The Aggregate Cost of Equity Underdiversification

Florentsen, Bjarne; Nielsson, Ulf; Raahauge, Peter; Rangvid, Jesper

The Aggregate Cost of Equity Underdiversification

Bjarne Florentsen, Ulf Nielsson, Peter Raahauge, and Jesper Rangvid

Journal article (Accepted manuscript*)

Uploaded to CBS Research Portal: May 2020

The aggregate cost of equity underdiversification*

December 2018

Bjarne Florentsen
Ulf Nielsson
Peter Raahauge
Jesper Rangvid

The aggregate cost of equity underdiversification

Abstract

We analyse equity diversification of all retail investors in a country (Denmark). We find that underdiversification is pervasive. We calculate the nationwide aggregate loss due to underdiversification and express it in absolute and expected-return terms. The aggregate loss is large. We find that investors with low education, low income, and low wealth are more likely to underdiversify. In spite of better diversification, the larger fraction of the aggregate loss nevertheless adheres to the top of the income/wealth distribution. Increasing financial literacy among the majority of stockholders holding few stocks can therefore only modestly cut the aggregate loss of underdiversification.

1. Introduction

Retail investors spread their equity investments over only a few different stocks. This equity underdiversification was first described in Blume & Friend (1975) and has since then been reconfirmed in numerous studies, see, e.g., Kelly (1995), Barber & Odean (2000), Polkovnichenko (2005), Ivkovic, Sialm, & Weisbenner (2008), and Goetzmann & Kumar (2008).[1] Equity underdiversification is surprising in light of the fact that idiosyncratic risks can be reduced – without sacrificing expected returns – by holding well-diversified portfolios. A number of papers have shown how standard deviations of equity portfolios can be reduced by diversification and how many randomly selected stocks are needed to eliminate idiosyncratic risks, starting with the by now famous Evans & Archer (1968) article, and followed by Bloomfield, Leftwich, & Long (1977), Elton & Gruber (1977), Statman (1987, 2004), Campbell Lettau, Malkiel & Xu (2001), Domian, Louton, & Racine (2007), Benjelloun (2010), and others. Suggested explanations for the underdiversification puzzle include preferences for lottery stocks (Barberis & Huang, 2008), margin and borrowing restrictions (Roche, Tompaidis, & Yang, 2013), solvency constraints (Hong, 2014), financial literacy (Gaudecker, 2015), and other channels. But how large are the aggregate costs retail investors suffer from underdiversification? If these are minuscule, we may not be that concerned. On the other hand, if they are sizeable, it is even more important that we continue our efforts to try to understand their underlying reasons.

The main contribution of this paper is that it calculates the aggregate cost of underdiversification of the total population of retail investors in a country, Denmark. We calculate the costs of underdiversification as the return foregone due to underdiversification, i.e. the return an investor foregoes by holding portfolios with idiosyncratic risk (that is not compensated by expected return) instead of portfolios with systematic risk (that is compensated by expected return). As an example, what is the expected return the investor foregoes by holding a portfolio with an idiosyncratic level of risk of, say, 30%, instead of a

---

[1] Retirement accounts are also underdiversified, see e.g. Benartzi & Thaler (2001) and Agnew, Balduzzi, & Sunden (2003).

portfolio with a systematic level of risk of 30%? We find sizeable costs of underdiversification. In our baseline calculation, we find that retail investors could potentially increase their expected return by up to three percentage points per year, without increasing risk, by shifting from concentrated portfolios to diversified portfolios. We are able to calculate the aggregate cost of underdiversification because we have access to data that contains the complete equity portfolios (outside retirement savings) of all retail investors in Denmark. We thereby differ from studies that examine subgroups of investors, such as on-line brokers, respondents to surveys, etc., as in Blume & Friend (1975), Kelly (1995), Barber & Odean (2000), Polkovnichenko (2005), Ivkovic et al. (2008), Bodnaruk (2009), and Goetzmann & Kumar (2008).

We first describe the portfolios of our retail investors, concentrating on the number of different stocks entering the portfolios. We find that the typical investor holds a very concentrated portfolio. Most investors, 66% of investors, hold one stock only. We also find that as little as 1.8% of investors hold more than ten stocks. This underdiversification resembles that found in other studies, such as those mentioned above.

We analyze a sample of Danish stocks for which we have monthly data for the last twenty years.[2] There are 91 such stocks. Based on this sample of stocks, we calculate the level of risk (standard deviation of returns) of portfolios including different numbers of stocks, i.e. the relation between the number of stocks in a portfolio and its standard deviation, in the tradition of Evans & Archer (1968) and the subsequent literature.[3] With one randomly chosen stock, we find that the expected standard deviation is 38%. When increasing the number of stocks in the portfolio, standard deviation of the portfolio naturally drops. We provide confidence bounds around the standard deviation, based on the empirical distribution. We find that with around 50 stocks in the portfolio, the standard deviation of the portfolio reaches its convergence point at around 15.5%. This is also the level of risk of

---

[2] Florentsen, Nielsson, Raahauge, & Rangvid (2017), who analyze the same investors as those we analyze here, show that Danish investors primarily invest in Danish stocks. They find home bias to be 86%, justifying our focus on Danish stocks.

[3] A by-product of our analysis, thus, is that we present results for the relation between portfolio standard deviation and number of stocks in a portfolio for a non-US market (Denmark, in our case).

the market portfolio, i.e. the portfolio of all 91 stocks. As much as (38% - 15.5%)/38% ≈ 60% of risk can thus be eliminated by diversification. Our finding that it takes around 50 randomly chosen stocks to eliminate idiosyncratic risk is in line with results reported in Campbell et al. (2001) for US stocks.

We propose a simple way to calculate the potential impact of diversification on expected returns. The model we use is the CAPM and its basic insight that investors get compensation for bearing systematic risk, but not for bearing idiosyncratic and diversifiable risk. Instead of holding, e.g., a one-stock portfolio with a 38% standard deviation, the investor can hold a well-diversified portfolio (with risk equal to 15.5%) and lever up this portfolio to 38% risk, i.e. move up the capital-market line and earn a risk premium. The historical risk premium on the Danish stock market is around five percent (Dimson, Marsch, and Staunton, 2011). We use this as our benchmark estimate of the expected risk premium (and provide robustness checks using alternative estimates for the expected risk premium). We find that an investor who moves up the capital-market line from 15.5% to 38% risk, should expect an increase in expected returns of slightly more than seven percentage points. This means that an investor foregoes at least seven percentage points expected return by holding a one-stock portfolio, compared to holding a well-diversified portfolio. We perform such calculations for portfolios with two stocks, three stocks, etc. up until 91 stocks.

From our data on the holdings of all Danish investors, we know the value of each investor's portfolio. We also know the loss, in terms of foregone expected returns, of holding a one-stock portfolio, a two-stock portfolio, etc., as just described. We multiply the value of stocks for investors with one-stock portfolios with their return foregone due to underdiversification. This gives the value of the loss due to underdiversification per investor who holds one stock, on average. We sum over all investors with one-stock portfolios. This gives the aggregate loss for investors with one-stock portfolios. We perform the same calculations for investors with two-stock portfolios, three stock portfolios, etc. Lastly, we sum across all portfolios and investors. This gives us the aggregate cost of underdiversification. We compare this to the total value of all investors' stock holdings. We

find that the aggregate loss from underdiversification of all investors is 3.1%. The interpretation of this number is that if all retail investors shifted from underdiversified portfolios to the market portfolio, the aggregate return they obtain could potentially be increased by 3.1 percentage points per year. Given that we have assumed an equity risk premium of five percent in these calculations, an increase in expected return of 3.1 percentage points just by diversifying is significant. We also calculate the fraction of losses born by different types of investors. As mentioned, 66% of investors hold one stock. We find that they bear 36% of the loss due to insufficient diversification. We also find that they own 15% of capital allocated to the stock market. Small investors with few stocks thus provide a relatively small fraction of the capital allocated to the stock market by retail investors, but bear a large fraction of the costs.

We use the fact that we have access to detailed data on the background characteristics of each investor to characterize those types of investors who are more likely to underdiversify. We distinguish investors according to their level of income, total wealth (sum of stocks, bonds, real estate, etc.), education, age, gender, etc. We calculate both the aggregate cost of underdiversification for different types of investors, as well as the likelihood that certain types of investors tend to be underdiversified using probit regressions. We find that investors with low income, low levels of wealth, and low levels of education are more likely to hold less diversified stock portfolios. The cost of underdiversification borne by these investors, relative to what these investors have invested in the stock market, is thus higher than for investors with high income, wealth, and education. This squares with results in e.g. Calvet, Campbell & Sodini (2009) that investors with low levels of income, wealth, and education are less financially sophisticated. We add to this literature by showing the fraction of the aggregate cost of underdiversification different types of investors bear. We find – not surprisingly – that investors with low levels of income, wealth, and education allocate a lower aggregate amount of wealth to the stock market. This means that even when investors with low wealth, income, and education bear a larger fraction of the costs of underdiversification relative to what these investors have allocated to the stock market, they bear a lower fraction of the aggregate cost of underdiversification. The implication of this result is that if financial

sophistication could be improved for investors with low levels of financial sophistication, the cost of underdiversification they individually bear could be reduced. The aggregate cost of underdiversification would be relatively less affected, though, because these investors allocate less capital to the stock market.

We discuss various aspects of our results. The first is that the typical diversified investor would probably not hold 91 stocks in her portfolio. The relevant alternative to holding few stocks is more likely an index mutual fund. The annual fee of an index fund tracking the Danish stock market index is around 50 basis points. The expected return gain from diversification, after costs, thus probably amounts to 2.6 percentage points. Still sizable. A second aspect we discuss is that some investors – 26% of stock market investors – hold mutual funds in addition to directly held stock. If the larger fraction of stock holdings of an investor is allocated to mutual funds, the investor might be well-diversified, even if holding only one or two stocks directly. Most of our investors do not hold mutual funds. In fact, 74% of stock market investors get stock market exposure from directly held stocks only, i.e. do not hold mutual funds. If we exclude investors holding mutual funds, we find that the loss from underdiversification is 3.7 percentage points per year. The conclusion from this is that most investors do not diversify using mutual funds, and suffer large losses due to underdiversification. Third, we discuss our assumption of treating all investor equal, conditioned on the number of stocks they hold. The point here is that some underdiversified investors might be doing well, i.e. pick few well-performing stocks, whereas others are doing very bad. We study the average investor. We discuss this approach in light of evidence in the literature regarding heterogeneity across investors. Fourth, our results illustrate the potential theoretical gain investors could achieve from replacing their current portfolios with fully-diversified portfolios. Whether such transitions would cause shifts in relative demands for stocks and thus have pricing effects that might affect the market depends on dynamics of shifts, how other investors react etc. It is outside the scope of this paper to deal with such transitional effects, but we discuss them. Finally, we provide a perspective on the hypothesis that investors (probably erroneously) believe that they can find the "best" stock. We do so by comparing the performance of a portfolio with n randomly chosen stocks to the

performance of the market portfolio. We find that the likelihood of choosing a stock that delivers an average return below the average return on the market portfolio is high, around 75%. When increasing the number of stocks in the portfolio, the likelihood of outperforming increases and the likelihood of underperformance decreases. Basically, the likelihood of picking a stock that outperforms the market portfolio is low ex ante. We discuss these results in relation to the stock-as-lotteries hypothesis of Barberis & Huang (2008).

The papers closest to ours are Goetzmann & Kumar (2008) and Calvet, Campbell, and Sodini (2007). Goetzmann & Kumar (2008) also find significant costs of underdiversification. They analyze a sample of investors at one U.S. discount brokerage house. The main difference between the study of Goetzmann & Kumar (2008) and ours is that we calculate the aggregate loss of a population. We can do so because we know the number of stocks each individual holds and how much each investor has allocated to stocks. Calvet, Campbell, and Sodini (2007), like us, use very detailed data to study all investors in a country, Sweden. They find that costs are low when including mutual fund holdings. We find that most Danish retail investors do not hold mutual funds. When we study the large subgroup of investors who hold no mutual funds, we still find significant losses due to underdiversification.

The remainder of the paper is structured as follows. In the next section, we describe the data we use, focusing in particular on the number of stocks investors hold in their portfolios. Section 3 describes the sample of Danish stocks we consider. Section 4 investigates the relation between the number of stocks in investors' portfolios and risks and returns. In Section 5, we calculate the nation-wide total cost of underdiversification of all investors in a country. We discuss robustness tests in Section 6. Section 7 analyses who bears the costs of underdiversification. Section 8 studies the likelihood of picking a stock that outperforms. A final section concludes.

2. Equity holdings of retail investors

In this section, we describe the equity holdings of retail investors, focusing on the number of different stocks in retail investors' equity portfolios. Our data cover all stock market investors (above the age of 17) in Denmark. The data are kindly provided by the Danish Tax Authorities. The data are based on direct reporting from Danish financial institutions to the Danish Tax Authorities and consist of each individual's end-of-the-year equity holdings.[4] We analyse data for the latest year in our sample; 2012.[5]

We know the ISIN code of each equity investment of each Danish individual investor, the number of stocks the investor holds, as well as the end-of-the year value of the investment in Danish kroner. Our data contain equity holdings outside retirement accounts, i.e. non-restricted and freely accessible voluntary equity investments.[6]

2.2.   Summary statistics
There are 4,442,855 individuals in our sample. 3,429,685 do not hold stocks, leaving 993,170 stock market investors in Denmark. The stock market participation rate is thus 22.5%.

Danish stock market investors hold very concentrated portfolios, like in the US (for US evidence, see, e.g., Barber & Odean, 2000; Polkovnichenko, 2005; Ivkovic et al., 2008; and Goetzmann & Kumar, 2008). The average investor holds two different stocks. The

---

[4] The data are hosted by Statistics Denmark. Statistics Denmark assigns an anonymous personal identification number to each individual.

[5] We have access to data for the years 2005-2012. In order to present the main results in a clearer manner, we conduct our main analysis on 2012 data. We note that there is a weak tendency that our investors hold less underdiversified portfolios during the later parts of our sample period. In 2005, investors held 1.67 stocks on average. In 2012, as mentioned, they hold two stocks on average. Similar, in 2012, 65.9% of investors hold one stock, compared with 73.2% in 2005. This implies that our estimates based on 2012 data represent a lower bound on the cost of underdiviersifcation.

[6] It might be that pension savings are well-diversified, but when the investors cannot get access to these savings, it is relevant to evaluate the cost of insufficient diversification of the part of savings that the investor determines himself-herself, and has access to. To be precise, we present estimates of the cost of underdiversification on that part of the portfolio that the investor can control and that is accessible to the investor. It is a different question (and one that we do not know anybody has attempted to answer) to estimate the costs of underdiversification of the true total portfolio of an investor. The true total portfolio includes pension savings, real estate, human capital, and other non-marketable assets. We therefore followed the norms of the related literature, which, as mentioned in the Introduction, analyzes either investments outside pension savings only, or pension savings only, ignoring savings outside pension accounts, and from that make conclusions about investors' diversification.

distribution is skewed, i.e. many investors hold only one stock and relatively few hold well-diversified portfolios. Figure 1 shows the distribution of stock holders across the number of different stocks they hold. More than half of stock market investors, 66%, hold one stock only. 16% hold two stocks. 94% of investors hold less than five stocks, 1.3% hold between ten and twenty stocks, and as little as 0.2% hold more than twenty stocks. Danish retail investors diversify their direct equity holdings to a very little extend only.

The average value of the holdings of investors with one stock is DKK 17,300 (≈USD 3,000).[7] In Figure 2a, we show the average value of investments in stocks, across the number of stocks investors hold.[8] The figure reveals, not surprisingly, that investors with more wealth have a tendency to diversify more. Investors holding two stocks, for instance, have basically invested four times as much in stocks (DKK 68,000 ≈ USD 11,000) compared to investors holding one stock only. The average value of stock holdings for investors with three stocks is DKK 135,000 (≈ USD 22,500), and so on. The increase is almost monotonic. The average value for investors with ten stocks is app. DKK 500,000 (≈ USD 80,000).

The total value of equity holdings by Danish retail investors is DKK 75bn (≈ USD 12bn).[9] The aggregate value of equity holdings of investors with one stock is DKK 11bn (≈ USD 2bn). This amounts to 11bn/75bn = 15% of total investments. In other words, a large majority (66%) of retail investors own one stock only, but their total holdings amount to app. 15% of total holdings by retail investors.

Investors holding few stocks are typically investors with relatively low wealth, too. The average net financial wealth (consisting of the value of what they have in cash, bonds, stocks,

---

[7] At the time of writing, USD 1 buys app. DKK 6, i.e. DKK 17,300 corresponds to slightly less than USD 3,000.

[8] Data on the value of total stock holdings of individuals are winzorized at the 1% and 99% fractiles in the value distribution, so that outliers do not influence our results. We winzorize by replacing values of holdings (below) above the (1%) 99% fractiles with the value at the (1%) 99% fractiles. I.e., observations are not dropped but the values are replaced by marking up holdings below the 1% fractile to the value at the 1% fraction and holdings above the 99% fractile are marked down to the value at the 99% fraction.

[9] The official aggregate wealth statistics from Statistics Denmark do not split between holdings inside and outside retirement accounts. The official number for individuals' total combined stock holdings is DKK 131bn (≈ USD 22bn) in 2012. Our number, based on data on individuals' holdings outside retirement accounts, as mentioned amounts to DKK 75bn (≈ USD 12bn).

mutual funds minus their debt) of investors with one stock is app. DKK 200,000 (≈ USD 33,000).[10] Hence, investments in stocks on average accounts for 9% of the total net financial wealth of investors who hold one stock only. In Figure 2b, we show the value of investments in stocks as a fraction of total net financial wealth for investors with portfolios containing different numbers of stocks. The fraction of financial wealth invested in stocks increases with the number of stocks in the portfolios of investors, up until ten stocks. Thereafter, the fraction of wealth invested in stocks is flat, at around 40%.

3. Sample of stocks

Danish investors invest predominately in Danish stocks. Florentsen, Nielsson, Raahauge & Rangvid (2017) analyse home bias of Danish investors using the same data on stock market investments of retail investors as the data used here. They show that the average degree of home bias (an individual's holdings of Danish stocks as a fraction of the individual's total stock market investments) of the same Danish investors as those we investigate here is 86%. Hence, to investigate the costs of equity underdiversification, we use a sample of Danish stocks.

We examine returns on all Danish stocks for which we have monthly stock market returns throughout the last twenty years. There are 91 stocks that fulfil this requirement. The sample period is April 1997 through April 2017. We list in Table 1 the names of the stocks, the average monthly returns (annualized), and the annualized standard deviations of monthly returns. The stocks are listed according to their average returns, in descending order. There is a wide dispersion in average returns. Two stocks (DSV and AMBU) have returned an impressive 22.4% per year on average for the last twenty years. This means that DKK 100 invested in one of these two stocks in April 1997 have turned into almost DKK 6,000 twenty years later. This appears from Figure 3 that shows the cumulative value of all 91 stocks throughout our sample period. At the bottom of the list, Torm has returned an average return of -22.8% per year. This means that DKK 100 invested in April 1997 had

---

[10] Data on the total wealth of individuals come from register data hosted by Statistics Denmark, to which we have also access.

turned into a microscope DKK 0.56 in April 2017. So, if choosing one stock randomly in April 1997, one could have ended up with anywhere between DKK 0.56 and DKK 6,000 in 2017, or an average return anywhere between -22.8% and 22.4%.

There are 91 stocks in total. We want to compare the risk of a portfolio with $n = 1, 2, 3, \ldots,$ 91 stocks to the risk of a well-diversified benchmark. The benchmark we compare with is one where the investor in the first month of our sample (April 1997) invests an equal fraction in each of the 91 stocks, and keeps an equal fraction is each stock throughout the sample period. In other words, diversifying the investment across all stocks. The average annualised return to this market portfolio is 10.9% per year and the annualised standard deviation is 15.5%. We highlight the return to this market portfolio (MP) in Table 1.

The 91 stocks listed in Table 1 are those stocks for which we have data throughout the last twenty years, as mentioned. This means that stocks that were available in 1997:4 but left the market due to mergers, acquisitions, bankruptcy etc. are not included. Similarly, those that enter the market after 1997:4 are not included either. Some of these stocks would lower the return to an investor who bought an equally-weighted portfolio of all available stocks in 1997:4 (for instance stocks that left the market due to bankruptcy) whereas others would have increased the return to an investor buying all stocks during the same period (for instance successful new entries after 1997:4). It turns out that the return on our market portfolio closely matches the return to the total Danish stock market, as measured by, e.g., MSCI Denmark. MSCI Denmark has returned 11.8% per year on average over the last twenty years. We consequently view our market portfolio as a reasonable representation of the return to a broad Danish stock market index over the last twenty years. In Section 6.4, we discuss consequences of assuming different market portfolios.

4. Risk reduction by diversification
In this section, we illustrate the reduction in risk that diversification makes possible. We do so by calculating the risk (standard deviation) of different portfolios with different numbers

of randomly selected stocks, following the tradition of Evans & Archer (1968), Elton & Gruber (1977), Statman (1987, 2004), Campbell et al. (2001), and Domian et al. (2007).

If the portfolio contains one randomly chosen stock, the expected risk of the portfolio is the average standard deviation of the 91 stocks. The average level of risk across the 91 stocks is 38%. We also calculate the 5% and the 95% fractiles. Five percent of the stocks have a lower standard deviation than 20% and five percent have a higher than 70%.

In order to calculate the average level of risk and confidence bounds when there are more than one stock in the portfolio, we make 10,000 random draws.[11] The level of risk decreases dramatically when increasing the number of stocks in a portfolio. For a portfolio of two stocks, the average level of risk drops eight percentage points, to 30%. This is a reduction in the level of expected risk by 8%/38% = 21%. The effect on standard deviation from increasing the number of stocks in a portfolio is shown in Figure 4. By showing the relation between portfolio risk and number of stocks in a portfolio for a non-US stock market (in this case Denmark), the figure contributes to the literature showing similar relations for the US stock market. In Figure 4a, the x-axis is numerical whereas Figure 4b shows results on a logarithmic x-axis in order to improve readability of the effects. From the empirical distributions of standard deviations, we can also calculate the 5% and 95% fractiles. For the two-stock portfolio, the confidence interval narrows to [18%; 51%] from the before-mentioned one-stock portfolio confidence interval [20%; 70%]. When further increasing the number of stocks, the average level of risk drops and the confidence bands narrows even further. In Figure 4a and 4b, we also show the level of risk for the market portfolio. When the portfolio includes around ten stocks (in fact a little less, eight stocks), the 5% percentile overlaps with the risk of the market portfolio. It takes around 50 stocks before the level of risk converges to the level of risk of the portfolio including all 91 stocks. In other words, when a portfolio includes 50 stocks or so, practically all idiosyncratic risk has been

---

[11] The simulations work as follows. For the portfolio of two stocks, we randomly choose two stocks. We calculate the standard deviation of this portfolio over the sample period (1997:4 – 2017:4). We then randomly choose two stocks again, and calculate the standard deviation. We repeat this exercise 10,000 times, i.e. consider 10,000 portfolios of two stocks. From this, we calculate the average and the confidence bounds. We proceed similarly with 10,000 random draws of portfolios with three stocks, four stocks, etc.

eliminated. As mentioned in the Introduction to this paper, 50 stocks or so is also what Campbell, Lettau, Malkiel & Xu (2001) find is needed to eliminate idiosyncratic risk in the US stock market.

4.1. Calculating the cost of underdiversification

Figures like Figure 4 have been presented in the related literature (for other markets and sample periods, of course). We express the gain from diversification in terms of an expected return. The innovation in this paper is that we use this to study the aggregate gain from diversification, respectively the aggregate loss from underdiversification.

We need a model to convert reductions in risk to implications for expected returns. At least since the CAPM was developed in the 1960s, it has been well understood that investors are compensated for bearing systematic risk. We also know from the CAPM that investors can increase their expected return by levering up the market portfolio, i.e. moving up the capital market line. They will not increase the risk-return relationship of a well-diversified portfolio, but the level of expected return itself can be increased by increasing the level of systematic risk.

We have shown that our market portfolio has a standard deviation of 15.5%. We have also shown that if holding a portfolio with one stock only, the average level of risk is 38%. Instead of holding one stock, investors could invest in the market portfolio, lever this up to 38%, and expect higher return as a result.

We want to calculate how much investors could have increased expected returns by holding diversified portfolios. In order to do so, we need an estimate of the expected equity risk premium on the Danish stock market. Dimson, Marsh, and Staunton (2011) calculate the long-term equity premium (over the 1900-2010 period) in Denmark to be around 5%.[12]

---

[12] We show in Table 1 that the realized return on a well-diversified portfolio of Danish stocks throughout the last twenty years is close to 11%. The interest rate on Danish long-term government bonds have been low through the last twenty years, steadily falling from around five percent in 1997 to close to below one percent in 2017. The average has been 3.8%. Short-term interest rate have been even lower. The realized equity premium over the last twenty years has thus been at least seven percent (11% - 3.8%). By using the historical long-run

Figure 5 illustrates how we proceed, using the one-stock portfolio as an example. We assume a risk-free rate of 2% and a risk-premium of 5%, i.e. an expected equity return of 7%. The market portfolio is the intersection of 7% return and 15.5% risk. The capital market line connects the assumed interest rate of 2% and the market portfolio. The one-stock portfolio has an expected standard deviation of 38%. If the investor, instead of holding the one-stock portfolio that contains a significant fraction of idiosyncratic, and hence a significant fraction of risk that is not compensated by expected return, levers up the market portfolio to 38% risk, the expected return is $0.02 + 0.38 \cdot (0.07 - 0.02)/0.155 = 14.3\%$. Given that we have assumed an expected equity return of 7%, this implies that if an investor – instead of choosing a portfolio with one stock – decides to hold the market portfolio with the same 38% level of risk, the investor should expect 7.3 percentage points (14.3% - 7%) higher return. This is the distance from point A to point B in Figure 5. Another way of saying this is that by holding only one stock in the portfolio, the investor foregoes 7.3 percentage points in expected excess return, for the same level of risk. Or, in again other words, the extra expected excess return that is needed on a portfolio with one stock to make the investor indifferent between this portfolio and the market portfolio is 7.3 percentage points.[13]

equity premium as the expected equity premium in our calculations, we provide a lower bound on the cost of underdiversification.

[13] It is beyond the scope of this paper to calculate actual returns on individuals' equity holdings, not least because our data cover a relatively short period. We could, as an alternative in order to come closer to actual returns on underdiversified portfolios, make 10,000 draws of one random stock from our sample of 91 stocks and calculate the average return on such one-stock portfolios. The average from such draws, however, will be equal to the return on the equal-weighted portfolio. Similar, if drawing two stocks randomly, the average will equal the equal-weighted portfolio return, and so on for three-stock, four-stock, etc. portfolios. For these reasons, we assume that investors' losses from underdiversified portfolios equal the difference between the return on the market portfolio and a portfolio where investors lever up the market portfolio. The assumptions we implicitly rely on here is that investors do not possess systematic stock-picking skills and do not get compensation for holding idiosyncratic risks. This means that we assume that investors randomly choose individual stocks (when choosing stocks randomly, investors on average hold the equal-weighted market portfolio). Instead, the robustness tests we make is to calculate the aggregate loss from underdiversification using different assumptions about the market risk premium (Tables 2 and 3). Furthermore, in Section 8, we show that underdiversified investors hold portfolios that on average resemble the market portfolio.

We can perform the same calculation for portfolios with two, three, four, etc. stocks, based on the levels of risk of these portfolios that we showed in Figures 4a and 4b. We can calculate confidence bounds as well. The results are shown in Figure 6a (numerical x-axis) and 6b (log x-axis). For portfolios of two randomly selected stocks, the excess return required to make the investor indifferent between the portfolio and the market portfolio is 5 percentage points, for three stocks 3.5 percentage points, etc. For 50 stocks or more, the excess required excess return is close to zero.

5. The aggregate cost of underdiversification

A key insight from the previous section is that an investor who holds one stock in her portfolio could have increased expected excess return on her equity portfolio (without taking on additional risk) by slightly more than seven percent by diversifying. The value of stock holdings for the average investor who holds one stock is DKK 17,300 (≈USD 3,000), as shown in Figure 2a. This means that the average investor can expect to lose close to DKK 1,260 (≈USD 200) per year due to underdiversification. An investor who holds two stocks can expect to lose five percent per year by holding an underdiversified portfolio (Figure 6b). The average value of stock holdings of investors with two stocks is DKK 68,000 (≈USD 11,000). This means that the expected annual loss from underdiversification is app. DKK 3,200 (≈USD 500). And so on for the investor with three stocks, four stocks, etc. We plot the annual loss (in DKK) for the average investor in Figure 7. The expected loss follows an inverted *u* shape. The reason is as follows. The average loss *in percent* is reduced when increasingly diversifying portfolios, as illustrated in Figure 6. The *average value* of the portfolio is increasing for increasingly diversified portfolios, i.e. for portfolios with a higher number of stocks (Figure 2). The increase in value when adding one stock to a portfolio is exceeding the decline in the loss in percent, for portfolios with less than eleven stocks. For portfolios with more than eleven stocks, the opposite is true.

There are app. 650,000 investors with one stock. They loose on average DKK 1,260 (≈USD 200) due to underdiverisfication, as just mentioned. The aggregate loss for all investors with

one stock is thus slightly more than DKK 800 million (≈USD 130mn). Figure 8 plots the aggregate loss for portfolios with different numbers of stocks.

We can aggregate these losses. The aggregate annual loss of the Danish population of stock holders due to underdiversification is DKK 2.3 billion (≈USD 400mn).

With 650,000 investors holding one stock, and each investing app. DKK 17,300 (≈USD 3,000) in stocks, the total value of stock holdings of all investors who hold one stock is DKK 11.3bn (≈USD 2bn). Repeating this calculation for all portfolios, the total value of stock holdings is DKK 75bn (≈USD 12bn). This means that the annual loss of DKK 2.3bn (≈USD 400 million) due to underdiversification corresponds to app. 2.3/75 = 3.1% of the value of stock holdings. Given that we have assumed an equity premium of five percent, a possibility to increase the expected annual return by around 3.1 percentage points per year without taking on additional risk, simply by moving from underdiversified portfolios to diversified portfolios, is remarkable.

We noted in Section 2 that 66% of investors own one stock only. We also noted that these are small investors, as they own app. 15% of total stock market investments of Danish retail investors. We can calculate the fraction of the aggregate loss due to underdiversification that these investors bear. When the aggregate loss amounts to DKK 2.3bn (≈USD 400mn) and the loss of these investors amounts to more than DKK 800mn (≈USD 130mn), they bear 36% of losses. In other words, investors with concentrated portfolios own a small fraction of invested wealth, but account for a large fraction of the losses.

We recognize that the relevant alternative for most retail investors is probably not to buy 91 stocks, but buying a mutual fund. An index fund that tracks the Danish stock market costs around fifty basis points in annual fees. This means that investors could improve their expected return by close to 3.1% - 0.5% = 2.6% per year, without taking on additional risk, if buying an index fund instead of holding underdiversified portfolios.

Our calculations are based on an expected return from the stock market of 7% p.a. If expected returns are higher (lower), the losses will be higher (lower), too. If, e.g., we assume

an expected risk premium of 6%, i.e. an equity return of 8%, the annual loss is 3.7%. If the expected risk premium is assumed to be 4%, and equity returns thus 6%, the annual loss is 2.5%. Table 2 presents these calculations for different values of the expected risk premium, ranging from 2% to 8%. In the table, we highlight the baseline case of 5%. Table 2 also shows results from using the 5% and 95% fractiles of the distributions of expected standard deviations of the portfolios with *n* stocks from Figure 4, instead of the mean standard deviations. These calculations thus provide confidence intervals for the estimates of the expected losses. In our baseline case with a risk premium of five percent, the 5% fractile of the aggregate loss is 0.4%, whereas the 95% fractile is 7.7%.

6. Discussions of results

6.1. Excluding investors holding mutual funds

26% of stock holders also hold mutual funds. Mutual funds are well-diversified per regulation. An investor who invests a small fraction directly in stocks but a large fraction in mutual funds would not be underdiversified, even if holding only, e.g., one stock directly. To make a perspective on this, we conduct the same calculations as those above for investors who do not hold mutual funds, i.e. investors who only hold stocks directly. We first stress that the major fraction of stock market investors do not hold mutual funds. 74% of stock holders do not hold mutual funds. These 737,782 stock holders, thus, do not gain stock market diversification via mutual funds. Conducting the same calculations as above for the sample of stock holders who do not hold mutual funds, we find that the annual loss due to underdiversification is close to DKK 1.4bn. The total amount of stock holdings of these investors DKK 38bn. This implies that these investors in aggregate could improve their return by 3.7% per year, if shifting from underdiversified portfolios to well-diversified portfolios.

Calvet, Campbell, and Sodini (2007) show that the average loss of Swedish retail investors due to underdiversification is low when taking into account their savings in mutual funds. We show here that most Danish retail investors do not hold mutual funds. We also show that when focusing on this large majority of investors, the losses from underdiversification

are still significant. This implies that a sensible conclusion based on the findings of Calvet, Campbell, and Sodini (2007) and those reported here is that they complement each other well: If many investors invest relatively large fractions in mutual funds, these investors end up holding portfolios with low idiosyncratic risks, even when they hold a low number of different individual stocks. On the other hand, if only few investors diversify via mutual funds, the aggregate costs of underdiversification can be sizeable. Most of our investors do not hold mutual funds.

6.2. Asset pricing implications of portfolio shifts

When calculating the aggregate loss, we implicitly assume that the expected risk premium remains at 5% when investors replace their current underdiversified portfolios with a levered market portfolio. There are four main reason why this is a reasonable assumption. First, it is the standard assumption in the CAPM. In the CAPM, investors can move up and down the capital market line (assuming no borrowing and lending constraints), given the expected risk premium. Our investors, as seen in Figure 2b, have substantial savings outside stocks, i.e. can easily replace some of their cash or bond holdings with stock holdings. Second, the marginal investor determines the market risk premium. Given that Denmark is a small economy with open capital markets, this marginal investor is the marginal world investor. Our calculations deal with Danish retail investors. Third, and related to the previous point, the market value of listed Danish stocks is app. DKK 1,300bn (≈USD 200bn) in 2012. Danish retail investors thus hold only app. $75/1,300 = 6\%$ of all listed Danish equity. This means that Danish retail investors constitute a relatively small fraction of total investments in the Danish stock market. Finally, if capital is elastic, such that firms issue more capital when demand for capital increases, the expected risk premium remains constant no matter the degree of leverage of the market portfolio (as in the CAPM).

6.3. Heterogeneity across individuals' performance

Conditioned on the number of stocks in investors' portfolios, we implicitly assume that all investors are identical. In other words, we assume that all investors holding, e.g., one stock would benefit from shifting from undiversified to diversified portfolios. It is outside the scope of this paper to examine in detail the performance of each individual's portfolio.

Related literature gives us some hints regarding heterogeneity among investors, however. Goetzmann & Kumar (2008) find results supporting our assumption that individual investors with better diversified portfolios perform better. Examining around 60,000 individual investors at a large U.S. discount brokerage during the 1991-1996 period, they report for example that (p. 457): "*On an annual basis, the most diversified investor group earns a 2.04% higher return than the least diversified investor group.*" Using the same data, but a different method, Ivkovic et al. (2008) find that there is no difference in performance between the average investor with a concentrated portfolio and the average investor with a diversified portfolio. When analyzing a small subset (less than 10%) of investors with high wealth allocated to the stock market (above USD 100,000), they find that investors holding concentrated portfolios outperform investors holding diversified portfolios.

6.4. Market portfolios

Our calculations of the aggregate costs of diversification are based on the assumption that the equal-weighted portfolio of our 91 individual stocks represents an efficient market portfolio. Other portfolios might represent a "market portfolio", however.

For instance, when we use the equal-weighted portfolio of the 91 stocks in Table 1, we implicitly assume that Danish investors diversify across Danish stocks only. There are other broad "market portfolios" based on Danish stocks, however, such as the MSCI Denmark or the FTSE Denmark market portfolios. If one instead assumes that Danish investors invest internationally, an international portfolio would be a better candidate for the market portfolio. This could for instance be a European portfolio (e.g., the MSCI Europe, the STOXX 600, or the STOXX 50) or a global portfolio (e.g., MSCI World, the MSCI AC World, or the FTSE All World). In this section, we discuss the consequences for our main estimates if using these other market-portfolio candidates.

What matters for our estimates of the aggregate cost of diversification, as explained in Figure 5, is the standard deviation of the assumed market portfolio. Table 3 shows annualized standard deviations from the equal-weighted portfolio of the 91 stocks that we use in our benchmark calculations, together with other relevant candidates for market portfolios of

Danish investors. All standard deviations are based on monthly returns in Danish kroner from April 1997 through April 2017.

Table 3 presents calculations of the aggregate loss from diversification if using different broad market portfolios. We find that other broad Danish market portfolios have a higher standard deviation than the equal-weighted average of the 91 stocks from Table 1, and consequently a smaller aggregate loss from underdiversification than in our baseline calculations. On the other hand, and perhaps theoretically more correct, if one assumes that Danes can invest internationally (which they of course can), the true market-portfolio would be an international portfolio. The standard deviations of the global portfolios are close to the standard deviations of the equal-weighted portfolios, implying that the aggregate loss is close to, but slightly lower, than our benchmark results. No matter the exact choice of market-portfolio, however, the conclusion is that the aggregate loss is substantial and close to three percent.

7. Who bears the cost of underdiversification?

A natural question to ask is who bears the cost of underdiversification? Or, in other words, whether certain types of investors are more likely to underdiversify? To help answering this, we split investors according to a number of criteria: gender, age, education, income, wealth, and investors holding mutual funds. Table 4 shows for each investor category the total loss from underdiversification, the total stock holdings of each type of investor, and the loss in percentage terms. We also show the number of each type of investor.

To explain the table, consider for instance gender of the investor. 458,024 (535,146) women (men) hold stocks directly. The total value of the direct stock holdings of women (men) is DKK 26,673 (48,066) mill. Women (men) loose DKK 951 (1,379) mill. from underdiversification. This corresponds to a loss in percentage for women (men) of 3.6% (2.9%). In total, this means that women allocate a smaller fraction of capital to the stock market, but lose a larger fraction of their stock holdings due to underdiversification, compared to men.

Turning to other investor types, Table 4 show that investors with low income, low wealth, low level of education, and investors who do not hold mutual funds lose more (as a fraction of their invested wealth). Simply stated, young women with low education, low income, low wealth, and no mutual funds bear higher costs of underdiversification than men with high education, income, wealth, and who hold mutual funds. Calvet, Campbell, and Sodini (2009) classify investors according to their level of financial sophistication. They find that investors with low education, low income, and low wealth are investors with a low level of financial sophistication. We, thus, also find that investors with low levels of financial sophistication tend to bear higher costs of underdiversification.

Investors with low income and wealth allocate less capital to the stock market. The 33% wealthiest investors have allocated app. DKK 55bn to the stock market, whereas the 33% least wealthy investors have allocated app. DKK 6bn. This means that even when rich investors hold better diversified portfolios, they still bear app. 2/3 of the aggregate loss due to underdiversification. Similarly, those with higher income hold better diversified portfolios, and bear lower costs relatively to what they have invested in the stock market, they still bear the larger fraction of the aggregate loss due to diversification. In total, rich investors diversify their stock holdings better, but still bear the majority of the aggregate loss from underdiversification, simply because they allocate more capital to the stock market.

7.1   Probit regression

Income, wealth, and education are correlated. For instance, investors with low income tend to be investors with low education, too. In order to control for background characteristics, i.e. separate out the effect of one variable while controlling for the rest, columns (1) and (2) of Table 5 present results from probit regressions where we relate the likelihood of holding 1 stock only, or no mutual funds, to background characteristics of the investors. Investors holding 1 stock only are obviously very underdiversified. Similarly, holding mutual funds in itself generates diversification (cf. Section 6.1). Hence, in Table 5, we identify types of investors who are more likely to hold only one stock or no hold mutual funds, i.e. investors holding portfolios with low levels of diversification, controlling for other characteristics of the investors.

To explain Table 5, consider the coefficient to the gender dummy, estimated at -0.07. This means that men are seven percent less likely to hold one stock only, compared to women, controlling for income, wealth etc. Similarly, the table shows that young investors, investors with high education, high income, high wealth, and those investors also holding mutual funds are less likely to hold one stock only. These types of investors are also more likely to hold mutual funds, as results in column (2) show. I.e., column (2) shows that investors with high income, high level of wealth, and a high-level education are more likely to hold mutual funds. Columns (1) and (2) together thus reveals that investors with low income, low levels of wealth, and low levels of education are more likely to hold underdiversified portfolios.

The effect of gender might seem surprising at first sight. Typically, the literature finds, men tend to take higher risk on the stock market, see, e.g., Barber and Odean (2001). We find that women are more likely to hold one stock only. However, as just mentioned, we also find (Table 5) that women are more likely to hold mutual funds. Hence, women are more likely to hold a well-diversified portfolio (mutual funds) in the first place. One interpretation is that women already hold a well-diversified portfolio, and then bet on individual stocks. This is from a diversification point of view not as harmful as betting on individual stocks but holding no mutual funds.

7.1.1. Employee stocks
Overall, we argue in this paper that surprisingly many investors hold very underdiversified portfolios. One potential reason why investors hold one stock only could be that they have received that stock "automatically", i.e. might not be interested in stocks as such. One such group of investors could be investors with employee stocks. A hypothesis is that some investors with employee stocks are not interested in buying stocks per se, but simply hold that single stock of their company because it is part of a compensation package.

In our data, around 100,000 (97,609) investors hold employee stocks. This means that ten percent of investors hold employee stocks. Given that nine out of ten investors hold less than five stocks, and only one out of ten holds employee stocks, we already learn from these number that employee stocks cannot explain underdiversification. Furthermore, it turns out

that those investors holding employee stocks are in fact less likely to hold one stock only, and more likely not to hold mutual funds, compared to other investors, as columns (3) and (4) of Table 5 shows.

7.1.2. Foreigners as investors in Danish stocks. The information channel

As mentioned, Calvet et al. (2009) argue that investors with high income, wealth, and education are sophisticated investors, i.e. investors with knowledge about financial markets. These investors diversify better because they have more information about financial markets.

We study Danish stocks in this analysis. A reasonable hypothesis is that investors being born and raised in Denmark know more about Danish stocks. Basically, if an investor has lived in Denmark all his/her live, this investor understands Danish and can thus follow the Danish press. This investor will automatically learn about Danish stocks, simply because stories about Danish companies appear in the Danish media. On the other hand, foreigners who have reallocated to Denmark most likely know comparably less about Denmark, Danish companies, and Danish stocks. We can thus proxy for the level of information about Danish stocks by identifying relocated foreigners in our data.[14] Our testable hypothesis is that foreigners do not buy as many Danish stocks, i.e. tend to be less diversified, because they have an information disadvantage (compared to Danish investors), as foreigners do not know as much about Danish stocks.

We test this hypothesis in columns (5) and (6) of Table 5. We extend the probit regression by adding a dummy that takes the value of one if the investor is a foreigner, defined as an investor immigrated to Denmark. The estimated coefficient to this dummy is positive. This means that foreigners are more likely to hold one stock only and less likely to hold mutual funds. Together with the finding that investors with high income, wealth, and education hold better diversified portfolios, this indicates that information plays a role for the degree of underdiversification that investors display.

---

[14] Florentsen, Nielsson, Raahauge and Rangvid (2018) use this idea to investigate home bias of relocated foreigners.

8. How lucky can you be?

One reason why investors choose to hold concentrated portfolios, could be that they believe that they can identify the "winner-stock". In this section, we give a perspective on how likely this is. We do so by comparing the returns obtained if choosing a certain number of stocks in 1997:4 to the return an investor would have obtained if she had invested in the market portfolio of our 91 stocks. We consider six cases. The first case is where the investor randomly chooses to invest in one stock. Next case is one where the investor randomly chooses two stocks. Then five stocks, ten stocks, twenty stocks, and finally the case of fifty stocks. For each case, when there is more than one stock in the portfolio, we make 10,000 random draws. We scale returns by the return on the market portfolio, and multiply by 100. The result is a distribution of scaled returns. A portfolio that delivers a scaled return below (above) "100" has delivered an average return below (above) the return of the market portfolio over the sample period. We show the resulting distributions in Figure 9a-8f.

The main point to notice from Figure 9 is that when moving from concentrated portfolios to diversified portfolios, i.e. when moving from Figure 9a to Figure 9f, the distribution changes from an asymmetric distribution towards a symmetric. This means that the figures together show how the likelihood of underperforming the market portfolio decreases when the portfolio contains more stocks, as the probability density gets less and less skewed when going from figure 9a to 9f. To explain the figures in more detail, consider Figure 8a first. This is the case of randomly choosing one stock.[15] The figure reveals that there is a high likelihood of obtaining a lower return than the return on the market portfolio, as the mass of the distribution is concentrated to the left of "100". In fact, there is a 74% chance that an investor ends up with a return that is below the return on the market portfolio. With one stock, this is the chance of picking one of the 67 stocks (out of the total of 91 stocks) that underperform the market portfolio (Table 1 reveals the 67 stocks that yielded lower return than the market portfolio). This is a remarkable finding in light of the fact that one stock is the typical number held by our investors.

---

[15] No simulations are done in this case. This is simply the distribution of returns of the 91 stocks, compared to the return on the market portfolio.

When investors start diversifying, the likelihood that they underperform the market decreases. If randomly choosing two stocks (Figure 9b), the likelihood of obtaining a return below the market return drops to 66%. This is still a high chance of underperformance but it is nevertheless a lower likelihood of underperforming than if choosing one stock only. If randomly choosing five stocks, the likelihood of underperforming drops further to 58%, and so forth when increasing diversification further. When randomly choosing 50 stocks out of the universe of 91 stocks, the return distribution approaches a symmetric distribution (Figure 9f) with a mean, median, and modal return very close the return on the market portfolio. With a well-diversified portfolio, there is an equal chance of underperforming and outperforming the market. The range of outcome narrows as well when the number of stocks in the portfolio increases. Whereas there is app. 2% chance of investing in stocks that outperform the market by more than 700% when investing in a single stock only, the maximum outperformance is around 40% when investing in fifty stocks. On the other hand, the probability of obtaining a very low return also drops when there are more stocks in the portfolio. The conclusion from this exercise is that when concentrating the portfolio to one stock, the chance of lousy performance is high. Diversification eliminates such lousy performance. It also eliminates the possibility of spectacular performance, but, when choosing few stocks, the likelihood of good performance is much lower than the likelihood of lousy performance. Using data from another market and in this sense "out-of-sample", these results provide support of the results in Bessembinder (2017) that less than three out of seven US stocks provide return that exceed the return from a short Treasury Bill, and that just 4% of US stocks have accounted for the entire US net stock market gain since 1926. The results here, and in Bessembinder, indicate that it is difficult to find the few well-performing stocks among the many stocks out there.

Do these results imply that our investors (who typically hold very few stocks) are worse off in utility terms? Not necessarily, as suggested by Barberis & Huang (2008). If investors have preferences for lottery stocks, then investors might take relatively large positions in few stocks in order to add skewness to their portfolios. Or, in other words, if investors view stocks as lottery tickets, and have preferences for such lotteries, investors might be fine

investing in few stocks with such lottery-like expected pay off. In this case, investors do not wish to diversify away this outcome. The point in this section, though, is that this outcome (finding a "winner stock") is a low-probability event.

9. Conclusion

We analyse equity holdings of all retail investors in a country. We find that the typical investor holds only one stock in her portfolio. The average is two stocks, and less than two percent of investors hold more than ten stocks in their portfolios. In essence, retail investors hold underdiversified equity portfolios. We calculate the reduction in risk an investor can achieve if moving from a portfolio of n stocks to a fully-diversified portfolio including all stocks. We find that if moving from a portfolio with one randomly selected stock to a well-diversified portfolio, 60% of risk can be eliminated.

We convert the reduction in idiosyncratic risk to a loss in expected return. Given that we know the equity portfolio of all investors in a country, we can calculate the total loss, in terms of foregone expected returns, due to underdiversification. We find that the loss amounts to 3.1 percentage points per year in our baseline calculations.

We study who bears the aggregate cost of underdiversification. We find that investors with low education, low income, and low wealth are more likely to underdiversify. We also find that investors in the top of the wealth and income distribution bear a larger fraction of the aggregate cost of underdiversification, in spite of being more diversified, as these investors allocate relatively more capital to the stock market. This means that rich investors, per dollar allocated to the stock market, have lower losses due to underdiversification, but when adding all the dollars they allocate, their aggregate loss is larger than the aggregate loss of investors in the lower end of the income/wealth distribution.

An interesting policy implication of our results is that financial education of the majority of stock holders who hold very concentrated portfolios will not significantly reduce the aggregate costs of underdiverisifcation. The reason is as follows. Financial education has the potential to make investors who otherwise hold underdiversified portfolios shift to

diversified portfolios. This will bring down the costs of underdiversifiation for these investors. But the majority of investors holding very concentrated portfolios at the same time hold a lower fraction of the aggregate wealth allocated to the stock market. To bring down the aggregate costs of underdiviersification, all investors – also those who are already reasonably, but not fully diversified – should shift to fully diversified portfolios

Another main policy conclusion is that many retail investors would benefit from shifting ownership of stocks directly to ownership via mutual funds. Mutual funds charge fees for their services, but our calculations show that the loss from underdiversification exceed the typical fee charged by mutual funds by a large margin.

References

Agnew, Julie, Pierluigi Balduzzi, and Annika Sunden (2003). Portfolio choice and trading in a large 401(k) plan. *American Economic Review* 93, 193–215.

Barber, Brad M., and Terence Odean (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *Journal of Finance* 55, 773–806.

Barber, Brad M., and Terence Odean (2000). Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics* 116, 261-292.

Barberis, Nicholas, and Ming Huang (2008). Stocks as lotteries: The implications of probability weighting for security prices. *American Economic Review* 98, 2066–2100.

Benartzi, S., and Thaler, Richard H. (2001). Naive diversification strategies in retirement saving plans. *American Economic Review* 91, 79–98.

Benjelloun, Hicham (2010). Evans and Archer − forty years later. *Investment Management and Financial Innovations* 7, 98-104.

Bessembinder, Henrik (2017). Do stocks outperform Treasury Bills? *Journal of Financial Economics*, forthcoming.

Bloomfield, Ted, Richard Leftwich, and John B. Long, Jr. (1977). Portfolio strategies and performance. *Journal of Financial Economics* 5, 201-218.

Blume, Marshall, and Irwin Friend (1975). The asset structure of individual portfolios and some implications for utility functions. *Journal of Finance* 30, 585–603.

Calvet, Laurent E., John Y. Campbell, Paolo Sodini (2007). Down or out: Assessing the welfare costs of household investment mistakes. *Journal of Political Economy* 115, 707-747.

Calvet, Laurent E., John Y. Campbell, Paolo Sodini (2009). Measuring the financial sophistication of households. *American Economic Review P&P* 99, 393−398.

Campbell, John Y., Martin Lettau, Burton Malkiel, and Yexiao Xu (2001). Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *Journal of Finance* 56, 1-43.

Dimson, Elroy, Paul Marsh, Mike Stauton (2011). Equity premia around the world. *Manuscript*, London Business School.

Domian, Dale L., David A. Louton, and Marie D. Racine (2007). Diversification in portfolios of individual stocks: 100 stocks are not enough. *Financial Review* 42, 557–570.

Elton, Edwin J. and Martin J. Gruber (1977). Risk reduction and portfolio size: An analytical solution. *Journal of Business* 50, 415–437.

Evans, John L. and Stephen H. Archer (1968). Diversification and the reduction of dispersion: An empirical analysis. *Journal of Finance* 23, 761-767.

Florentsen, Bjarne, Ulf Nielsson, Peter Raahauge, and Jesper Rangvid (2017). Turning local. Home-bias dynamics of relocating foreigners. *Manuscript*.

Gaudecker, Hans-Martin von (2015). How does household portfolio diversification vary with financial literacy and financial advice? *Journal of Finance* 70, 489-507.

Goetzmann, William N. and Alok Kumar (2008). Equity portfolio diversification. *Review of Finance* 12, 433–463.

Ivkovic, Zoran, Clemens Sialm, and Scott Weisbenner (2008). Portfolio concentration and the performance of individual investors. *Journal of Financial and Quantitative Analysis* 43, 613–655.

Kelly, M. (1995). All their eggs in one basket: Portfolio diversification of U.S. Households. J*ournal of Economic Behavior and Organization* 27, 87–96.

Liu, Hong (2014). Solvency constraint, underdiversification, and idiosyncratic risks. *Journal of Financial and Quantitative Analysis* 49, 409–430.

Polkovnichenko, Valery (2005). Household portfolio diversification: A case for rank dependent preferences. *Review of Financial Studies* 18, 1467–1502.

Roche, Hervé, Stathis Tompaidis, Chunyu Yang (2013). Why does junior put all his eggs in one basket? A potential rational explanation for holding concentrated portfolios. *Journal of Financial Economics* 109, 775–796.

Statman, Meir (1987). How many stocks make a diversified portfolio? *Journal of Financial and Quantitative Analysis* 22, 353–363.

Statman, Meir (2004). The diversification puzzle. *Financial Analysts Journal* 60, 44–53.

Table 1. List of stocks.

The table shows the list of stocks included in the analysis, together with average monthly returns (annualized) and annualized standard deviation of monthly returns. **MP** is the market portfolio.

Name	Return	STD	Name	Return	STD
DSV	22.4%	32.1%	TIVOLI	6.4%	17.1%
AMBU	22.4%	33.2%	INVSTSSL.LUXOR	6.4%	46.0%
GABRIEL HOLDING	19.3%	26.9%	GYLDENDAL 'A'	6.1%	41.3%
COLOPLAST	19.0%	21.3%	FLSMIDTH & CO.	6.0%	37.2%
RINGKJOBING LANDBOI	18.8%	23.3%	RIAS	5.8%	23.7%
DMPKBT.NORDEN	18.5%	41.9%	HVIDBJERG BANK	5.6%	35.4%
NOVO NORDISK	18.5%	26.4%	BRODRENE HARTMANN	5.6%	26.7%
DFDS	17.5%	33.5%	GYLDENDAL 'B'	5.4%	26.9%
TOPDANMARK	16.9%	24.5%	CARLSBERG 'B'	5.4%	28.5%
BRD KLEE	16.4%	35.1%	CARLSBERG 'A'	5.1%	27.3%
KOBENHAVNS LUFTHAV	16.1%	26.7%	SOLAR	5.1%	155.4%
AKTKT.SCHOUW & CO.	15.6%	35.7%	DANTAX RADIO	4.2%	27.0%
NKT	15.3%	48.6%	ARKIL HOLDING	4.0%	27.9%
WILLIAM DEMANT	14.4%	29.9%	FLUGGER	3.9%	26.7%
GRONLANDSBANKEN	14.4%	24.6%	SKJERN BANK	3.7%	29.2%
ROYAL UNIBREW	14.1%	43.6%	SALLING BANK	3.6%	26.1%
SYDBANK	13.5%	28.8%	SANTA FE GROUP	3.5%	35.8%
UNITED INTL.ENTS.	13.2%	32.4%	TOTALBANKEN	3.3%	47.8%
JEUDAN	12.3%	17.5%	ALM BRAND	3.0%	33.1%
SPAR NORD BANK	11.5%	25.3%	MONBERG & THORSEN	2.4%	37.2%
JYSKE BANK	11.1%	29.1%	MIGATRONIC	1.9%	30.7%
DANSKE BANK	10.9%	30.9%	HARBOES BRYGGERI	1.8%	25.5%
ROBLON	10.8%	26.9%	SKAKO	1.3%	37.6%
GN STORE NORD	10.8%	43.6%	GREENTECH ENERGY SY	1.2%	63.9%
			IC GROUP	-1.4%	46.2%
MP	*10.9%*	*15.5%*	SCANDINAVIAN BRAKE S	-1.7%	44.9%
			BANG & OLUFSEN	-1.8%	47.2%
LOLLANDS BANK	10.4%	26.5%	H&H INTERNATIONAL	-2.2%	43.3%
TDC	9.8%	29.2%	NORTH MEDIA	-3.0%	47.7%
DJURSLANDS BANK	9.7%	19.6%	VESTJYSK BANK	-4.0%	41.4%
NORDJYSKE BANK	9.2%	21.5%	NEWCAP HOLDING	-6.1%	78.3%
ROCKWOOL	9.2%	32.2%	TK DEVELOPMENT	-6.2%	51.9%
KREDITBANKEN	9.0%	18.2%	PARK STREET NORDICOI	-6.4%	56.5%
ALK-ABELLO	9.0%	33.1%	GLUNZ & JENSEN HOLDI	-6.6%	42.9%
ROCKWOOL	8.9%	31.7%	NTR HOLDING	-6.9%	31.1%
PER AARSLEFF	8.7%	29.5%	SANISTAL	-7.7%	32.1%
MONS BANK	8.6%	20.4%	INTERMAIL	-8.8%	37.7%
NORDFYNS BANK	8.3%	22.3%	DALHOFF LAR.& HORNE	-9.3%	46.8%
HOJGAARD HOLDING 'A'	8.1%	32.7%	OSTJYDSK BANK	-12.2%	46.9%
HOJGAARD HOLDING 'B'	7.9%	31.7%	SILKEBORG IFS	-12.8%	49.6%
A P MOLLER - MAERSK 'B	7.6%	32.8%	BRONDBY IF	-12.8%	55.8%
A P MOLLER - MAERSK 'A	7.4%	32.2%	BIOPORTO	-14.0%	70.3%
EGETAEPPER	7.0%	27.0%	ARHUS ELITE	-14.9%	117.5%
SP GROUP	6.7%	43.5%	CEMAT	-17.1%	74.8%
LAN & SPAR BANK	6.7%	12.1%	NEUROSEARCH	-18.6%	69.9%
FE BORDING	6.6%	28.0%	VICTORIA PROPERTIES	-18.7%	64.1%
ANDERSEN & MARTINI	6.5%	33.1%	TORM	-22.8%	62.8%

Table 2. Aggregate return needed to justify holding underdiversified portfolios. Varying the expected risk premium.

The table shows the aggregate loss from underdiversification as a fraction of total investment in stocks for different levels of the expected risk premium. The table shows the expected loss as well as the 5% and 95% fractiles of the distribution of standard deviations (from Figure 4) used when calculating the losses.

Risk premium	2%	3%	4%	**5%**	6%	7%	8%
Expected	1.2%	1.9%	2.5%	**3.1%**	3.7%	4.4%	5.0%
5% percentile	0.2%	0.2%	0.3%	**0.4%**	0.5%	0.6%	0.6%
95% percentile	3.1%	4.6%	6.2%	**7.7%**	9.3%	10.8%	12.3%

Table 3. Aggregate loss with different choices of market portfolio

The table statistic includes all Danish stockholders in 2012. STD is annualized standard deviation and *Aggregate loss* is the estimated annual loss in DKK millions due to underdiversification based on different proxies of the "market portfolio".

Portfolio	STD	Aggregate loss
Equally-weighted portfolio of 91 stocks	15.5%	3.1%
FTSE Danish market	18.6%	1.8%
MSCI Denmark	19.3%	1.5%
MSCI Europe	17.0%	2.4%
STOXX EUROPE 600	16.6%	2.6%
STOXX EUROPE 50	16.5%	2.6%
MSCI WORLD	16.0%	2.9%
MSCI AC WORLD	16.2%	2.8%
FTSE ALL WORLD	16.3%	2.7%

Table 4. Who foregoes the most from underdiversification?

The table statistic includes all Danish stockholders in 2012. *Loss from underdiv.* is the estimated annual loss in DKK millions due to underdiversification, aggregated across types (e.g. across all males). *Foregone return* is this aggregate loss as a fraction of the value of *Stock holdings*. Young/old are those below/above the median age of 55 years. Higher education is defined as those with a university degree, which applies to 12% of stockholders. The income variable is defined as total income before taxes. This includes regular salary, pension, public income transfers, irregular income (e.g. honorary income, consulting income, etc.), income from self-owned firm, capital income, foreign income, etc. Low/medium/high income refers to those with income below/between/above the 33%/33-67%/67% fractile. Same is true for wealth categories, where wealth is defined as net wealth (gross wealth minus debt) at year-end, excluding pension savings. All other registered wealth is included, such as the value of cash, stocks, bonds, mutual funds, net house wealth, etc. (value of motor vehicles and boats is unregistered, though). 76% of stockholders do not hold mutual funds. The last row shows the total results across each category, e.g. 951m+1,379m losses across both genders accumulated to 2,330m. Variables in DKK terms are winsorized within each year at the 1st and 99th percentiles. Statistical inference is not reported since the figures are not estimated from a sample but represent actual values calculated from the full population.

	Number of stockholders	Loss from underdiv. (mill.)	Stock holdings (mill.)	Foregone return
Female	458,024	951	26,673	3.6%
Male	535,146	1,379	48,066	2.9%
Young	483,457	910	27,258	3.3%
Old	509,713	1,420	47,481	3.0%
No higher educ.	874,748	1,848	56,841	3.3%
Higher education	118,422	482	17,898	2.7%
Income < 33 fractile	327,741	554	15,496	3.6%
Income in 33-67 fractile	337,668	628	18,849	3.3%
Income > 67 fractile	327,766	1,148	40,389	2.8%
Wealth < 33 fractile	327,737	287	6,189	4.6%
Wealth in 33-67 fractile	337,669	549	13,155	4.2%
Wealth > 67 fractile	327,736	1,494	55,390	2.7%
No mut. funds	737,782	1,405	38,102	3.7%
Holding mut. funds	255,388	925	36,636	2.5%
Total across each category	993,170	2,330	74,739	3.1%

Table 5. Probit Regresssions: What predicts underdiversification?

The table shows the marginal effects estimated from probit regressions, where the dependent variable is an indicator for i) holding 1 stock only or ii) holding no mutual funds. Marginal effects of the explanatory variables are calculated at their means. Variables are defined in Table 4. The sample consists of all stockholders in 2012. Z-statistics testing for zero marginal effect are reported in parentheses, i.e. the ratio of the marginal effect estimate to the standard errors of the respective predictor. Star-marked coefficients are statistically significant at the 1% level.

	Main results		Robustness: Employee stocks		Informational: Foreigner	
	(1) Holding 1 stock	(2) No mutual funds	(3) Holding 1 stock	(4) No mutual funds	(5) Holding 1 stock	(6) No mutual funds
Male	-0.07* (71.83)	0.03* (-32.49)	-0.07* (75.19)	0.03* (-35.91)	-0.07* (75.08)	0.03* (-35.83)
Old	0.05* (46.51)	-0.13* (-134.20)	0.04* (40.51)	-0.12* (-124.39)	0.04* (40.87)	-0.12* (-123.28)
Higher educ.	-0.05* (-28.69)	-0.08* (-47.89)	-0.05* (-29.51)	-0.07* (-46.51)	-0.05* (-29.60)	-0.07* (-46.71)
Income/100,000	-0.03* (-110.73)	0.00* (16.39)	-0.02* (-102.04)	0.00* (5.08)	-0.02* (-102.02)	0.00* (5.10)
Wealth/100,000	-0.00* (-121.06)	-0.01* (-218.05)	-0.00* (-124.24)	-0.01* (-210.59)	-0.00* (-123.89)	-0.01* (-210.03)
Holding mutual funds	-0.15* (-125.72)		-0.16* (-131.42)		-0.16* (-131.06)	
Employee stock			-0.14* (-65.68)	0.19* (151.15)	-0.14* (-66.01)	0.19* (150.17)
Foreigner					0.04* (9.19)	0.07* (17.10)
Obs.	985,127	985,127	985,127	985,127	982,399	982,399
Pseudo R2	0.07	0.09	0.07	0.10	0.07	0.10

Figure 1. Distribution of investors across number of stocks

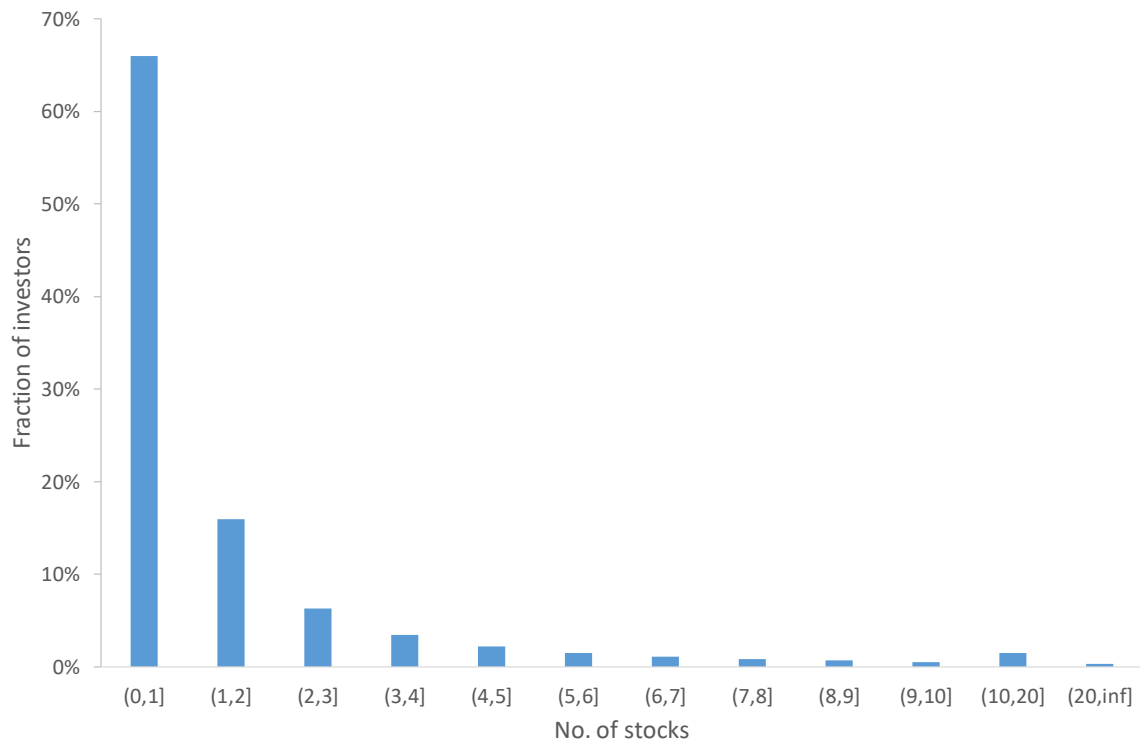The figures show the distribution of the number of stocks investors hold in 2012.

Figure 2. Distribution of the value of investments in stocks across number of stocks

Fig. 2a. The figures show the distribution of the value of stock market investments in DKK in 2012, conditioned on the number of stocks the investors hold.
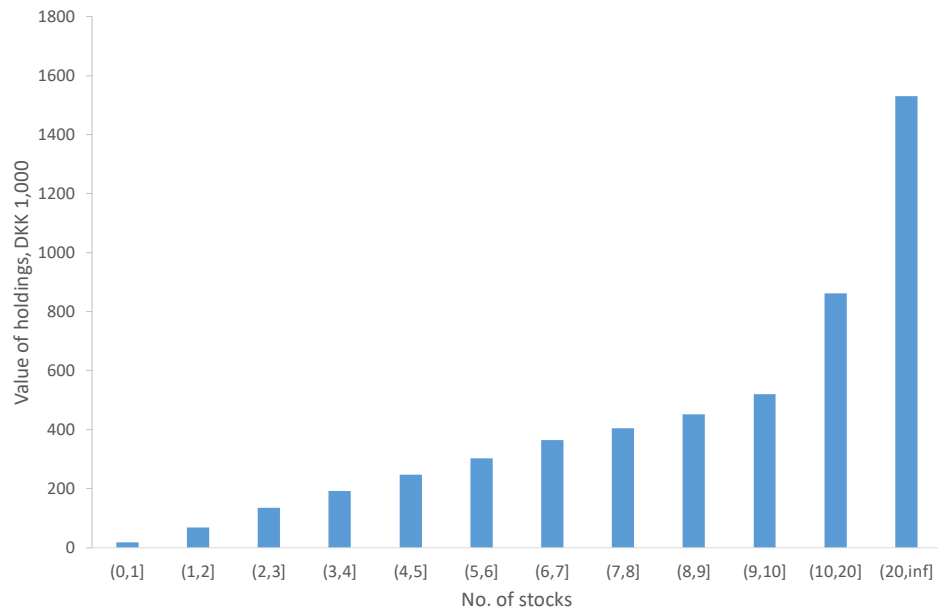

Fig. 2b. The figures show the distribution of the fraction of stock market investments out of total net financial wealth of investors in 2012. Financial wealth is the combined value of cash, bonds, stocks, mutual funds minus total debt.
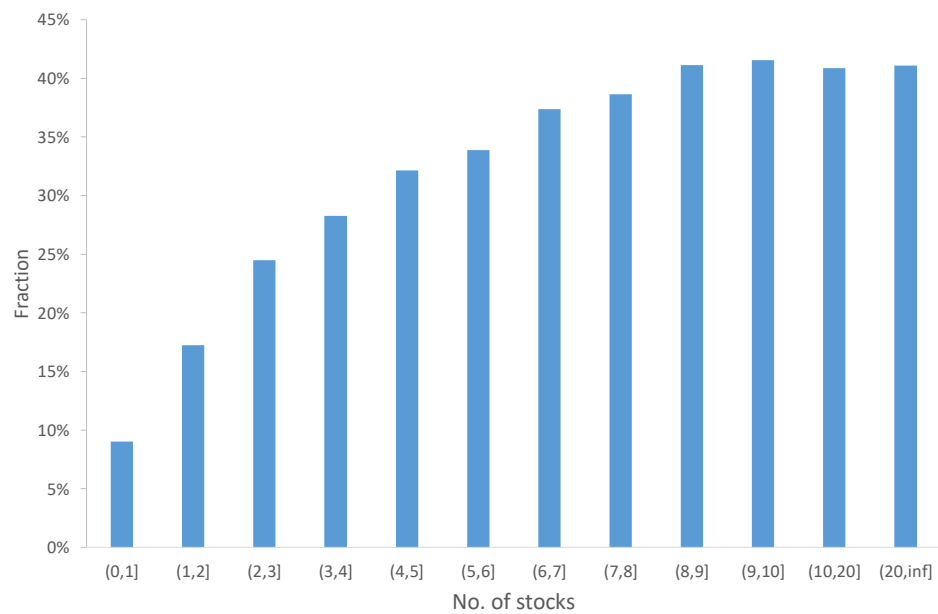
Figure 3. Cumulative returns of 91 Danish stocks. 1997:1-2017:1

The figure shows the development in cumulative returns for 91 Danish stocks.
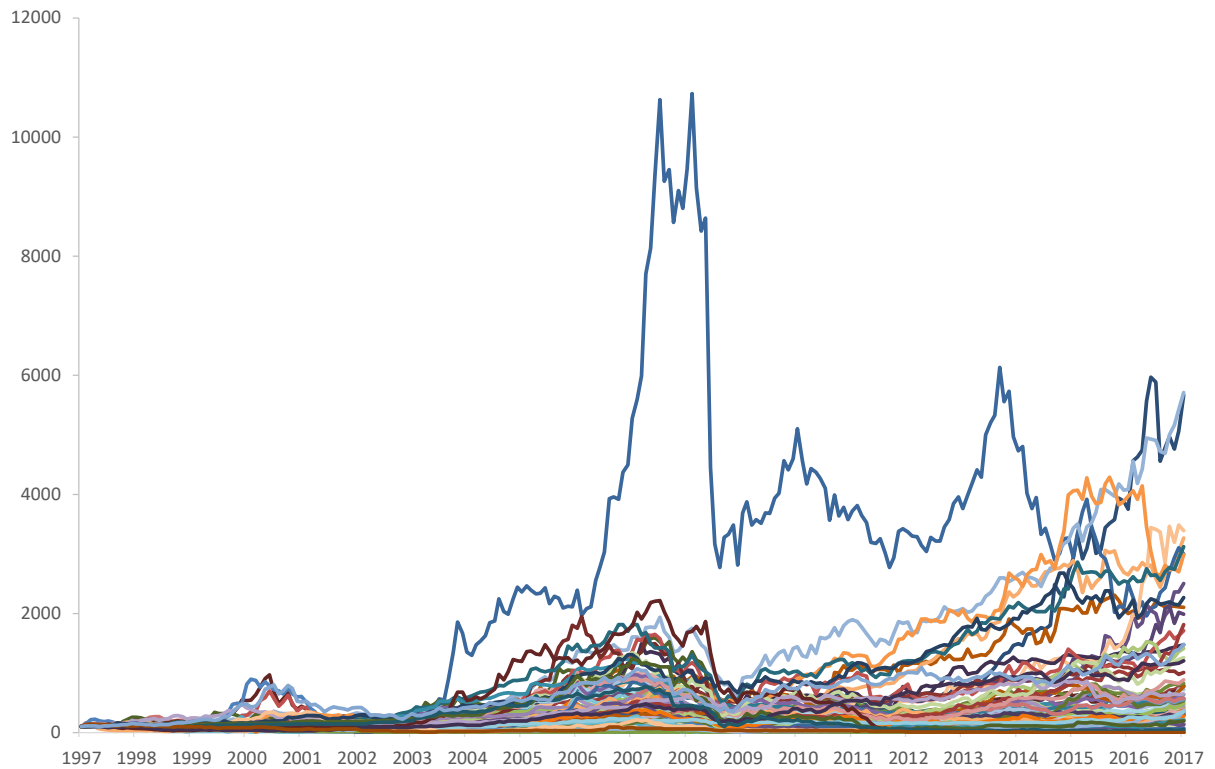
Figure 4. Risk reduction by diversifying

Fig. 4a. The figure shows the level of risk (annualized standard deviation) for portfolios consisting of different numbers of stocks.
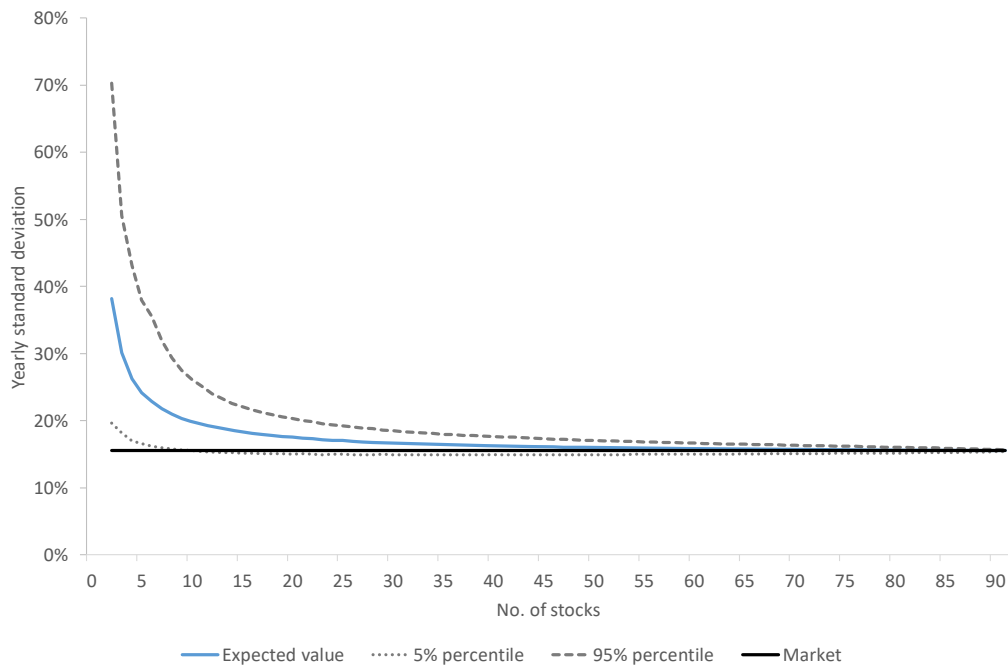
Fig. 4b. The figure shows the level of risk (annualized standard deviation) for portfolios consisting of different numbers of stocks, logarithmic scale for number of stocks (x-axis).
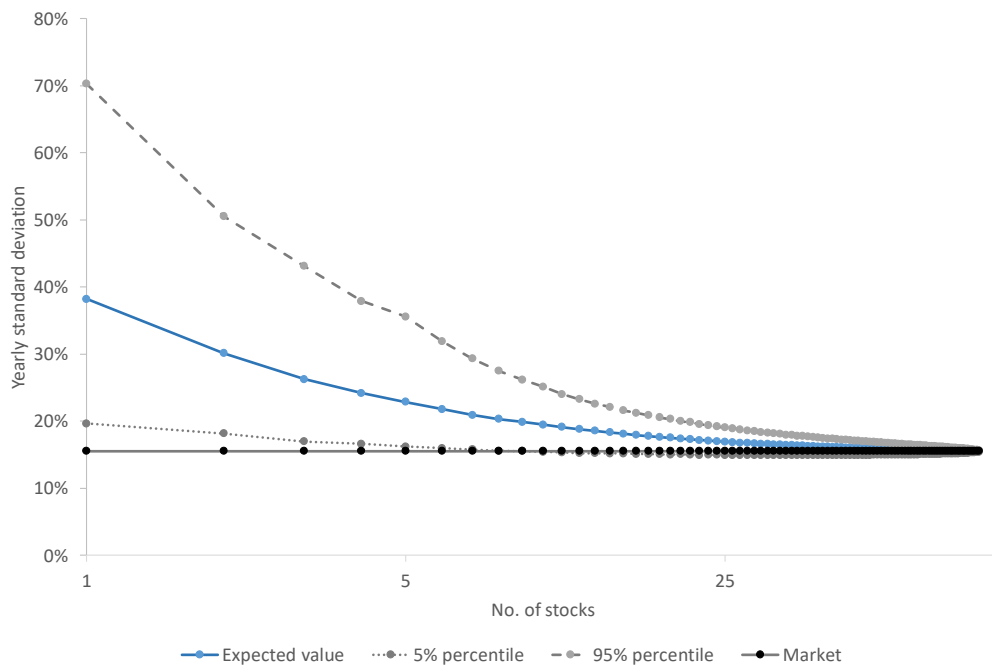
Figure 5. Calculating the loss of underdiversification.

The figure, which plots the efficient frontier and the capital market line, illustrates the foregone return from holding an underdiversified portfolio. The figure is based on the one-stock case. We assume a risk-free rate of 2% and a risk premium of 5%. The standard deviation of the market portfolio is 15.5% and the standard deviation of the one-stock portfolio is 38%.
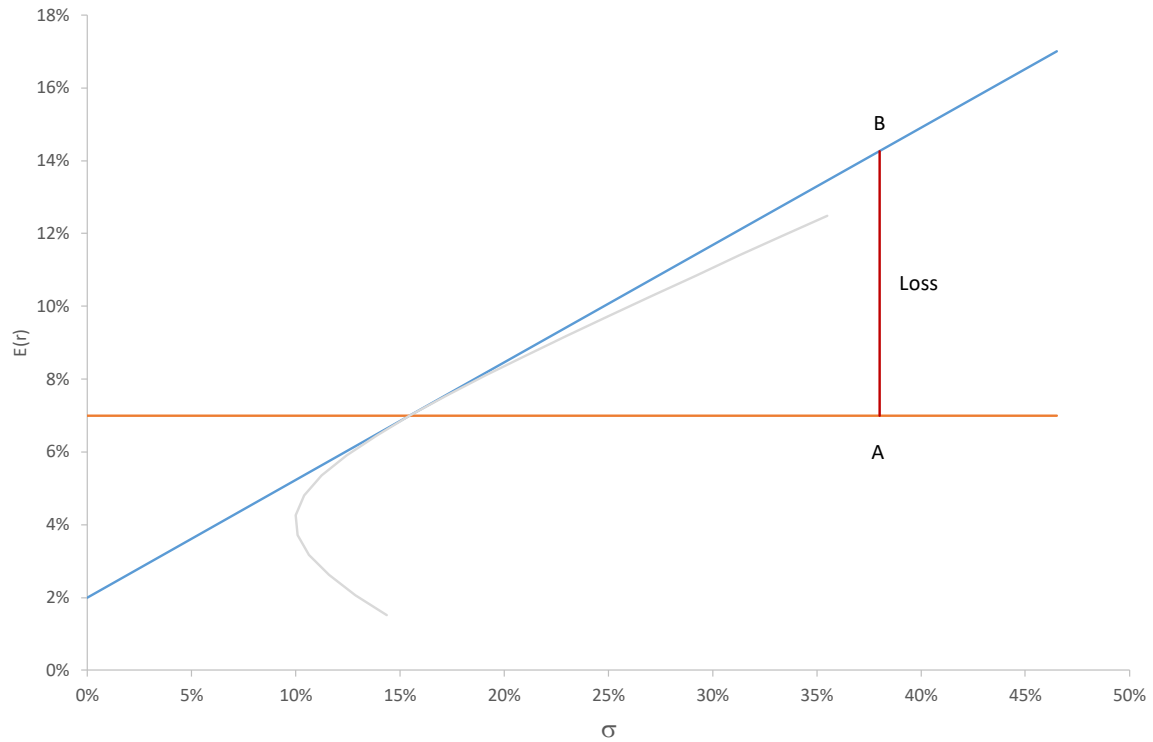
Figure 6. Excess return needed to justify holding *n* random stocks.

Fig. 6a. The figure shows the level of excess return needed to make the investor indifferent between holding *n* randomly selected stocks or holding the market portfolio, for portfolios consisting of different numbers of stocks. Calculations based on a risk-free rate of 2% and a risk premium of 5%.
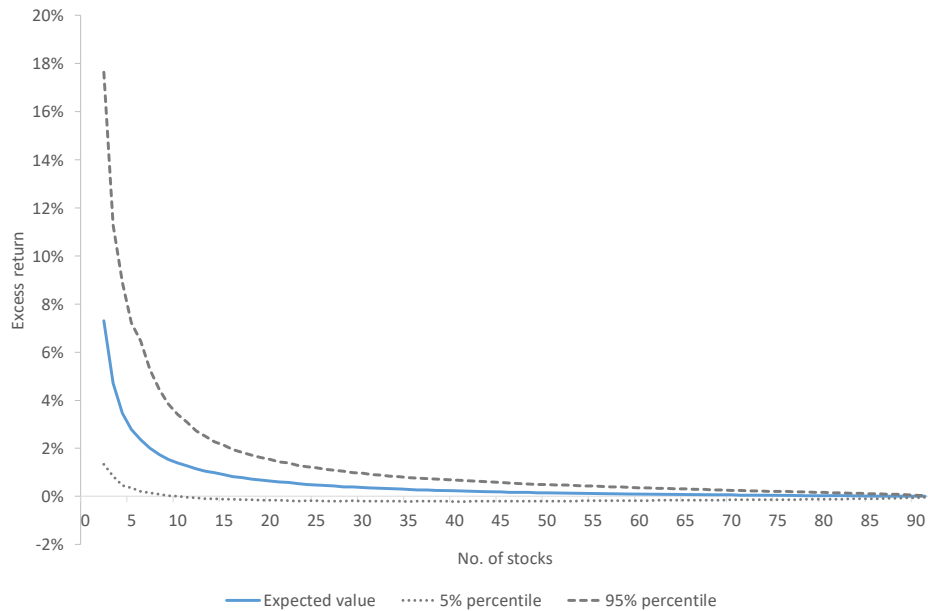

Fig. 6b. The figure shows the level of excess return needed to make the investor indifferent between holding *n* randomly selected stocks or holding the market portfolio, for portfolios consisting of different numbers of stocks, logarithmic scale for number of stocks (x-axis). Calculations based on a risk-free rate of 2% and a risk premium of 5%.
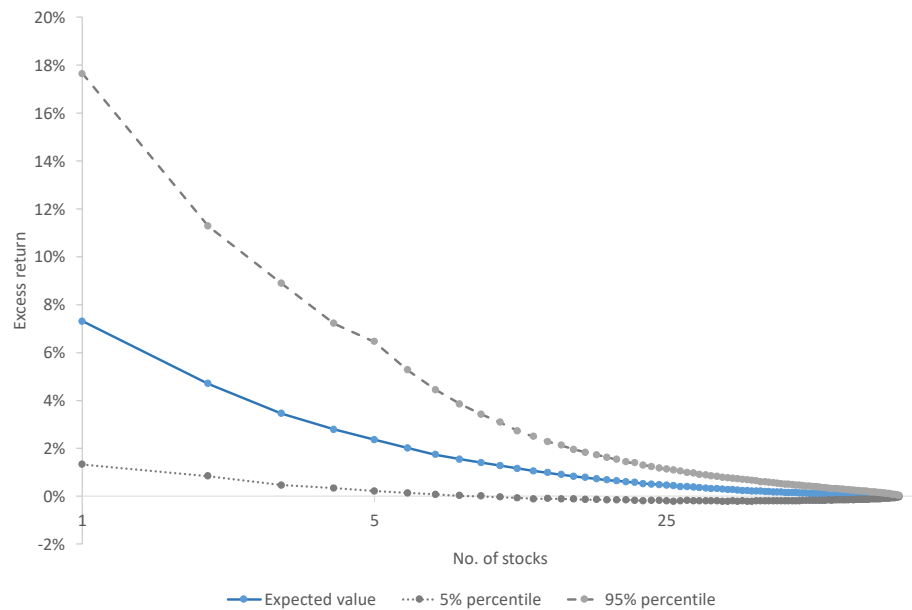
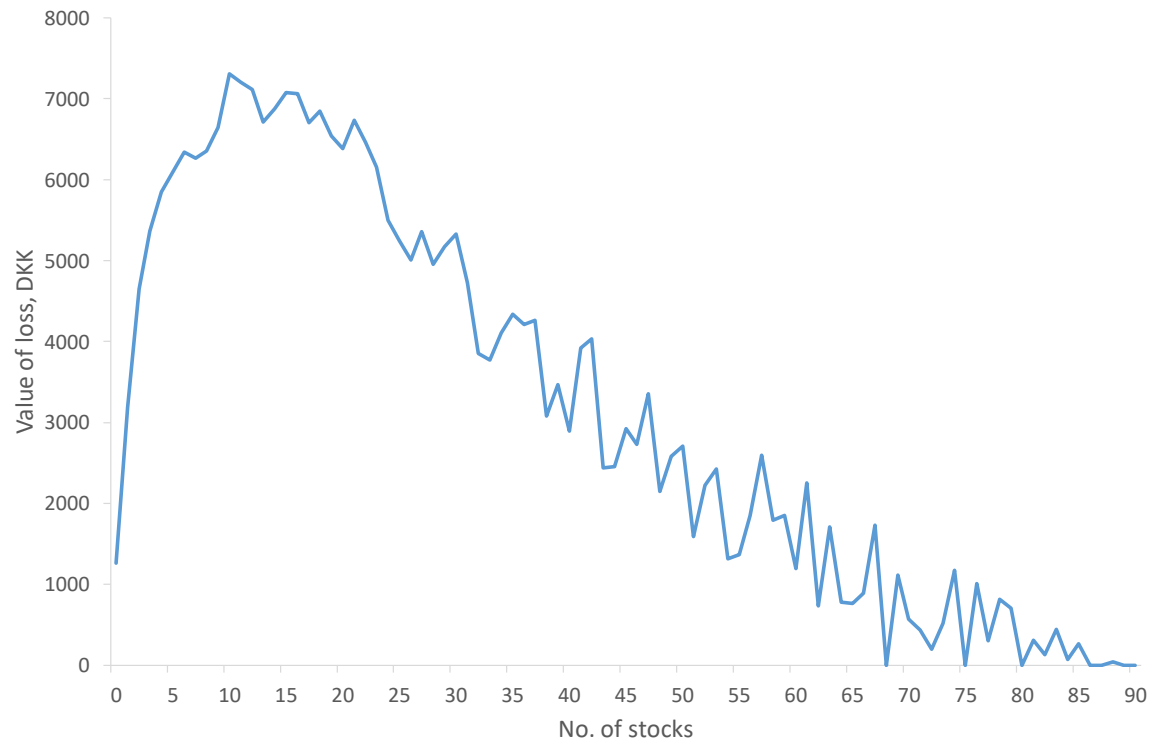**Figure 7. Loss in DKK from underdiversification for portfolios with *n* random stocks**

Figure 8. Aggregate loss in DKK from underdiversification for portfolios with *n* random stocks. DKK millions.
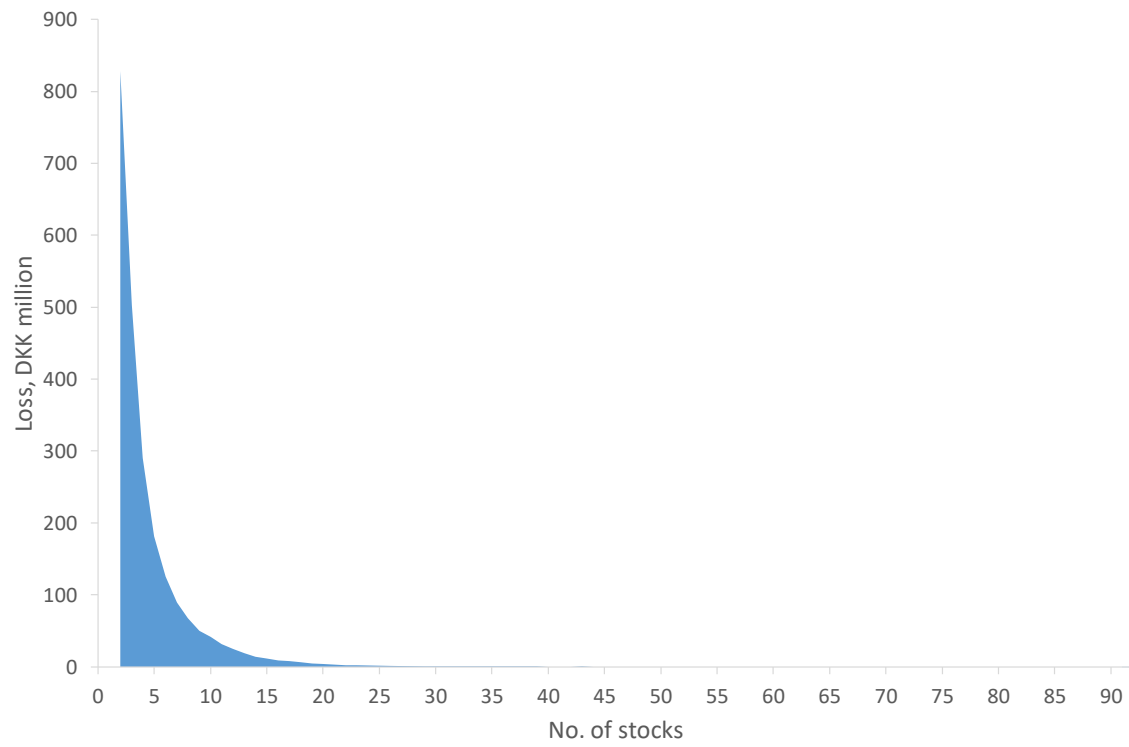
Figure 9. Likelihood of obtaining returns below and above return on the market portfolio

The figure shows the likelihood of obtaining returns below or above the return of the market portfolio. We consider situations where the investor in 1997:4 randomly selects *n* different stocks. We consider situations where the investor chooses one stocks (A), two stocks (B), five stocks (C), ten stocks (D), twenty stocks (E), and fifty stocks (F). We scale portfolio returns by returns on the market portfolio, such that portfolios to the left (right, in blue) of "100" has performed worse (better) than the market portfolio.

A)   One stock

B) Two stocks

C)   Five stock

D) Ten stocks

E)   Twenty stock

F) Fifty stocks

43