VC investors comprise venture capital (traditional, corporate or government-affiliated) funds, early-stage investment funds associated with traditional PE groups, and incubators.

In the purchase file, we drop share buybacks by the firm itself, resulting in 97.7% of all purchases being ordinary shares. Next, we identify duplicate entries in the tax records if we observe multiple subsequent share purchases of the same purchase type by the same investor made in the same firm on the same date in the same share class and reported in the same calendar year. Each entry has a time stamp, and the distribution of time between two entries reveals that the entries are made either a couple of seconds apart or on the next day. If two transactions are entered subsequently after each other with a difference of fewer

than 10,000 seconds (the median difference is 12 seconds), we assume they are two entries representing two parts of one transaction. In this case, we aggregate the two entries into one transaction. Otherwise, we assume the later one to be a correction and keep it in a given reporting year. Because investors must report all unrealized share purchases annually, we keep the latest observable reporting year of each transaction. Finally, we drop entries with a negative or zero number of shares (0.8% of the share purchases), any share purchases after the year of the firm's first exit event (47.5% of observations) and also in the exit year itself in case of an IPO, merger or acquisition (3.1% of observations). The realization file is processed analogously. In addition, we drop seemingly erroneous transactions for which the realization date falls before the purchase date or the realization amount is negative. In addition, here, 99.0% of all realizations are made in ordinary shares. Since the realization file also contains share purchase information, we double-check whether any purchases are not entered in the purchase file. This is not the case.

The firm-level corporate reporting provides the number of shares outstanding of each firm at the beginning and end of each calendar year. We drop entries with a missing or negative number of shares. In cases with several entries in the same reporting year, we keep the latest one. In this file, firms also report 98.5% of shares as ordinary shares. We conduct a couple of robustness checks here. First, we compare the number of shares outstanding from these tax records with those annually reported to the business registry. Second, we manually calculate the end-of-year number of shares outstanding by adding all newly issued shares from the purchase file, as reported by investors, to the beginning-of-year number of shares outstanding from the firm-level corporate reporting file, as reported by the firms.

Finally, we combine the beginning-of-year number of shares from the firm-level

corporate reporting file with the information on newly issued shares from the purchase file to calculate the number of shares outstanding after each primary share purchase. This allows us to calculate the purchased ownership in each individual transaction and the firm-level ownership sold in a financing round.