Norwegian School of Economics
Bergen, Spring 2021

Predicting Private Equity Enterprise Multiples using Coarsened Exact Matching

Ole Bø-Pedersen & Henrik Finsrud

Supervisor: Carsten Bienz

Master thesis, Econonomics and Business Administration,
Financial Economics

NORWEGIAN SCHOOL OF ECONOMICS

Contents 2

1. Abstract

This thesis investigates if it is possible to predict accurate and unbiased Net Asset Values for private equity (PE) portfolio companies using multiple valuation. The study is motivated by PE research that has found that General Partners (GPs) under certain circumstances have incentives to exert opportunistic valuations, made possible by the structure of institutional PE where Limited Partners (LPs) rely solely on the self-reported interim Net Asset Values (NAVs) from GPs.

First, we construct a novel time series dataset with quarterly company level data for 141 exited portfolio companies in Argentum's Nordic buyout portfolio from 2002-2020. Second, we gather equivalent data for publicly traded companies in the Nordics and ultimately consolidate the two datasets. We then match portfolio companies with comparable public peer's contingent on PE selection criteria, using the matching algorithm Coarsened Exact Matching (CEM). The objective is to test if statistical matching methods in combination with prediction models are able to identify representative Nordic peers and enterprise multiples that can be used to indicate unbiased Fair Market Values for portfolio companies given underlying market conditions. We measure the performance of predictions against each portfolio company's corresponding exit transaction value.

Our findings show that particularly one of our prediction models exhibit consistency and seems to predict NAVs with similar accuracy as the GP when moving further than six months prior to exit. There is a large increase in the GPs prediction accuracy between twelve and six months before exit, which is in line with our expectations given GPs informational advantage near exit. In summary, our results suggest that our best performing specification using CEM may provide a consistent and valid second opinion on the Enterprise Value of portfolio companies.

In the final section, we explain model limitations and discuss applicableness. Although the peer median model predicts enterprise values with similar aggregated accuracy as the GP in certain periods, it is still frequently inaccurate on company level, and contingent on relatively strict criteria that prune observations. Further, there are confounding variables that we are unable to capture during matching, which would likely have facilitated better prediction accuracy had they been included.

2. Acknowledgements

This thesis marks the conclusion of our Master's degree in Economics and Business Administration at the Norwegian School of Economics. The topic of private equity buyouts was chosen due to our mutual interests for the asset class and for corporate valuation. We extend our sincere gratitude to Argentum who provided us with a unique opportunity to learn more about private equity in all aspects, especially from an institutional standpoint as we gained access to their unique database. It was fascinating to delve into the specifics of fund management and the numbers behind the portfolio companies that comprise Argentums successful buyout portfolio. We also thank Argentum and particularly their secondary team for their valuable insight throughout the process.

We would also like to offer our gratitude to Associate Professor Carsten Bienz, who helped cultivate the thesis' focus and guided and supported us along the way. His feedback has been invaluable for the thesis.

3. Introduction

Value is arguably the ultimate measure in financial economics since it addresses the most important question for all investments, the relationship between risk and reward (Koller, Goedhart and Wessels, 2010). Investors expect to be compensated for the level of risk they take on and are thus met with the fundamental questions of what value is and subsequently how to measure it. Further, value estimates from market participants are important for the functionality of capital markets as they influence portfolio decisions and consequently asset prices. Although some are convinced that value lies in the eyes of the beholder, market participants generally agree that intrinsic value, the present value of future cash flows, is the relevant measure for financial assets (Damodaran, 2011)

The emphasis of asset valuation in academia is primarily on intrinsic (absolute) valuation approaches that determines the value of an asset by the present value of its expected future cash flows. The most common models for absolute valuation are the Discounted Cash Flow (DCF) model and the Dividend Discount Model (DDM) (Berk and DeMarzo, 2017). The difficulty associated with valuing an asset varies substantially across all securities and the process is often a mixture of art and science. Absolute valuation approaches are often sensitive to substantive assumptions, especially since they deal explicitly with the uncertainties of the future. As a consequence, it is often complemented or replaced by relative approaches that determines the value of a firm by comparing it to comparable firms, "comps". Instead of valuing a firm's cash flows directly, the relative approach estimates the value of a target firm based on the value of comps that are expected to generate similar cash flows in the future (Berk and DeMarzo, 2017). Relative multiple valuation can be described as taking the ratio of either equity- or enterprise value to a value driver like earnings or sales and applying it to a comparable firm.

Although relative valuation may not be the primary focus in business schools, a study by Pinto, Robinson and Stowe (2015) shows that using multiples is the most common approach for professionals to evaluate individual equity securities. The beauty lies in the simplicity: it is convenient, easy to understand, and it reflects the current market sentiment, which may be valuable to get a feel for. Whilst the concept is simple on the surface, identifying proper peers and using them correctly is a profoundly complex process. Both methods have strengths and weaknesses, and a prudent investor should strive to perform both to "*[...] form your own* *opinion, and then test it against the market*" (Metrick and Yasuda, 2011)Whereas public equities are traded on exchanges and priced every day, privately held assets are unquoted and often more challenging to value. In institutional private equity context, the quarterly reports from General Partners (GPs) update investors (Limited Partners, LPs) on the outlook and value of their portfolio. From initial investment until realisation, these estimated interim Net Asset Values (NAVs) together with accumulated fund distributions, make up the key measure for investors evaluating the performance of their private equity portfolio. The *actual* performance is only known at final realisation, which may be a decade into the future (Metrick and Yasuda, 2011). The quarterly valuations are self-reported and subject to considerable discretion from the General Partners. This subjective component motivated Jenkinson, Sousa and Stucke (2013) to investigate the *fairness* of private equity valuations, and found that there are certain conflicts of interests that can give rise to "opportunistic valuations" by GPs.

Private equity professionals in most cases have equity stakes in their own funds, helping to align incentives between investors and fund managers (Ivashina and Lerner, 2016). Furthermore, private equity payoff structures incentivize performance beyond the equity stake, as common schemes such as the "Two and Twenty[1]" boosts fees when a pre-specified hurdle is surpassed (Metrick and Yasuda, 2011). While this is true, the conflict of interests referred to is not in the context of *actual* performance, but in the context of *reported* performance. As each fund has a finite lifespan, private equity firms need to continuously raise new ones to ensure vital future revenue. A study by Chung et al. (2012) shows that the performance of a current fund has a direct effect on the GPs ability to raise a successor fund. This is important because follow-on funds are typically raised well before current funds are fully realized. The marketing of follow-on funds is thus based on performance measured partly by the unrealized assets, which oftentimes make up the majority of the portfolio value at that time (Jenkinson et al., 2013). This creates a conflict of interest because it is favourable for fund managers to present positive interim performance numbers during this fundraising period. Other potential instances of opportunistic valuation include limiting asset impairments during market turmoil, or smoothing returns by consequently providing conservative estimates as a strategy to avoid negatively surprising investors at realization (Jenkinson et al., 2013).

[1] Two and Twenty compensation agreement: 2% annual management fee of committed capital and 20% of profits ("carried interest") over a minimum return rate ("hurdle rate").

3.1.1 Argentum Asset Management and Research Question

This thesis is written in collaboration with Argentum, a specialised private equity fund investor designated to manage the Norwegian Government Wealth Fund for investments into unlisted equity in primary and secondary markets, as well as through direct co-investments. Argentum has invested in private equity since 2001 and has thus evolved parallel to the asset class and become a leading fund investor in Northern Europe[2]. The specialised PE investor has EUR 1.6 Bn in committed net capital across more than 180 funds, split approximately into 81% buyout capital and 19% venture capital (VC). Their core focus is small and mid-cap funds in the Nordics, with an expanded investment area in Northwestern Europe.

We investigate if it is possible to calculate accurate and unbiased Net Asset Values for portfolio companies using relative valuation. The research question was developed in collaboration with Argentum, who were particularly interested in testing the interim NAVs from GPs against an unbiased market-based estimate. We have signed non-disclosure agreements and been granted access to Argentum' entire database with historical reporting from all their 180+ fund investments. This entails that the data is strictly confidential, and that descriptions and results are anonymized.

The scope of the thesis is limited to one of Argentum's core focuses, Nordic buyout funds[3]. We have constructed a novel time series dataset that consist of company level data for all Nordic buyout portfolio companies held by Nordic GPs from 2002-2020. We manually extracted quarterly trading data, capital structure details, NAVs and various qualitative data for the entire holding period of all 141 successfully exited companies in our sample held in 33 different funds. The database consists of two subsamples since some funds only report on an annual basis[4] (45 of the portfolio companies). The construction of a detailed private equity database represents a substantial part of our contribution. The second dataset contains equivalent quantitative data for publicly traded companies in the Nordics from 2002-2020 gathered from Refintiv Eikon Datastream, which was ultimately consolidated with the private

---

[2] For more about this, see "*The state of Nordic private equity 2020"* (Argentum, 2020).

[3] See section "Private Equity Segments" for rationale.

[4] See the data section where we explain the differences in financial reporting.

equity database.

We evaluate if statistical matching methods in combination with two separate prediction models are able to identify representative Nordic peers and enterprise multiples (EV/EBITDA) that can be used to indicate an objective Fair Market Value (FMV) given underlying market conditions. We deploy the matching algorithm Coarsened Exact Matching (CEM) to assign publicly listed peers to each portfolio company in our sample. The matched peers' median enterprise multiple is then used to predict the portfolio company's interim and exit mark-to-market value. We also employ a regression model where the matched peers' enterprise multiple is OLS regressed on relevant predictors to estimate an equation for each portfolio company to predict its enterprise value.

The ultimate test is naturally at exit when the companies are realized, but we expect biased results when evaluating our predictions against the GPs at exit, because their near-exit estimates will often be based on indicative offers from potential buyers[5]. Therefore, we test our interim predictions against the estimates from GPs with particular focus on the quarters within one year of exit, since these are more likely unbiased while also testable against the actual transaction price (due to their proximity to exit). For the annual subsample we focus on the two years prior to exit. To evaluate the performance of the predictions we use the *Mean Squared Error* (MSE) measure, in line with e.g., James, Witten, Hastie and Tibshirani (2013). Logically we expect consistency in the median model performance as it relies exclusively on the quality of the CEM matches and inherently deals well with outliers. The same is not true for regressions which we expect to perform substantially better when we implement measures to deal with outliers. In line with Bernström (2014), we expect that applying a systematic marketability (liquidity) discount to the estimated multiples increases the accuracy of the predictions.

Our research question is *"Private equity research has found that General Partners have significant discretion in determining asset values, thus we investigate if a multiple based statistical approach is able to deliver unbiased and accurate valuation results."*

[5] This is frequently communicated in quarterly reports.

The primary motivation of the study is to create a practical tool that Argentum can use on live portfolio companies. The tool itself is comprised of the private and public equity databases, a Stata do-file with matching specifications, 22 sub-routines comprising 1720 lines of code written from scratch in Microsoft's programming language "Visual Basic for Applications" (VBA) used to process and transform the datasets for analysis, and finally an additional do-file with regression specifications and Stata packages for each portfolio company and its peers.

There are not many similar studies in academia as detailed proprietary PE datasets are not readily available, primarily due to confidentiality considerations. Furthermore, portfolio valuation tools that can be used for in-house purposes are naturally not publicly available. Argentum has been invested in Nordic buyout funds for a long time and our sample is thus thorough, covering an estimated 20% of the universe of our study[6]. We hope to contribute operationally for Argentum and academically by studying the uncharted territory of data-driven private equity valuations. The thesis is interesting for anyone interested in private equity, corporate finance, valuation, prediction and portfolio management.

3.1.2 Results

We calculate the mean squared error for each prediction in order to evaluate our results. The MSE is then averaged and aggregated for each model specification and iteration. Mean squared error is not meaningful in isolation and must be evaluated relative to the performance of other model specifications or GP estimates. The results show that for the quarterly subsample, our median model exhibits consistency in its predictions. It also seems to predict at least at the level of GP accuracy when moving further than six months prior to exit. In line with our expectation, there is a large spike in GP accuracy between twelve and six months before exit, likely because the GP receives an indicative offer or has entered negotiations. This finding is supported by the annual sample, which also shows that GP predictions are more accurate closer to exit. The median model exhibits consistency for the annual subsample as well, however it should be noted that since this sample is relatively small it is more prone to random noise. Further, for both subsamples the regression models performs seemingly

[6] Our defined universe = Exited deals from Nordic buyout funds (managed by Nordic GPs) from 2002-2020. Calculated with data from Preqin.

consistent after we implement measures to deal with outliers, although not particularly accurate in relative terms. We also find support for our expectation that leveraging knowledge of variables when *coarsening* for matching seems to help avoid assigning poor matches. We emphasize the results from our quarterly sample as this is the largest sample and has the most frequent financial reporting. In summary, the results suggests that our best performing specification using substantive knowledge on variable coarsening may provide a consistent and valid second opinion on the Enterprise Value of portfolio companies.

The thesis is structured as follows. In the next section we outline background information on private equity and limited partnerships, and how our study relates to existing literature. Section III explains the peer selection frameworks we employ. Section IV provides a detailed explanation of the data. Following the data section, we provide our methodological analysis. Finally, section VI discusses results and concludes.

4. Background and Related Literature

4.1.1 Private Equity and Limited Partnerships

The universe of our study is institutional private equity, which excludes the majority of private companies since they are not investable for an institutional investor due to e.g., size and financial constraints (Døskeland & Strömberg, 2018). Private equity in institutional context refers to equity investments in unlisted firms by professional investors. The common structure of institutional PE is through a private limited liability partnership, with capital invested by Limited Partners and managed by a General Partner. The GPs represent professional financial intermediaries often referred to as private equity firms. The contractual term of the limited partnership (fund) is typically ten years with fund extension options ranging from one to three years (Kaplan and Sensoy, 2015). The LPs commit capital which is typically drawn by the GP over a five-year investing period while attractive target firms are identified. As previously mentioned, GPs need to raise follow-on funds to secure future revenue. The interval between any subsequent funds depends on the success of the predecessor and typically ranges from two to seven years (Jenkinson et al., 2013). The opportunistic valuation issues that may arise during fundraising in this interval is part of the motivation for our study.

Despite the potential opportunistic valuation issues that may arise due to the illiquidity of PE, industry advocates argue that the asset class facilitates an advantageous long-term value creation perspective, absent from the short-term pressures of liquid markets (Koller et al., 2010; NY Times, 2012). Furthermore, PE focuses on alignment of interests between ownership and management as the latter are also expected to invest in the portfolio companies (Ivashina and Lerner, 2016).

4.1.2 Private Equity Segments

Private equity can be categorized into *buyout*, *venture capital* and *growth capital*. The thesis focuses on buyouts, which is by far the largest category of private equity (Metrick and Yasuda, 2011). Buyout funds seek to execute control and co-control investments typically in mature mid to large-sized companies in leading market positions with solid cash flows, but with potential for revenue and earnings growth, predominately financed by leverage (Kaplan and Strömberg, 2009). The “buyout” term stems from the tendency for buyout funds to acquire the majority stake in the target company (buying out the owner), thus gaining control over it. Although the objective is to capitalize on the untapped potential of a target company, the model focuses primarily on scaling and margin improvement rather than turning unprofitable businesses around (Kaplan and Strömberg, 2009).

Venture- and growth capital are earlier stages of private equity, with VC being the earliest and growth capital often entailing late-stage VC in profitable firms financed by subordinated debt (Metrick and Yasuda, 2011). Although venture and growth capital make up a substantial part Argentum's portfolio, they are not included in our study. This is because valuing companies using earnings multiples requires a high degree of stability in earnings and cash flows. Due to the early-stage characteristics of venture and growth capital, the companies have not reached maturity and stable growth, and as a consequence are often not fit for comparable multiple valuation (Metrick and Yasuda, 2011).

4.1.3 General Partner Valuation Policy

The International Private Equity and Venture Capital (IPEV) Valuation Guidelines describe broad yet important private equity concepts for the industry to lean on regarding the value of assets held by Limited Partnerships. The guidelines are endorsed by all the GPs in our sample, ensuring at least a theoretical resemblance in the estimation of interim values. Although the guidelines facilitate transparency and standardization in valuation framework principles and emphasizes consistency and comparability, they do not guarantee unbiased valuations. Private equity research by Brown, Gredil and Kaplan (2018) describe opportunistic valuation challenges and find that some underperforming managers inflate reported returns during fundraising of follow-on funds. Using a dataset from the largest U.S. investor in private equity, Calpers[7], Jenkinson et al., (2013) find similar results and warn investors of basing investment decisions regarding follow-on funds on the reported returns of a current fund. Even though fund managers comply with IPEV Guidelines, there is still room to exert subjectiveness that can affect valuations in the GPs desired direction. As a Limited Partner who frequently reinvests in follow-on funds, this is highly relevant for Argentum.

[7] Calpers is short for California Public Employees' Retirement System.

Although observable market prices are not available for unlisted private equity investments General Partners must report interim company values and portfolio updates, as per the limited partnership agreement. Fund valuation policy is typically communicated in quarterly reports where The IPEV methodologies are generally applied. IPEV defines Fair Value in accordance with International Financial Reporting Standards (IFRS) 13 as "*… the price that would be received to sell an asset in an Orderly Transaction between Market Participants at the Measurement Date*". Transparent, independent and credible valuations are increasingly being requested from limited partners (FW, 2014). In spite of this, most of the GPs in Argentum's buyout portfolio value in-house seemingly without consultation from third parties other than adjustments made at fiscal year-end audits. This may however be inaccurate since it is not certain that funds would disclose third-party consultation. Some GPs disclose their valuation principles, although to varying extent, with some observed commonalities. We outline these commonalities and how General Partners value assets given their practical implementation of IPEV Valuation Guidelines, as gathered from their quarterly reports. For the guidelines themselves we refer to IPEVs December 2018 version.

Investments are typically valued at cost for at least the first year, taking the taking the bid spot exchange rate as at the last day in the quarter into consideration (IPEV, 2018). However, if the newly acquired portfolio company's trading is significantly below expectation or there has been adverse changes in market or economic conditions, fund managers state that they write down asset values to reflect the impairments. Following the first year's holding period, a variety of valuation methodologies are applied depending on the asset's characteristics and its market.

GPs valuation of portfolio companies beyond the first year of holding is usually comprised of either a peer group multiple, a sum-of-the-parts calculation, a Discounted Cash Flow, or a combination. Comparable multiples are appropriate earnings or sales multiples for public companies that are comparable especially in industry and size to the investee company (Metrick and Yasuda, 2011). GPs often adjust multiples before applying them to the target firm's relevant accounting measure, which is also often substantively normalized (IPEV, 2018).

Although relative valuation is common, GPs state that they base them on fundamental analysis where company performance, revenue and earnings growth outlook, changes in cash flows and recapitalizations and other capital structure transactions are evaluated. The fundamental analysis serves as the basis from which material impacts on the selection of comparable peers or the applied multiple are considered. Further, corporate transactions executed or pending in the company or for comparable issuers are considered either directly as a multiple or for adjustment, but in line with e.g., Bernström (2014), these are not applied together with quoted companies' multiples. GPs also consider offerings in equity or debt, together with the overall solidity situation, to reflect capital structure characteristics.

Theory suggests that the value of shares in a quoted company should be higher than that of shares in an equivalent private company due to investors preference for liquidity (Bernström, 2014). As a consequence, one would expect the lack of marketability to constitute a discount for private companies to peers. IPEV Guidelines (2018) describe the risk associated with the lack of liquidity and suggests calibrating the applied market multiple with regard to liquidity and other risk factors. In line with findings from Harjoto and Paglia (2010) who investigated the discount for lack of marketability (DLOM) for private companies, we observe that most GPs who disclose their valuation methodologies indicate a discount, typically ranging from 10-20% to the weighted peer enterprise multiple. Some GPs employ marketability discounts on portfolio level, whereas most assess the risk associated with lack of marketability individually[8].

On the other hand, it is theorized by control premium theory that firm value increases when owners acquire a controlling share, which is especially relevant in the buyout context. Control is advantageous because there is arguably value in being able to run a company differently and better than comparable companies (Damodaran, 2005).

To demonstrate the extent of GP discretion in valuation using multiples, we portray a simple example of how a portfolio of buyout companies was actually valued as disclosed in an extended quarterly report. The GP chose 3-15 peers for each portfolio company with earnings multiples gathered from the two most recent years. The implied multiples and years were then assigned weights, which were then adjusted for a marketability discount that resulted in a final

[8] This is consistent with the research from Harjoto and Paglia (2010) who found that discounts vary substantially across industries.

weighted enterprise multiple. The implied Enterprise Value was ultimately adjusted for Net Debt to derive the weighted Equity Value[9]. As the approach was consistent with their previous valuations and satisfied general principles, it is thus complaint with IPEV Valuation Guidelines. This demonstrates that there is a vast range of equity values for a given investee company. In another instance two GPs from our sample had jointly invested in a portfolio company and valued it substantially different, where the difference in estimated value at one point exceeded the total estimation from the conservative GP.

Further, regulators are concerned about the discretion that GPs has in picking the comparables, particularly that they "cherry pick" comparable public companies (Grant Thornton 2021; Clifford Chance 2020). This entails that GPs could theoretically create comparable sets based on performance in trading and stock price, pulling out poor performers to boost the applied group multiple. Comparable transaction multiples could in theory also be subject to cherry picking, which can be as impactful as multiples from public companies.

[9] Equity Value = Enterprise Value – Net Debt

5. Peer Group Identification

The selection of comparable sets to draw inference from is likely to be the most critical step in terms of impact on ultimate prediction from the valuation tool. This poses important questions regarding what comparability is, as well as the overall process of selecting comparable quoted companies. Before we identify the peers, we thus need to outline comparability in the context of relative valuation. This is done by discussing multiples and reformulating the enterprise multiple to address its value drivers, such that we can discuss the characteristics that public companies have to exhibit similarity in for comparable valuation to be accurate. We then move on to matching. For this purpose, we employ matching algorithms with several iterations and recalibrate to improve the ultimate tool performance. Under *Matching* we first explain data preprocessing through matching, its goals, our methodology for peer identification - Coarsened Exact Matching (CEM) and ultimately variable selection.

5.1 Muliple Selection and Value Drivers

Mark-to-market peer group multiples is the most common methodology used to indicate the value of mature private equity assets, at least in the context of quarterly reporting. This is arguably because of its simplicity and due to the financial structuring in private equity. Although discounted cash flow analysis is more accurate and flexible, it is often a tedious task that requires high precision since it relies heavily on forecasts (Koller et al., 2010). Taken into the context of our study, present discounted valuation is not a viable option since it has the inherent bias from forecasting and cannot be fully automated.

Further, it is important to emphasize the manner of which comparable multiples are calculated. The point is that there must be consistency between the numerator and denominator to avoid bias: value must be paired with the corresponding income source. For instance, the numerator of enterprise multiples must include the market value of both equity capital and debt (i.e., all investor capital), while the denominator must include income to all investors, both shareholders and debt claimers (Bernström 2014). If consistency is not achieved, the multiple will be biased in a direction depending on the over- or understatement of either the numerator or denominator (Koller et al., 2010).

Many industry professionals use price multiples, especially the price-to-earnings multiple, calculated by:

$$\frac{P}{E} = \frac{Price\ per\ share}{Earnings\ per\ share}$$

For our purpose, P/E multiple has two primary flaws: first, that it can only be used for entities that have similar capital structure, and second that it is calculated before nonoperating items such as one-off losses or gains, which may be significant, and are important to capture (IPEV, 2018; Koller et al., 2010). Thus, it may provide values that does not reflect the financial reality.

Enterprise Value multiples, on the other hand, removes the influences of capital structure, a feature that is essential for our study with an observational dataset of leveraged buyouts. EV multiples take the entity market value in relation to an appropriate base metric, such as *revenue*, *EBITDA* (earnings before interest, taxes, depreciation and amortization), *EBITA* (earnings before interest, taxes and amortization) or *FCFF* (free cash flow to firm) (Bernström 2014). Literature generally agrees that for most purposes EBITDA or EBITA, which are calculated after nonoperating items, in combination with EV are the most appropriate metrics to compare peer valuations e.g., because they minimize accounting differences (Koller et al., 2010). Using EBITA over EBITDA may be argued since depreciation might be essential to comprehend certain companies' value. This is especially important for industries where depreciation represents a precise predictor of a firm's capital expenditure (capex) in the future (Koller et al., 2010). However, EBITDA is generally more accurate when depreciation does not provide consistent estimates of future capex. In the equity valuation paper by Pinto, et al., (2015) EBITDA is further supported by the fact that it is the most widely used metric in combination with EV by business professionals. This is also anecdotally supported by our findings, where all 141 portfolio companies include EBITDA in the financials, and only a few include EBITA, indicating the former as the preferred choice also from a practical standpoint.

The issue of bias from forecasting should also be addressed for multiples, since many industry professionals substitute the latest fiscal year earnings with forecasts for the following year(s). In line with Metrick and Yasuda (2011), we use recent historical financials since we want to maintain the notion that the relative valuations should reflect the markets opinion of company value. We therefore ensure that the financials we gather are not forecasts, which are often also included in the private equity fund reports (see data section). By gathering Enterprise Value as the sum of market value of Equity and Net Debt, and the most recent historical EBITDA for public companies, we ensure that the derived enterprise multiple reflects the market opinion. Koller et al., (2010), on the other hand, suggests the opposite, that forward-looking multiples should rather be applied. The main argument is that forecasted estimates have better empirical evidence of accuracy (also noted by Metrick and Yasuda), which although true, is not a compelling argument for our application since we aim for unbiasedness. The somewhat self-explanatory, but still relevant downside of focusing on market valuations, is that the market sentiment may be wrong, thus leading to over/undervaluation. Such questions are addressed when we evaluate our predictions against actual exit realizations (see Methodological Analysis and Results sections). The market multiples are calculated from the price quotations of EV and interim earnings gathered at the relevant last day of quarter/year from Datastream. These are applied such that portfolio company Enterprise Values are calculated by:

$$Enterprise\ Value_{portco} = \frac{Enterprise\ Value_{public}}{EBITDA_{public}} * EBITDA_{portco}$$

To better grasp the dynamics of the multiple, we show how it derives primarily from the value drivers profitability, growth and risk. This is done by reformulating the equation under guidance of Damodaran (2012). Starting with EV:

We can reformulate this into:

$$\frac{EV}{EBITDA} = \frac{(1-T)}{WACC-g} + \frac{Depr/EBITDA}{WACC-g} - \frac{Capex/EBITDA}{WACC-g} - \frac{\Delta Working\ Capital/EBITDA}{WACC-g}$$

Where,

$WACC = Weighted\ Average\ Cost\ of\ Capital$

$g = Growth\ rate$

$T = Corporate\ tax\ rate$

$Capex = Capital\ expenditure$

$Depr = Depreciation$

$\Delta Working\ Capital = Change\ in\ Working\ Capital$

From the last equation it is evident that the key value drivers of EV are **profitability** (through EBITDA, corporate tax rate and depreciation), **growth** (growth rate) and **risk** (cost of capital) (Damodaran, 2006). This is important to recognize because it clarifies the factors that companies have to exhibit similarity in for enterprise multiple valuations to be applicable. The derivation of the value drivers of EV/EBITDA thus has consequences for how we make predictions, which we emphasize through both matching and model specifications.

5.2 Matching

Matching is a method used to control for confounding influences of pre-treatment covariates and thus addresses selection bias to enable causal inference from an intervention. Selection bias refers to bias in predictions or estimates caused by endogenous sample selection (Woolridge, 2013). We pair non-treated participants with treated participants conditional on similarity in important characteristics, such that the differences in outcome between the groups can be attributed to the treatment. These characteristics, referred to as pre-treatment covariates or confounders, are covariates that have to influence both participation (PE selection) as well as the outcome variable (some valuation metric), without being affected by treatment (thereof “pre-treatment”) (Rosenbaum 1984; Caliendo and Kopeining, 2008). The matches are therefore preferably assigned at entry (which varies for all portfolio companies) and kept until exit. Matching methods are preprocessing algorithms, and statistical estimation is the typical route post-matching to make causal inferences. Regular estimation is however not our objective, our purpose is to provide unbiased matches whose values are used for prediction. Through matching we ideally identify peer companies that are identical to the portfolio companies in all regards relevant for predicting enterprise multiple, except not being acquired by private equity. We evaluate the quality and applicableness of the matched peer’s ex post, by running tests tailored for our special application. (Iacus, King and Porro 2012).

5.2.1 Goals of Matching

The principal goal of matching is to prune observations from observational data to achieve more balance between the treatment and control group, entailing that there is more similarity in the empirical distributions of the variables between the groups (Iacus et al., 2012). We employ CEM and match on substantively coarsened variables which creates a perfectly balanced dataset. Coarsening represents a trade-off between the number of observations and model dependence, because reduced model dependence achieved by perfect balance can imply fine-grained coarsening which prunes a lot of observations (Iacus et al., 2012).

5.2.2 Notation and Quantities of Interest

In our dataset T_i is a dichtomous treatment variable for unit i $(i = 1, \ldots, n)$ which has value 1 if it is part of the treatment group and 0 if its is untreated and part of the control group. The treatment is whether the company i in our merged database is a PE portfoltio company. The dependent variable Y_i represents a valuation metric such as EV/EBITDA, although not practically important since we are not estimating a treatment effect. Nontheless it is useful to outline the notation and theoretical quanteties of interesent to better explain the framework. X_i represents the relevant pre-treatment covariates that we have extracted from quarterly fund reports and Datastream, such that the theoretical esimated Treatment Effect for treated (T_i=1) observation i would be equal to $TE_i = Y_i\ (T_i = 1) - Y_i\ (T_i = 0)$, where the unobserved counterfactual $Y_i\ (T_i = 0)$ is estimated from the matched X_i controls. The total Sample Average Treatment Effect on the Treated is equal to $SATT = \frac{1}{n_T} \sum_{i \in \{T_i = 1\}} TE_i$.

To get unbiased estimates post-matching, we also require the ignorability assumption to be satisfied, that there are no omitted variables correlated with both the dependent and independent variables. (Iacus et al., 2012).

5.2.3 Coarsened Exact Matching

Coarsened exact matching was introduced by Iacus, King and Porro in their article[10] on causal inference from matching that was published in *Political Analysis*. CEM approximates a fully blocked experiment and thus achieves an exactly balanced data pool without necessarily requiring a large sample[11]. This can be conceptualized as follows. Whereas the standard experiment design, complete randomization, flips a coin for each observation to determine

[10]See "*Causal Inference without Balance Checking: Coarsened Exact Matching*", latest version released in 2011.

[11] Randomized experiments often require large samples to achieve balance (King, 2018)

treatment, a fully blocked experiment matches (blocks) on the covariate(s) of interest and then flips a coin for each pair (King, 2018).

CEM handles well the “the curse of dimensionality”, that exact matching on several variables may produce few matches because two observations are not likely to be identical across all of them (Ho et al., 2007). We temporarily coarsen into substantively meaningful bins determined at our discretion or by the CEM default binning algorithm, then exact match on the coarsened data, and ultimately move on to prediction with the original uncoarsened data for the observations that were matched. CEM is a monotonic imbalance bounding method, entailing that reducing imbalance (more coarsening) on one covariate has no impact on others. Making balance decisions ex ante is preferable to the manual process of reestimating and adjusting the model to achieve a certain maximum imbalance, which is the case for other matching methods. We do however experiment with varying levels of coarsening to evaluate differences in ultimate prediction power.

5.2.4 Alternative Methods

Propensity score matching is a common way of identifying the counterfactual in private equity literature. Introduced by Rosenbaum and Rubin (1983), PSM estimates the conditional probability of treatment given specified covariates (Caliendo and Kopeinig, 2008). Instead of matching on coarsened covariates, it calculates the probability of treatment given a vector of the covariates, often assigning nearest-neighbor[12] matches conditional on the score. The logic of covariate selection and time of measuring is the same for PSM as CEM, but the method does not leave room for substantive decisions on covariates from expert knowledge. Furthermore, since we are creating a tool to be used for live deals, it is advantageous that *“CEM is faster, easier to use and understand, requires fewer assumptions, is more easily automated, and possesses more attractive statistical properties for many applications than existing matching methods”* (Blackwell et al., 2010). We refer to the implementation section for a thorough explanation of the practical implications of using CEM.

[12] There are several methods for matching propensity scores other than nearest-neighbor, see Caliendo and Kopeinig (2008).

6. Data

The novel dataset constructed from manual collection of data from Nordic PE deals done by Nordic GPs, combined with a dataset of public companies, create the basis for our analysis. In this section, we explain the datasets by providing an overview of the data gathering process and adjustments made to collected variables with the purpose of making data operational. First, we introduce the PE portfolio of Argentum and explain how data from quarterly reports were used to create our buyout dataset. Second, we show how Refinitiv Eikon Datastream was used to gather a sample of listed companies to be used for matching. Finally, we evaluate the two datasets in context of each other, identifying additional areas for cleaning and the creation of new metrics in both datasets. For this reason, both datasets will be presented in the end of the chapter.

6.1 Private Companies

6.1.1 Source

Argentum

The source for our buyout sample is the quarterly reports of PE funds currently or previously in the PE portfolio of Argentum as of February 2021, which makes Q3 2020 the latest reports accessible. In the database we have gained access to, we identify 182 different funds managed by 94 GPs. On their website, Argentum reports being invested in 187 funds, meaning we are able to cover nearly all of Argentum's fund investments [13]. Of our total, 148 are identified as buyout and 34 as venture. The majority of funds are focused on the Nordics and Northwestern Europe, while there is also exposure to funds with focus on global opportunities and Southern Europe.

Portfolio Company Information in Private Equity Quarterly Reports

While quarterly reports of public companies are relatively uniform, quarterly reports of private equity companies are more diverse in both form and content. This has made the data gathering

[13] There is only one fund in the online portfolio list which we are unable to find in our database, suggesting that the rest of the difference is due to funds not reported either due to being too new or for other reasons

challenging as not all metrics or information that we would prefer to collect for our analysis is available for all funds. Thus, in order to aggregate the data, we resort to metrics which are common for the majority of funds. For this reason, we give a brief explanation of what information about the portfolio companies is usually found in a PE quarterly report.

Most reports start with a fund summary page, explaining the fund terms and focus of the fund, notably mentioning geographic scope and investment stage of the fund. This is followed by a valuation section providing the stated FMV of the specified ownership stake the fund has in each portfolio company. FMV includes the value of all equity, including common shares, preference shares and shareholder loans.

Later, there is usually a section for each unrealized portfolio company providing qualitative information about geography, industry and market outlook, in addition to more in-depth information about its valuation, capital structure and trading for the current quarter or year[14], and often for a few periods back. Information about trading is typically not available in the report for the quarter of the respective portfolio company's entry and exit. Similarly, pre-entry metrics are rare, with relatively few GPs providing revenue and EBITDA, and close to none provide Net Debt.

Valuation and capital structure related metrics include FMV, Net Debt and EV. Often, one of the mentioned metrics will be missing, but there is usually sufficient information to calculate it. For trading, a quarterly or Year To Date (YTD) sales metric such as revenue or sales is included, as well as quarterly or YTD profit metrics such as EBITDA or EBITA, but occasionally only one is provided. Both net debt and the trading metrics are often normalized, but it differs between GPs to what extent this is elaborated on and if there is information about it at all. Some GPs also include information about the portfolio companies' balance sheets, relevant peer groups, estimates of future sales and margins and the valuation methods used. If provided, the balance sheet may be used to find information about potential shareholder loans,

[14] Potentially non-standard Fiscal Year

as well as backing out Net Debt subtracting cash from total debt in cases where Net Debt is not explicitly stated elsewhere.

While most GPs release quarterly reports with updated metrics for all portfolio companies, some release more infrequently. This can be either once every year or half-year, the latter at either odd or even quarters. There are also cases of GPs releasing fund reports quarterly, but updating the portfolio companies' valuations more infrequently.

Alternative Sources Considered

The restrictions in accounting data posed by the use of quarterly reports exclusively, meant that we were limited to key financial metrics and relatively few balance sheet items. Some GPs might include more detailed data about pre-deal information in confidential data rooms that Argentum is granted access to during fundraising, but investigating this would be outside the scope and time constraint of the thesis and likely introduce bias as the depth of the data rooms might vary substantially.

For this reason, we have researched both academic and commercial external sources to find sources that could provide more financial metrics and particularly financials for the portfolio companies pre-entry. For Norwegian private companies registered in the company registry Brønnøysundregisteret, it is possible to get detailed information of annual accounting data from Proff Forvalt back to the 1990s. Similarly, Rakner & Rasmussen (2013) also working with Argentum data, were able to use the SNF's and NHH's database of accounting and business information to extract accounting information for Norwegian buyout companies. However, equivalent sources for a Nordic sample were not as readily available. For example, the Swedish company register Bolagsverket charges a fee per annual report retrieved.

We also investigated the databases Amadeus and Orbis which contains information on private companies, both by business information provider Bureau Van Dijk, but these databases were highly inconsistent. Both were incomplete for Swedish and Danish companies, only containing the last five years of accounting data for Denmark and in general few observations beyond the last nine years. In correspondence with Bureau Van Dijk, we learned that the limitation was due to contractual reasons from their data providers. Finally, the data provider S&P Compustat Global was considered, but it did not seem to have information on medium to small private companies.

As we were unsuccessful in finding adequate alternative data providers for private company information beyond Norway, we decided for consistency reasons to solely use the quarterly reports from Argentum and the financial metrics found there for our analysis. While finding an alternative source for accounting information would have been useful for retrieving pre-entry metrics, it would not fully solve the challenge of few financial metrics as the metrics would be derived from annual reports, thus creating a discrepancy in the frequency of information since most data would still be gathered quarterly.

6.1.2 Data Gathering

Starting with 182 funds identified in Argentum's database, we used fund and company characteristics to narrow the scope to identify the deals relevant for our analysis. On fund level, we used information found in the quarterly report summary together with Argentum's classification of their portfolio on their website[15] to identify and exclude all VC funds. This includes growth, seed and expansion funds. This excluded 34 funds. Of the remaining funds, we used the same approach to identify which funds were Nordic based[16] and thus relevant for our analysis. From a start of 148 buyout funds, 66 funds were left after excluding non-Nordic funds.

After having identified relevant funds, we went through over 800 reports[17], working our way through company level data from the oldest report of a fund and forward. This approach was chosen as we wanted to collect data in a manner of which the tool would be used. This is also important because we frequently encountered situations where newer reports adjusted previous financials slightly, as they for instance later identified one-offs. Not all deals were applicable to our analysis, thus the following considerations were made:

Nordic deals only: In line with our desired focus, non-Nordic deals were excluded based on headquarters, such that only companies with headquarters in Norway, Sweden, Denmark and

[15] https://argentum.no/nb/portfolio/

[16] Defined as having a Nordic headquarter

[17] Only counting the reports where we followed one or more portfolio companies. There were also many reports investigated which did not have any relevant portfolio companies, as per the following exclusions.

Finland were included. We have not excluded companies based on the markets they are operating in, meaning that it is possible for some to have most of their revenue stream outside of the Nordics and in that regard be less of a Nordic company. While information about markets served is relevant for matching purposes (IPEV, 2018), this proved difficult to filter on as there were large differences between whether or not GPs provided this information.

Realized deals only: Our tool can be applied on live deals to match the portfolio companies with public peers, but we need market value realizations to be able to compare how well the peer matching and models perform in predicting exit multiples. Deals not realized by Q3 2020 are thus excluded.

Excluded growth companies: Some of the buyout funds also invest in growth companies, which is mentioned in the fund objective and summary part of the quarterly reports. For consistency in our analysis, as we want stable buyout companies where the EV/EBITDA multiple is applicable, growth deals are excluded as revenue multiples are more relevant for these deals (IPEV, 2018). We identified these portfolio companies by the GPs' classification of the deal, while also checking the equity stake percentage and characteristics to substantiate our choices.

Excluded bankrupt companies: Companies approaching bankruptcy will typically stop being valued by EV/EBITDA multiple and start being valued at other metrics such as NAV (IPEV, 2018), such that our tool is not fitting to value. Our exclusions include both write-offs where the FMV is set to zero as well as where the equity is sold for a symbolic sum.

Excluded companies without sufficient information: Some companies had to be excluded because it was not possible to extract sufficient information about either valuation metrics or trading that were needed for them to be included in the analysis. There were predominately two reasons for this. First, especially older funds showed inconsistencies in what metrics were reported, such that for example Net Debt was missing or EBITA was used instead of EBITDA with no way of calculating the missing metric. Second, a proportion of the funds in Argentum's portfolio originate from secondary market transactions, such that quarterly reports prior to the acquisition were missing. As quarterly reports often show valuation and trading for a few periods back, we were able to extract the necessary information for some companies even without the older quarterly reports, but several were excluded because we were not able to get data for the full lifespan of the companies.

6.1.3 Variables

In accordance with our methodological framework seen in the context of available information from quarterly reports, the following variables were collected:

Table 1 – Variables for PE companies

Variable	Description
Date	Date for the observation reported on a quarterly basis
Fund and Fund Manager	Fund and Fund Manager
Company Name	The newest name of the company
Industry	Industry as reported by the GP
Fair Market Value	The value of all equity in the portfolio company
Net Debt	Market value of Net Debt, often proxied by book value of Net Debt
Enterprise Value	Enterprise Value
Revenue	Quarterly and annually
EBITDA	Quarterly and annually
Currency	Currencies used for valuation and trading

Other variables occasionally used to indirectly acquire the variables in the table were also collected. These include the fund's ownership stake in the portfolio company, shareholder loan, YTD revenue and YTD EBITDA. The combination of often having to find variables indirectly and having many assumptions made which might have implications for the analysis, motivates a more in-depth discussion of selected variables in order to understand how the data is treated.

6.1.4 Qualitative Information

Date

Collecting data for the quarterly sample was always done on a quarterly basis dating the observation to the last day of the corresponding quarter. For the annual sample, we chose to report data as per the last day in Q4, as annual companies tended to be reported in the Q4 annual reports. However, this creates some differences in the exit date as an exit valuation happening earlier in the year will in this way be reported as being exited in Q4. Similarly, entry dates for annual companies is set to Q4 of the entry year.

At entry, we rarely observed trading and capital structure data prior to buyout. As the search for alternative sources to provide this information was unsuccessful, we did not include pre-entry observations in the sample. Thus, the observation period for each company starts at their respective entry periods.

At exit, the GP sometimes keep the company on their balance sheet after company realization if the exit was by IPO or if they expect an earn-out, as there is still value in the company for the fund. We regard the period of the realization event as the last observation in our dataset, as this is the point at which we get an unbiased market value for the company. Partial realizations during the lifespan of the portfolio company are not treated as exits.

Company Name

The portfolio company might change name through its lifespan, either because of a rebranding or due to merging or carve-out. In cases of a rebranding we have used the newest name, but noted and kept track of the older names for reference as the dataset is to be operational for Argentum at a later stage.

Industry

GPs usually report the industry of their portfolio companies, often at two levels where one is the sector and the other one is a more specialized industry designation. We retrieved both, but for comparison purposes, these would later need to be converted to a common industry code (see Industry in Additional Data Processing).

6.1.5 Valuation

FMV, Net Debt and EV

If not stated directly the FMV of each portfolio company was gathered indirectly by subtracting Net Debt from EV. For the majority of companies, EV was stated directly, or alternatively found using the fund's valuation multiple for the company, typically EV/EBITDA multiplied with the corresponding EBITDA used as basis by the GP. In most cases it was possible to verify the FMV of the company by dividing the FMV of the fund's investment in the company by the fund's ownership percentage to get the full FMV. However, this check was not reliable if it was not possible to adjust the fund's FMV for potential shareholder loans, as the fund's ownership percentage usually referred to the ownership of shares in the company on a fully diluted basis. For example, if the majority of the FMV of the fund's investment in the company derived from a shareholder loan, the FMV for the company would be overstated when divided by the ownership percentage. Rather, it would be necessary to subtract the shareholder loan from the fund's FMV, divide the latter by the fund's ownership percentage and then add back the value of all shareholder loans to the company.

If Net Debt was not stated directly and we had the necessary balance sheet information, the metric was found subtracting cash from total debt. In a few cases where both EV and FMV for the company was known, but not the Net Debt, Net Debt was found as the difference between the two. For most companies, the reported Net Debt figure was stated to be normalized, but this was not always clear from the reports. For comparison to be possible, we assume the net debt to be normalized if not explicitly stated.

6.1.6 Trading

Revenue & EBITDA

While some GPs provide quarterly numbers directly, the revenue and EBITDA were most often found through the difference in the stated YTD figure, subtracting the previous period's YTD metric from the current one. In cases where the portfolio company used a non-standard Fiscal Year for its reporting, it was generally possible to use a combination of annual figures, YTD and previous non-standard quarterly figures to get standard quarterly figures.

Similarly to Net Debt, a potential issue regarding both trading metrics is that there are differences in whether they are normalized by the GP or not, and to what extent information about this is given on a quarterly basis. For example, some GPs include one-offs in the reported trading and write a footnote about it, others normalize the trading and write a footnote about how trading is normalized, while others do not comment on it at all. The IPEV (2018) guidelines, which all funds in our sample follow, suggests that the GPs should *"*[…]*represent a reasonable estimate of maintainable earnings, which implies the need to adjust for exceptional or non-recurring items*[…]*"*. Thus, we assume that trading is adjusted unless explicitly communicated otherwise and use what is reported in the trading figures without adjusting and investigate potential outliers at a later stage.

6.1.7 Currency

Currency used for reporting in PE is not standardized, thus the GPs reporting currency differ, even between the fund itself and its portfolio companies. Typically, one currency is used for valuations across the fund's investments, while local portfolio company currency was used for trading. We collected figures in both EUR, NOK, DKK and SEK for our samples, noting the valuation and trading in their original currencies and converting everything into EUR for comparison purposes. This was done by using the exchange rate at the end of day of the respective quarters, thus the end of Q4 for the annual sample. The currency data was retrieved from Yahoo Finance.

The figures were converted from their local currencies into EUR as this was the currency most frequently used by both the funds and the portfolio companies. An issue with this approach is that companies reporting in a currency more heavily impacted by macroeconomic events might have their valuation and trading slightly misrepresented from period to period by the exchange rate. We find converting the local currencies into EUR as the preferable option as this minimizes the number of companies being subject to this effect.

6.2 Public Companies

6.2.1 Source

Refinitiv Eikon Datastream

The source for our sample of listed companies is Refinitiv Eikon Datastream, which is a financial data analysis platform with access to accounting data for listed companies in all Nordic countries back to our desired starting point of 2002. Datastream was chosen not only because it is one of the world's leading providers of financial markets data, but also since it is the commercial data provider of choice for Argentum. This satisfied our priority that the database and tool created is operational for Argentum and can be updated and developed further if desired.

Data Gathering

From the database of Datastream, we filter on public companies headquartered in one of the Nordic countries between 2002 and 2020, retrieving a list of 1565 unique entities. For each entity, we retrieve the following variables:

Table 2 – Variables for Public companies

Variable	Description
Company Name	
Country of Headquarters	
Exchange Name	
Bank Total Revenue	Quarterly and annually
Total Revenue EUR	Quarterly and annually
EBITDA	Quarterly and annually
Net Debt	Quarterly
Enterprise Value to EBITDA	Quarterly
GICS Sub-Industry Code	8-digit GICS Sub-Industry Code
Instrument Type	

All figures in Datastream are reported in EUR as of the final day in the quarter or year.
We started the data cleaning by using Instrument Type to remove non-companies from the list. This includes ETFs, American Depository Receipts and Open-Ended Funds. As we only want public companies listed on Nordic stock exchanges, companies on the OTC markets and foreign exchanges are also removed. Investigating the dataset, it becomes clear that the majority of companies listed on non-major stock exchanges, such as Oslo Axess and Merkur Market, frequently have missing values for our desired metrics. To ensure sufficient data quality, we remove companies from these exchanges, ending up with companies from Oslo Bors ASA, Nasdaq Stockholm, Nasdaq Copenhagen, Nasdaq Helsinki and Nasdaq Iceland. While we have not registered any Icelandic buyout deals, we include it in our public sample as it is a Nordic country.

6.3 Additional Dataset Processing

With one dataset for private companies and one for potential public peers, there were some additional preparations which were done for both datasets.

6.3.1 Industry Classification

As explained in Peer Group Identification, we need each company's industry classification in order to match private companies with public companies. We employ Global Industry Classification Standard (GICS), a four-tiered system developed by Morgan Stanley Capital International and Standard & Poor's, which has been shown to better explain variations in for instance valuation multiples, forecasts and growth rates, and financial ratios compared to other systems (Koller et al., 2010; MSCI, 2021). The four tiers are, from the broadest to the narrowest: Sector, Industry Group, Industry and Sub-Industry.

Most public companies are classified in this system such that each code can be retrieved through Datastream, but this is not the case for most private companies. In order to assign relevant GICS codes for our private companies, we compare the name and descriptions of all GICS codes to the industry classification assigned by the GP in the quarterly reports[18].

6.3.2 New metrics created

In order for the datasets to be operational for matching, we used the gathered data to create relevant metrics as mentioned in Peer Group Identification. These metrics were made for both portfolio companies and public peers.

As a basis for the metrics, LTM revenue and LTM EBITDA for the quarterly sample were constructed at a quarterly basis by adding up the four latest quarters of respective trading. Subsequently, revenue growth and EBITDA growth were created by calculating the growth in LTM for each quarter. For the annual sample, annual revenue and annual EBITDA already functioned as LTM figures, so we only had to create their annual growths. As a metric for profitability, EBITDA margin was created by dividing LTM EBITDA by LTM revenue. Finally, as a proxy for size and thus risk, created Log Sales by taking the logarithm of LTM revenue.

6.3.3 Missing values

For both the quarterly and the annual sample, there were missing values for some metrics at entry and exit, which had to be dealt with in order for the dataset to be operational.

Frequently, there was no trading data in the quarterly reports for the entry period itself, with financial metrics for the company first being available for the following quarter. This problem was also present for the annual sample, but was not as prevalent, possibly due to on average longer time from entry to the next reporting. While information about the entry period was often given ex-post in later reports, this was not always the case. For this reason, LTM revenue and EBITDA were first entered for a company in the first period where it was possible to construct LTM figures. Similarly, there were occasionally no trading data for the reporting period of the exit, if the GP only reported the investment as realized together with the corresponding transaction value. Thus, we used the previous period's LTM revenue and LTM

---

[18] https://www.msci.com/documents/1296102/11185224/GICS+Methodology+2020.pdf/9caadd09-790d-3d60-455b-2a1ed5d1e48c?t=1578405935658

EBITDA as proxies for the trading of the exit period. In the few cases where we have data for a quarter neither at entry nor exit is missing, we also use the preceding LTM metric instead.

6.4 Final Datasets

6.4.1 Private Companies

While the dataset for our quarterly and annual private companies are one, we split the set up in a quarterly and an annual sample to better be able to describe the datasets. This is also in line with how the quarterly and the annual samples will be matched independently in the matching process.

Table 3 – Descriptive Data on Private Companies - Quarterly

		Private Companies				
		Norway	**Sweden**	**Denmark**	**Finland**	**Total Sample**
Quarterly Sample						
Number of companies		38	40	12	6	96
Number of observations		953	911	229	117	2210
Number of industries		12	11	8	4	17
Average holding period (months)		51	45	42	42	47
	Max	59%	21%	19%	43%	59.0 %
Revenue Growth at entry	Median	5%	2%	1%	9%	3.1 %
	Min	-10%	-67%	-66%	-5%	-67.3 %
	Max	45%	129%	27%	62%	128.6 %
EBITDA Growth at entry	Median	6%	0%	0%	13%	3.0 %
	Min	-24%	-205%	-75%	-22%	-205.3 %
	Max	47%	32%	33%	23%	46.5 %
EBITDA Margin at entry	Median	17%	13%	20%	17%	14.4 %
	Min	4%	0%	6%	6%	0.4 %
	Max	9.0	9.0	8.9	8.7	9.0
Log Sales at entry	Median	7.8	7.9	8.2	7.4	7.8
	Min	6.9	7.0	7.5	7.3	6.9

In the quarterly sample we see that most of our companies and observations are from Norway and Sweden, which is in line with how Norway is Argentum's home market and Sweden is the biggest economy in the Nordics attracting the majority of buyout activity[19]. We see that

[19] https://argentum.no/wp-content/uploads/sites/73/2019/06/Argentum_The-state-of-Nordic-private-equity-2018_digital.pdf

the average holding period for buyout companies are similar across all countries. Both revenue growth and EBITDA growth seems volatile, while EBITDA margin at entry has no negative values. Through the logarithm of sales at entry, we get an impression of the company sizes, which is relatively similar for all countries.

Table 4 – Descriptive Data on Private Companies - Annually

		Private Companies				
		Norway	**Sweden**	**Denmark**	**Finland**	**Total Sample**
Annual Sample						
Number of companies		10	12	9	14	45
Number of observations		54	79	59	104	296
Number of industries		7	7	4	9	17
Average holding period (months)		41	55	53	63	47
	Max	89%	87%	81%	145%	145.1 %
Revenue Growth at entry	Median	57%	15%	15%	19%	19.2 %
	Min	13%	-6%	-5%	-13%	-12.5 %
	Max	95%	199%	81%	104%	198.7 %
EBITDA Growth at entry	Median	30%	17%	4%	15%	14.3 %
	Min	-166%	-650%	-207%	-66%	-650.0 %
	Max	68%	15%	45%	22%	68.0 %
EBITDA Margin at entry	Median	7%	10%	11%	9%	9.7 %
	Min	2%	2%	-6%	2%	-6.3 %
	Max	8.3	8.3	8.7	8.5	8.7
Log Sales at entry	Median	7.9	8.0	8.1	7.9	8.0
	Min	7.2	7.3	7.4	7.1	7.1

Compared to the quarterly sample, the sample of annual portfolio companies is smaller, both in number of companies and observations. However, we see that Finland has a larger number of companies in this sample. The revenue growth and EBITDA growth at entry is volatile here as well, with a higher median growth. The reason for this probably that the growth metrics in the quarterly sample are based on the growth in LTM from one quarter to the next, while it in the annual sample are based on growth from one year to the next, implying higher growth. The EBITDA margin is somewhat lower, and the size as represented through logarithm of sale is similar to the quarterly sample.

6.4.2 Public Companies

Table 5 – Descriptive Data on Public Companies

Public Companies	Norway	Sweden	Denmark	Finland	Iceland	Total
Quarterly Sample						
Number of companies	165	311	105	118	18	717
Number of observations	8588	16119.00	5679.00	6431	604	37421
Number of industries	21	22	19	24	10	24
Median LTM Revenue	131425765	157529368	210618608	248556500	200683387	210618608
Median LTM EBITDA	25295503	21041045	28836748	26404500	36457968	28836748

[1]Norway is Oslo Børs, Sweden is Nasdaq Stockholm, Denmark is Nasdaq Copenhagen, Finland is Nasdaq Helsinki and Iceland is Nasdaq Iceland

In the public sample, we see that Sweden as the largest economy has both the largest number of companies and by far the largest number of observations. Most industries, based on the four-digit GICS code, seems to be present across all countries, except Iceland which is an outlier in given the low number of companies and observations. The median LTM revenue is somewhat lower in Norway and Sweden compared to the rest of the countries, but the median LTM EBITDA seems similar. While the quarterly portfolio companies will potentially be matched with any of the observations, the annual portfolio companies will only be matched with the observations in Q4.

7. Empirical Analysis

As shown in the previous section, company value can be expressed as a function of three value drivers, profitability, growth and risk, respectively. We predict Enterprise Values by ensuring that the comparable companies exhibit similarity in these drivers, in addition to other factors that ensure comparability. The factors are emphasized in our models both through matching and additionally by the regression model we employ.

7.1 Coarsened Exact Matching

Coarsened exact matching is used to address the problem in our observational data that portfolio companies and public companies generally are not identical before private equity entry, and thus not initially suitable for comparable valuation (Iacus et al., 2012). In this section we explain how we address this by using the "*cem*" Stata package to execute coarsened exact matching (Blackwell, Iacus, King and Porro, 2010).

7.1.1 Matching Characteristics

To ensure comparability of public peers and portfolio companies, we control for key differences in value drivers. As explained in the thesis introduction, multiple valuation rests largely on the expectation that the comparable companies generate similar cash flows in the future (Berk and DeMarzo, 2017). Put in the statistical context of our application, this means that we need to control for variables that are important for PE-selection, which we find to be concurrent with factors known to identify comparables in relevant literature. Hahn (2009) concluded in his doctoral thesis on PE target selection that profitability, company size and performance trends are especially important criteria. The findings are consistent with the comparability factors outlined in e.g., Koller et al., (2010) and IPEV Valuation Guidelines (2018), which emphasize the same criteria in addition to leverage and minor qualitative factors.

Since it was not possible to supply the PE dataset with public sources (see Data) we rely exclusively on information from quarterly reports. This has consequences for the number of matching variables we can use and leaves us unable to fully satisfy the ignorability assumption of independence between treatment effect and matching covariate, because GPs generally do not provide pre-entry financials. This is especially challenging for leverage which is an important factor for value that should preferably be included as a covariate. Matching on capital structure data from the quarterly reports would however be a major violation of the pre-treatment assumption due to the immediate and substantial changes in debt brought on by a leveraged buyout. This is fortunately not the case for the remaining matching metrics that we use.

Financial performance is a key factor for assigning matches, preferably measured by trend/growth metrics such as *revenue growth*, *EBITDA growth*, while profitability can be measured by *EBITDA-margin* (Koller et al., 2010; Hahn, 2009). The only absolute measure we match on is the logarithm of revenue, *Log revenue*, which in accordance with Hahn (2009) is a fitting measure for company size that enables simple/standardized comparisons that are difficult to draw from raw revenue numbers. This is important since we are unable to capture the pre-entry capital structures/enterprise values which would arguably have been the best size measures. As mentioned in the Peer Group Identification section, *risk* is a key value driver that we seek to capture in matching, which is proxied by precisely company size, since portfolio company cost of capital is not available.

Revenue and EBITDA growth are measured on a LTM basis, with a downside being that we violate the pre-treatment ignorability assumption to some extent as a consequence, since growth measures require time-series data. However, we argue that PE ownership has not had a considerable effect on these metrics since they are gathered at max one year after entry. These metrics aims to capture the importance of *growth* in matching, while the EBITDA-margin reflects company *profitability*. Together with time and industry covariates, these variables are used as input to CEM.

We process the PE data before employing CEM into a matching database with only entry observations for the portfolio companies. The database is supplied with the entire public set with no restrictions on timing, since the portfolio companies have varying entries and must be able to search for Nordic public peers in their unique entry period. The table below shows the distribution of treated and untreated observations in the database, with the 141 treated representing the entry data for the PE companies, and the 37,558 untreated representing the

Nordic public companies with quarterly entries from 2002-2021. The matching universe for the annual PE companies is a quarter of the total public sample, 9,388 observations.

Table 6 – Matching Database

Matching Database	
Treated	Frequency
Portfolio Company	141
Public Company	37,558

7.1.2 Coarsening

We utilize the advantageous design of CEM and coarsen into bins set at our discretion, maintaining control of the trade-off between the number of matched portfolio companies and model dependency (King, 2018). We test these results against the predictions we get when using CEMs default binning algorithm to compare results.

In line with the sequence of comparable selection described in Koller et al., (2010), we start with industry as a prerequisite in matching. We run both four and six-digit GICS codes in CEM which corresponds to respectively 25 different industry groups and 71 industries[20]. To implement GICS exact matching, we use the cutpoints syntax in *cem* and indicate that we require a perfect match. The same logic applies for time, since we require an exact match for the corresponding entry dates (which is “quarter-year” for quarterlies and “year-end” for annuals). Since we require a perfect match on industry group/industry and time for all iterations, remaining variables are necessarily coarsened quite broadly to avoid excessive pruning. Financial performance measures *Revenue growth*, *EBITDA growth*, *EBITDA margin* and size measure *Log revenue* are thus coarsened into relatively broad bins:

[20] See MSCI GICS (2021) and sample GICS frequency table in the appendix.

Table 7 – Revenue Growth

	Private Equity Selection	
Revenue Growth	0	1
Negative	16,437	34
Positive	18,785	107
Very negative/error	122	
Very positive/error	1,253	

Table 8 – EBITDA Growth

	Private Equity Selection	
EBITDA Growth	0	1
Negative	17,068	48
Positive	15,595	85
Very negative/error	1,921	3
Very positive/error	2,665	

Table 9 – EBITDA Margin

	Private Equity Selection	
EBITDA Margin	0	1
Negative	4,994	2
Positive	31,588	139

Table 10 – Log Revenue

	Private Equity Selection	
Log Revenue	0	1
Large sized	7,883	
Medium sized	25,738	141
Small sized	761	

- *Revenue growth* is coarsened into principally negative and positive bins, however with cutpoints above negative and positive 200% to ensure that extreme values are pruned. The reason why we use such loose thresholds for cut-off is to allow companies with for instance one-off fluctuations in trading to still be matched.
- *EBITDA growth* is coarsened similar as revenue growth, with corresponding cutpoints.
- *EBITDA margin* is coarsened into negative and positive bins.
- *Log revenue* has its purpose in ensuring that very large(small) public companies are not matched with our medium sized sample of PE portfolio companies. Thus, we

coarsen it into bins that achieves this purpose, which we found with cut-offs at *log revenue* of below six and above nine, respectively.

If we apply more fine-grained coarsening, there simply are not enough matches due to the curse of dimensionality. As a consequence, the performance of our predictions relies largely on industry and the broad financial trends of the companies. See appendix for table of observation frequency for industries. Additionally, we run all iterations using CEMs default coarsening on the variables, with the exception of *Industry* and *Entry time*, because the default coarsening is unable to coarsen them. The general methodology is to run several iterations for all the model specifications and register and compare results. These are ultimately compared in the results section, where we also describe each limitation/rule that we apply beyond coarsening. After coarsening we get matched strata, which are then passed on to prediction in their respective uncoarsened format (Iacus et al., 2012). The table below illustrates what a matched pair may look like. The "*Positive*", "*Negative*" and "*Medium*" values are for illustration purposes only, the matches are assigned with their corresponding numeric values which are used post-matching.

Table 11 – Matching Illustration

Company	Entry Quarter	GICS	Industry Name	Revenue Growth	EBITDA-Margin	EBITDA Growth	Log Revenue
Private Equity	31/12/2010	101010	Energy Equipment Services	Positive	Positive	Negative	Medium
Public #1	31/12/2010	101010	Energy Equipment Services	Positive	Positive	Negative	Medium

7.2 Prediction Models

After we identify the matched peers in the respective strata from CEM, we process the data and move on to enterprise multiple predictions. The post-matching processing essentially ensures that pruned observations are removed from the set and that portfolio companies assigned to the same stratum are separated to ensure that they do not affect each other. The prediction models we employ are 1) Matched peers median multiples, and 2) Multiple regression. The performance is measured by mean squared error, using the transaction price as the true observation.

7.2.1 Median Prediction

The peer median model is relatively intuitive as it is explicitly contingent on the quality of the matching, extracting the median of the matched peers' enterprise multiple in each relevant time period. If we obtain accurate aggregated results by using the peer median, it is likely because the matching process using CEM is successful in identifying comparable firms that are able to indicate an accurate multiple at exit, when assigned at entry. We expect that the closest quarters up to and including twelve months prior to exit is most fitting (and most unbiased) to evaluate predictions against exit transaction prices, because as we observed in some quarterly reports, GPs will likely have had offers to base valuations of if it gets closer to realization.

7.2.2 Regression Model

The multiple regression model is pragmatically used as a test to see if a regression of the matched public company's enterprise multiple and trading can be used to predict portfolio companies' values near exit and at exit. We regress the chosen peers' enterprise multiple for each portfolio company using OLS in the respective holding period, with predictors created from peer data. It is not a strong model from an econometric perspective, primarily since the four[21] assumptions for unbiasedness are not satisfied. The assumption that the error term u has a conditional mean of zero (ZCM) given any values of the independent variables is certainly violated due to omitted variable bias, as there are many unobserved variables that are correlated with both the predictors and the enterprise multiple (Woolridge, 2013). We can thus conclude that there is an endogeneity problem since the ZCM assumption is violated. Nonetheless we run the regression as a point of reference, with concept and specification as follows.

$$Enterprise\ Multiple = B_0 + B_1 Revenue + B_2 EBITDA + B_3 Quarter\_Year + u$$

Where,

$u = error\ term, unobserved\ factors$

$B_0 = constant, intercept$

$B_j = slope\ paramter\ associated\ with\ variables$

[21] Linearity in parameters, random sampling, sample variation in predictors and zero conditional mean. See e.g., Woolridge (2013).

We run the model separately for each portfolio company and divide the results into their respective frequency of reporting, quarterly and annually. For the quarterlies, we run regressions and obtain predictions in the three nearest quarters to exit and at exit. For the annual sample we regress the two years prior to exit as well as at exit. To ensure that each portfolio company is predicted only by the financials from its relevant matched peers, we run one regression for each portfolio company (totaling 141 regressions) in all relevant time frames using the Stata command *statsby*, which allows us to streamline the process. The variables we use are *EBITDA*, *Revenue*, *Quarter_Year* (time dummy), and *Industry* (which is accounted for through CEM). We multiply the portfolio companies' trading by the regression coefficients and add the relevant dummy coefficient from each regression to obtain the predicted multiple for any given quarter/year.

7.2.3 Outliers and Rules

Relative valuation for any asset class will usually include a method to prevent outliers from skewing results, sometimes handled at database level by the reporting services that provide the financials. Such services may exclude outliers when computing averages or constrain the accepted intervals for multiples in general (Damodaran, 2011). One of the most regular method of dealing with outliers first-hand is using medians, which are more meaningful than averages that are subject to the distortions of extreme values. We want to identify the most appropriate way of addressing outliers in our public dataset, so that we are able to implement a "set of rules" to the tool. After testing which coarsening level and GICS specification predicts the most accurate results, we address outliers by implementing exclusion rules based on results, which we account for thoroughly in the results section.

We test the model specifications in an iterative process with much of the same logic as with coarsening: first we evaluate prediction results without exclusions or constraints, and then step-by-step implement rules that is thought to increase prediction accuracy, although typically at the expense of exclusion. This constraining process is implemented equally for both models and subsamples, which we generally expect to be more significant for the regression since it does not inherently deal with outliers.

7.2.4 Mean Squared Error

In order to draw inferences from the prediction results we need a measure that quantifies how well the predictions match the true values we observe. We evaluate prediction results in line with James et al., (2013) using the mean square error. As seen in the formula below, the MSE is calculated by squaring the difference between true and predicted response. Since MSE is squared, it is by nature highly sensitive to extreme values/outliers, which dominate the calculation if present. Thus, we expect MSE for the regression to be far off initially and substantially improved when we address extreme values. If CEM does a good job of matching, we expect that the median model will produce consistently low MSEs leading up to exit. We calculate and evaluate the MSE for all three predictions in the periods nearing exit, i.e., for the four closest quarters for the quarterly subsample, and for the two nearest years for the annual sample, in addition to at exit for both samples.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{f}(x_i)\right)^2$$

Where,

$n = number\ of\ predictions\ in\ the\ sample$

$y_i = true\ enterprise\ multiple\ for\ observation\ i$

$\hat{f}(x_i) = predicted\ enterprise\ multiple\ for\ observation\ i$

8. Results and Discussion

In this section, we will first compare the predicted exit multiple for each portfolio company given by our two models with the actual exit multiple. This provides us with a mean squared error for each prediction, which is then aggregated by calculating the average MSE for all predictions. The MSE is hard to interpret in isolation, for our purpose it must be evaluated relative to other model specifications. After evaluating which of our prediction models perform the best relatively, we move on with the best specifications and compare their predictions to the GP estimates in the quarters and years leading up to exit. We thus evaluate how applicable the models would be as a tool in practice.

As the results originate from an iterative process where model specifications are tested, adjusted for what is observed, and then re-tested, we will present the results and outline our process simultaneously, providing insight and discussion of the choices made. This will provide context to our discussion of the best performing specifications compared to GP estimates in the time leading up to exit. Finally, we discuss all the results seen together and consider limitations.

8.1 Exit Prediction and Internal Model Evaluation

Exit is the first point in time where the actual transaction value is observable, referred to as the true response (James et al., 2013). In the following table, the predicted exit responses for portfolio companies are subtracted from the true response and squared, each number thus represents an MSE corresponding to a specified variant of GICS coarsening and set of exclusion rules (that are increasingly strict). This is with the exception of the cursive numbers which indicate the number of portfolio companies that are predicted, with initial sample size *n* of 96 for the quarterly sample and 41 for the annual sample. The difference in the number of predicted portfolio companies and the respective sample size thus represents the amount of portfolio companies that were pruned.

Table 12 – Mean Squared Error Predictions at Exit

Mean Squared Error of Predictions				
	Quarterly (n=96)		Annual (n=41)	
	Custom Bins	**Default Bins**	**Custom Bins**	**Default Bins**
Narrow GICS				
Median Peer Multiple Model	704	127	170	208
Regression Model	7 605 026	87 168	594 904	2 576
# Portfolio Companies	*59*	*75*	*21*	*31*
Wider GICS				
Median Peer Multiple Model	48	44	82	138
Regression Model	864 495	3 140	442 575	43 804
# Portfolio Companies	*86*	*88*	*32*	*37*
Excluding Extreme Peer Values[1]				
Median Peer Multiple Model	28	37	26	115
Regression Model	1 256	1 340	11 994	386
# Portfolio Companies	*86*	*88*	*32*	*37*
Excluding Portcos With ≤ 2 peers				
Median Peer Multiple Model	28	38	20	129
Regression Model	1 331	159	451	342
# Portfolio Companies	*71*	*66*	*25*	*34*
Excluding Portcos With ≤ 4 peers				
Median Peer Multiple Model	23	45	27	35
Regression Model	335	123	29	55
# Portfolio Companies	*50*	*44*	*19*	*20*
Excluding Banking and Insurance				
Median Peer Multiple Model	24	45	27	35
Regression Model	62	123	29	55
# Portfolio Companies	*48*	*44*	*19*	*20*

[1]Excluding Enterprise Multiples > 50 and < 0

We start by evaluating whether to use a narrow or wider variant of the GICS classification. It is concluded that the wider GICS seems better suited for our purpose, as the narrower GICS excludes too many portfolio companies with no gain in predictive performance of neither the median nor the regression model. This is true for both custom and default bins. We thus move forward with the wider specification and implement exclusion measures based on common features of the specification results.

First, we observed that extreme values in the enterprise multiples from peers substantially distorted the predictions. For example, EBITDA is an item calculated after nonoperating expenses, and thus may be subjected to substantial one-offs that distort the public companies' enterprise multiples, which we are unable to normalize. Both our prediction models were distorted by extreme values, but especially the regression model since its predictions is not inherently protected from outliers, compared to the median. This is the most prevalent for large values of EV/EBITDA, as Datastream already removes negative EV/EBITDA values if caused by negative EBITDA (see Data). Since we cannot deal with these cases explicitly, we find the best alternative is to remove them. Thus, all positive enterprise multiples above 50x are removed, together with the remaining negative enterprise multiples originating from negative EVs. After removing these values from the public dataset, the regression MSE for the quarterly custom bin specification went from 864,495 to 1,256, which is a considerable improvement in relative terms, but still notably worse than the median peer multiple. We observe similar effects for all other specifications as well.

Second, we noted that the exit predictions of portfolio companies with few peers performed far worse than the average, which was especially true for the regressions where portfolio companies became too reliant on the company-specific development of a few companies. Thus, we implemented rules indicating that portfolio companies with equal to or less than two peers should be excluded, afterwards increasing this threshold to four peers. These rules improved regression MSE for all regression specifications, especially when demanding more than four peers.

Finally, due to the unique characteristics of capital structure in the banking and insurance industry, we observed that the few portfolio companies in this industry were consistently predicted poorly on the basis of the industry groups' unique enterprise multiples. This was however only applicable for the quarterly custom bins specification, as neither the other specifications nor the annual sample contained any portfolio companies in this industry.

Ultimately, after having implemented a four-digit industry code as well as three rules of exclusion, we had addressed all apparent unreasonable distortions. The improvement was substantial for the regression model, since the rules were able to address the outliers that severely influenced its initial predictions. The median model was not enhanced much beyond the improvement from exclusion of extreme values of peer enterprise multiples and remained consistent afterwards. It is thus evident that the median model is applicable without having to implement many explicit measures.

The results seem to suggest that defining custom bins based on variable knowledge compared to using default bins is beneficial to identifying well suited matches[22], after having accounted for distortions from outliers. The default algorithm often seems to perform better in the first iterations, but after excluding what we view as distortions in the data, it seems like the custom bins are more able to capture the underlying relationship between the matching variables and EV/EBITDA. This is in line with our findings from the literature that similarity in growth, profitability and risk are tied to predicting similar EV/EBITDA multiples, which we capture by basing our custom bins on broad trends in these factors.

The custom bins chosen for the matching variables are broad, and one could argue that more relevant matches, and thus better predictions, could have been assigned if the bins were finer and a larger public sample such as all listed European companies was used. An issue of setting finer cut-off points is that one might cut matches that are close to perfect, e.g., when trying to categorize between "medium" and "high" revenue growth. If the cut-off is at 50%, then a portfolio company with 51% revenue growth will never be in the same stratum, and thus never match with a public peer which is identical in all variables except having a 49% revenue growth (see curse of dimensionality under Matching).

[22] This is in line with the preferred use of CEM, to coarsen covariates into substantive groups that preserve information. See Blackwell et al., (2010).

8.2 Comparing Predictions to GP Estimates

Having compared and identified the most accurate model specifications, we compare the model predictions to the respective GP estimates in the periods leading up to exit

Quarterly Sample

Table 13 – Mean Squared Error Predictions Prior to Exit - Quarterly

Mean Squared Error Predictions - Quarterly				
	3 Months Prior	**6 Months Prior**	**9 Months Prior**	**12 Months Prior**
Custom Bins				
Median Peer Multiple	17,70	20,02	21,22	23,93
Regression	67,48	56,79	51,70	50,55
General Partner Estimate	5,06	9,18	29,41	46,23
Default Bins				
Median Peer Multiple	69,11	65,01	123,25	39,81
Regression	117,02	110,47	97,30	101,62
General Partner Estimate	11,05	8,45	68,68	68,06

For the quarterly sample with matches from custom bins, we observe that the MSE of the median model is quite consistent in all quarters in the twelve months before exit, but not as accurate as the GPs' estimation when the exit is six months or less away. The large spike in GP accuracy between twelve and six months is likely because the GP often receives an indicative offer or has entered negotiations. The median model exhibits consistency in its predictions, and also seems to predict at least at the level of GP accuracy when moving nine months or further away from exit.

The GP estimates for the portfolio companies that are present when applying default bins exhibits similar dynamic to the GP estimates for custom bins between nine and six months prior to exit, where accuracy is notably increased. The default specification for the median peer multiple is performing worse, especially from six months and earlier. Investigating this data, the relatively high MSE of 123,25 is due to a large prediction miss from a singular portfolio company. The company in question was not assigned any matches by CEM when applying custom bins, suggesting that leveraging knowledge of the variables we coarsen helps avoid poor matches, however this is difficult to say for certain. This statement is also supported by the outperformance of the custom bins versus the default bins for the median model in all periods for the quarterly subsample. The regression model is relatively inaccurate but is however somewhat consistent.

8.2.1 Annual Sample

Table 14 – Mean Squared Error Predictions Prior to Exit - Annually

Mean Squared Error Predictions - Annually

	1 Year Prior	2 Year Prior
Custom Bins		
Median Peer Multiple	25,91	27,89
Regression	32,50	36,82
General Partner Estimate	4,02	12,59
Default Bins		
Median Peer Multiple	25,50	32,05
Regression	47,63	45,31
General Partner Estimate	11,98	40,08

For the annual sample, the GP prediction becomes more accurate closer to exit, supporting our findings for the quarterly sample which indicated that accuracy close to exit stems from unique deal information. It should be noted that the smaller annual sample size of 41 companies compared to 96 in the quarterly sample makes it relatively more prone to random noise and is thus more difficult to draw inferences from, especially when pruned down to about half of its initial sample size.

In summary, a median model with relatively few exclusion rules seems to perform quite consistent, although not more accurate than the GPs when approaching realization. This is in line with our expectations given the GPs information advantage close to exit. However, moving further than six months away from exit, our results suggests that our best performing specification using substantive knowledge on variable coarsening may provide a consistent and valid second opinion on the Enterprise Value of portfolio companies.

8.2.2 Limitations

While it seems like our median model is able to predict exit values with similar accuracy as the GP further than six months away from exit, the predictions at company level are still frequently inaccurate, and the model aggregated MSE is highly sensitive to singular but large prediction misses. One explanation could be that the matching algorithm assigns peers that are similar to the portfolio company at PE entry, but the companies might not necessarily be comparable at realization after typically 3-8 years (Kaplan and Sensoy, 2014). Portfolio companies are often acquired by a private equity firm with the intent of implementing measures such that the business can be changed to attract higher valuations, or the company can be used as a platform for acquisitions facilitating inorganic company growth, both entailing substantial change. Ideally, the matches would thus be updated more frequently, however this does not seem feasible in our current framework since our methods assumes that buyout treatment is the only difference between the public and private companies. This assumption would be severely violated with frequent matching updates, since buyout would have already occurred. At the extreme end, there is thus reason to believe that our models are not able to properly value abnormal over- or underperformers deviating from peer median multiples. This is unfortunate as many PE funds will have a few portfolio companies that are transformed into markets leaders.

Sample size

A key limitation for the inference from our findings is the relatively small and non-random PE sample sizes. The results of our study are for instance strictly speaking only applicable to Nordic portfolio companies in mid-market funds that are managed by top quartile private equity firms. The limited sample size might also influence how the rules for exclusion are implemented. In our results, we saw an example of how portfolio companies within the banking sector only showed up in the wider GICS custom bin specification, where it became apparent that EV/EBITDA is not a good valuation tool for this industry. With an even smaller sample size, this rule might not have been implemented, underscoring that the model can only be relied on to work for similar companies it has predicted in the past.

Covariates

Another issue during matching is that the portfolio companies are paired with public peers based on metrics from a single quarter or year, which may not be representative for its longer-term trend. While basing financial metrics on LTM figures mitigates this issue somewhat, the metrics might still be volatile. As we have already mentioned, there are confounders that are relevant for PE selection and enterprise multiple prediction that we were unable to match on. It is plausible that these would have increased the quality of matches, especially if we had used a larger public sample so that the "increased" curse of dimensionality introduced by new covariates would be somewhat offset by an increase in sample size. This further suggests that

that Argentum should request more accounting information in the quarterly reports they receive in order to get information about confounders, especially Net Debt at entry, which might provide better matches and thus better predictions. The results might also give LPs incentives to push the GP harder for precise exit dates, as this will give them a better view of what time to trust GP valuations over the peer model. Such information might be beneficial for Argentum's secondary team in knowing which prediction of the exit value is the most probable when valuating fund stakes in interim periods.

Mean Squared Error

While MSE is the desired measure for evaluating predictions, an issue with solely regarding MSE at the aggregated level is that large deviations are weighed heavily. Thus, we are implicitly unable to remedy situations where we are correct in most predictions but miss substantially on a singular one. This could e.g., be addressed by implementing measures that made sure that certain values are investigated and addressed subjectively. As we identified when investigated our results, one large prediction miss was able to heavily influence one of the datapoints in our results.

Marketability Discount

In line with our expectation and literature on marketability, we consistently overestimate valuations relative to the GPs when applying peer multiples without discounts to private companies in periods before exit (Harjoto and Paglia, 2010). Predictions would thus be more accurate with a discount up to a certain percentage, but it seems somewhat arbitrary to apply it without investigating the derivation of each portfolio companies predicted multiple thoroughly, which is beyond the scope of this thesis.

8.2.3 Conclusion

In this thesis we set out to investigate if a multiple based statistical approach would be able to deliver unbiased and accurate valuations. Compared to the GP, our models are generally less accurate, with the expectation of the median model that performs at least at the level of the GP when further than six months prior to exit. This model is however frequently inaccurate at company level. Few, but large prediction errors are penalized heavily across all predictions due to the nature of MSE, which squares the difference between predicted and true responses.

This is particularly evident when we address outliers for the regression model, where MSE moves drastically down as a few large misses are pruned. We also note that the GP valuations are consistent across all our samples, exhibiting no signs of bias other than at the two quarters closest to exit, where we suspect that they learn transaction details from potential buyers.
The limitations to our results make the peer model somewhat difficult to implement successfully at the present for an LP like Argentum, as there are many boundaries to the applicableness of the tool that. Even so, the sample that the tool is successfully tested on, Nordic buyout companies, is one of the core focuses of Argentum. Thus, as long as the predicted company does not change too much from its original peers and is not heading for bankruptcy, the model might already be usable for a relatively large part of Argentum's portfolio.

Otherwise, the peer model can be regarded as in experimental stages, where Argentum can apply their expert knowledge about the PE industry to tweak the bins to better catch nuances in the buyout segment, implement other rules, or differentiate in how predictions from different GPs are regarded given their repeat business with them. A substantial advantage with the model design through CEM is that it facilitates and benefits off the use of expert knowledge.

9. References:

Argentum. (2020). The state of Nordic Private Equity 2020.

Berk, J., & DeMarzo, P. (2017). Corporate Finance (4th ed.). Harlow: Pearson Education.

Bernström, S. (2014). Valuation: The Market Approach. Chichester: John Wiley & Sons.

Blackwell, M., Iacus, S., King, G., & Porro, G. (2010). cem: Coarsened Exact Matching in Stata.

Brown, G., Gredil, O., & Kaplan, S. (2019). Do private equity funds manipulate reported returns? *Journal of Financial Economics, 132*(2) 267-297

Caliendo, M., & Kopeinig, S. (2008). Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys, 22*(1) 31-72.

Chung, J., Stern, L., Sensoy, B., & Weisbach, M. (2010). Pay for Performance from Future Fund Flows: The Case of Private Equity. *The Review of Financial Studies, 25*(11) 3259-330

Clifford Chance. (2020). Recent SEC Enforcement Actions Highlight the Importance of Sound Valuation and Disclosure Practices by Investment Managers.

Damodaran, A. (2005). The Value of Control: Implications for Control Premia, Minority Discounts and Voting Share Differentials. *SSRN Electronic Journal.*

Damodaran, A. (2006) Valuation Approaches and Metrics: A Survey of the Theory and Evidence. *Foundations and Trends in Finance, 1*(8), 693-784.

Damodaran, A. (2011). The Little Book of Valuation. New Jersey: John Wiley & Sons.

Damodaran, A. (2012). Investment Valuation: Tools and Techniques for Determining the Value of Any Asset (3rd ed.). New Jersey: John Wiley & Sons

Døskeland, T., & Strömberg, P. (2018). *Evaluating Investments in Unlisted Equity for the Norwegian Government Pension Fund Global (GPFG).*

Financier Worldwide. (2014). Q&A: Valuations for the private equity industry.

Grant Thornton. (2021). Report on Transactions and Comparables.

Hahn, M. (2009). *Essays on Private Equity Value Creation* [Ph.D. thesis]. Ludwig-Maximilians-Unviversität München

Harjoto, A., & Paglia, J. (2010). The Discount for Lack of Marketability in Privately Owned Companies: A Multiples Apporach. *Journal of Business Valuation and Economic Loss Analysis, 5*(1) 5-5.

Iacus, S., King, G., & Porro, G. (2012). Causal Inference without Balance Checking: Coarsened Exact Matching. *Political Analysis, 20*(1) 1-24

IPEV. (2018). *International Private Equity and Venture Capital Valuation Guidelines*.

Ivashina, V., & Lerner, J. (2017). Pay Now or Pay Later?: The Economics within the Private Equity Partnership. *Journal of Financial Economics, Forthcoming.*

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning with Applications in R (2nd ed.). Springer.

Jenkinson, T., Sousa, M., & Stucke, R. (2013). How Fair are the Valuations of Private Equity Funds? *SSRN Electronic Journal.*

Kaplan, S., & Sensoy, B. (2015). Private Equity Performance: A Survey. *Annual Review of Financial Economics, 7*(1) 597-614.

Kaplan, S., & Strömberg, P. (2009). Leveraged Buyouts and Private Equity. *Journal of Economic Perspectives*, *23*(1) 121-146.

King, G. (2018 May 30). *Simplifying Matching Methods for Causal Inference*. [Conference presentation]. International Conference about Innovations in Political Methodology and China Study, Taiwan.

Koller, T., Goedhart, M., & Wessels, D. (2010). Valuation: Measuring and Managing the Value of Companies (5th ed.). New Jersey: John Wiley & Sons.

Lattman, P. (2012, February 12) Private Equity Industry Attracts S.E.C. Scrutiny. *Dealbook*

Metrick, A., & Yasuda, A. (2011). Venture Capital & the Finance of Innovation (2nd ed.). New Jersey: John Wiley & Sons

MSCI. (2021). The Global Industry Classification Standard (GICS).

Pinto, J., Robinson, T., & Stowe, J. (2015). Equity Valuation: A Survey of Professional Practice. CFA Institute.

Rakner, C., & Rasmussen, A. (2013). *The Bullshit Effect in Private Equity Fundraising*. [Master's thesis]. Norwegian School of Economics.

Refinitv Datastream. (2020). *The world's most comprehensive historical database.*

Rosenbaum, P. (1984). The Consequences of Adjustment for a Concomitant Variable That Has Been Affected by the Treatment. Journal of the Royal Statistical Society, *147*(5) 656-666

Rosenbaum, P., & Rubin, D. (1983) The central role of the propensity score in observational studies for causal effects. Biometrika, 70(1) 41-55

Wooldrige, J. (2013). Introductory Econometrics a Modern Approach (5th ed.). Ohio: Cengage Learning.

10. Appendix:

Frequency of GICS Industries

GICS Industry Group Name	Public Companies	Private Companies
Automobiles & Components	402	.
Banks	3266	.
Capital Goods	6265	26
Commercial & Professional Services	1937	15
Consumer Durables & Apparel	1322	13
Consumer Services	603	14
Diversified Financials	1819	3
Energy	1655	5
Food & Staples Retailing	249	2
Food, Beverage & Tobacco	1408	3
Health Care Equipment & Services	1862	13
Household & Personal Products	160	2
Insurance	559	1
Materials	2122	10
Media & Entertainment	1061	2
Pharmaceuticals, Biotechnology & Life Sc.	2304	1
Real Estate	2461	.
Retailing	1097	6
Semiconductors & Semiconductor Equip.	210	.
Software & Services	2237	18
Technology Hardware & Equipment	2368	4
Telecommunication Services	353	2
Transportation	1591	1
Utilities	247	.