Topic 5: Risk and Return

Eddie Shore

Corporate Finance – ECONS4280

July 18, 2022

So far, we've only hinted at the central role of risk. If we have mentioned it, it has been in a hand-wavey fashion when referencing opportunity costs. Now that we have an understanding of some of the basic tools, we can turn to risk and investigate it more thoroughly. In this set of notes, we'll first discuss why risk and return are so intimately linked. We'll then talk about how this relationship bears out when we look at actual financial assets. We'll then talk about how we measure risk. We'll discuss how we can disentangle different sources of risk, and finally, we'll look at a specific model, CAPM, that is used to estimate the riskiness of stocks. With these concepts in place, we can then turn to how we apply them to the problems involved with project selection.

1 Risk and Reward

You've probably heard me say a few times that more risk means more reward. Why is this true? The best way to illustrate this is by recalling that that financial projects (e.g. investment projects, stocks, bonds, etc.) exist in a *marketplace*, and so their returns are subject to *market pressures*. Let's illustrate with an example: suppose we place all projects along two dimensions of risk and return, much in the same way as shown in the picture below:

Here, projects can either be high risk or low risk, and either high return or low return. The 'high risk-high return' and 'low risk-low return' quadrants make sense: this is the maxim behind the 'no pain no gain' slogans we are no doubt familiar with. But suppose you were offered a project that had high risk and low return.
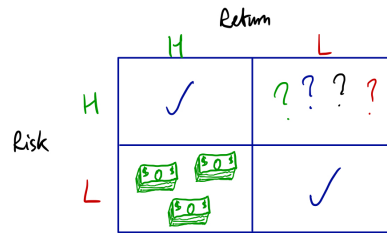
Would you take it? Similarly, suppose there were a project with low risk and high return: would you take that one?

I suspect that your answers would be no and YES respectively. So, what does that mean for the *marketplace* of these projects? Well, if I were the supplier offering the 'high risk-low reward' project, I would see that nobody wants to take me up on my offer. So, *to induce people to buy my project*, I *have* to raise the returns that I offer, *to compensate buyers for the additional risk.* So, the forces of *demand and supply* will push the returns of the project up, moving me into the upper left quadrant.

Similarly, if I am offering a 'low risk-high reward' project, what will happen? Everyone will want it! Punters will flood my office desperate to get their hands on my project! This excess of demand will have the effect of raising the price to buy into my project. In other words, if people now have to pay a lot to buy my project, due to high demand, *then the returns are pushed down.* In other words, the market pushes us into the lower right quadrant. I summarise this idea in the following picture:
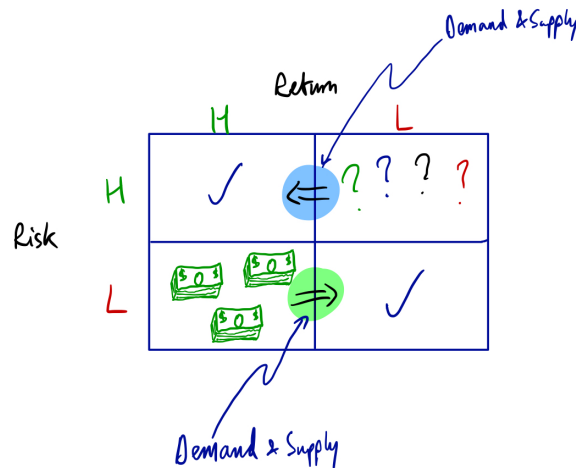
So in well-functioning marketplaces, where buyers and sellers respond quickly and effectively to the forces of supply and demand, projects in the upper right and lower left quadrants ought not to exist, at least in the medium-term. *Financial markets are very fast-moving and efficient marketplaces*, where prices are constantly adjusting and agents are highly motivated to act efficiently. As such, we ought not to see many, if any, projects that are in these quadrants.

1.1   Does this mean it never happens?

No! Sometimes there are projects that turn out to be 'low risk-high return'; you just have to get there first. If you identify something as in that lower left quadrant, you better buy it quick.

That said, hedge funds, banks, and supercomputers are constantly trying to identify these projects, and aim to do so using *vast* quantities of data, both contemporaneous and historical. You should not expect to beat them easily! Maybe you have access to information that the market doesn't have? Trade on that and you could enjoy huge returns! Of course, there is another risk associated with this approach... prison! Yay!!

In general, we should operate with *intense humility* when assessing the risk-return profile of projects. *With near certainty*, the project you're considering is in one of the upper-left or lower-right quadrants. There are many *extremely* smart, and *extremely* informed agents operating in these markets, and you should expect them to do a better job than you at performing these identifications. Thus, always operate under two critical, and inter-related, maxims:

1. If something seems too good to be true, it probably is

2. 'I wouldn't want to be a member of any club that would have someone like me as a member' - Groucho Marx

With respect to this second point, if someone is selling you on a great investment, ask yourself why they're doing that.

1.2   An example: Cryptocurrencies

Cryptocurrencies have been *repeatedly* marketed to *mostly* less well off households as a *near-guaranteed* way to make huge returns.

The narrative around bitcoin is something like, 'This is the future. Economists, bankers, finance researchers, governments, and business leaders are just **squares** who don't get it. You're an idiot if you don't buy bitcoin *right now*.' This is a narrative that strongly indicates that bitcoin is a lower-left quadrant opportunity. The forecasts for where bitcoin is headed have been similarly outlandish:

> **"There are 3 eras of currency: Commodity based, politically based, and now, math based." - Chris Dixon**, Co-founder of Hunch now owned by Ebay, Co-founder of SiteAdvisor now owned by McAfee

> **"Bitcoin will do to banks what email did to the postal industry" - Rick Falkvinge**, Founder of the Swedish pirate party

For a long time, things looked good for the cryptocurrency advocates. Bitcoin in particular seemed to defy the predictions of economists everywhere, lending credence to the view that we were/are 'ignorant'. Here is a graph of the value of Bitcoin from 2013 to the end of 2020, and then from Jan 2021 to June 2022

From a peak of just over $60,000 in September of last year, the value of Bitcoin is now hovering around $20,000, so a **66%** drop. **This is a colossal drop in value**. To give perspective, consider the figure next page showing the drop in the DOW Jones associated with the Wall Street Crash - stocks didn't drop by anything *like* that amount. The crypto crash has not just affected Bitcoin. Below I show the snapshot of the change in price from January to July 2022 of a large number of cryptocurrencies.

5
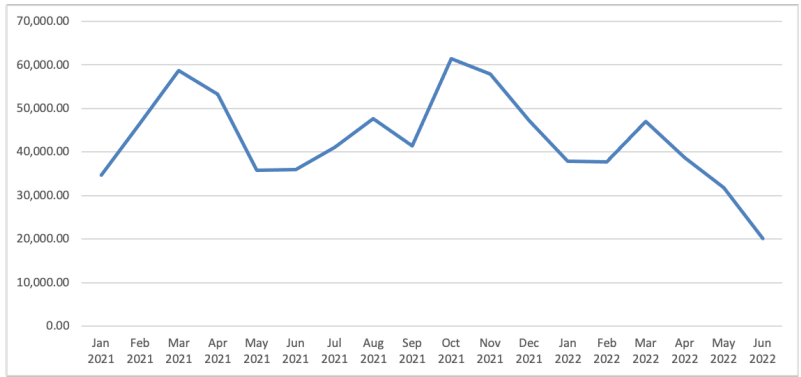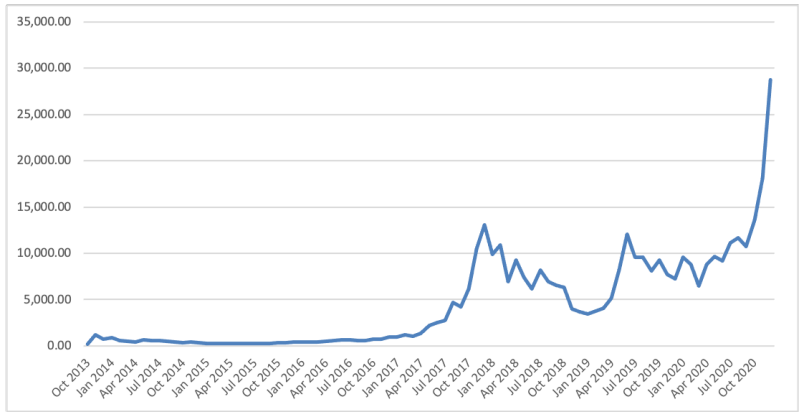
Wall Street Crash on the Dow Jones Industrial Average, 1929
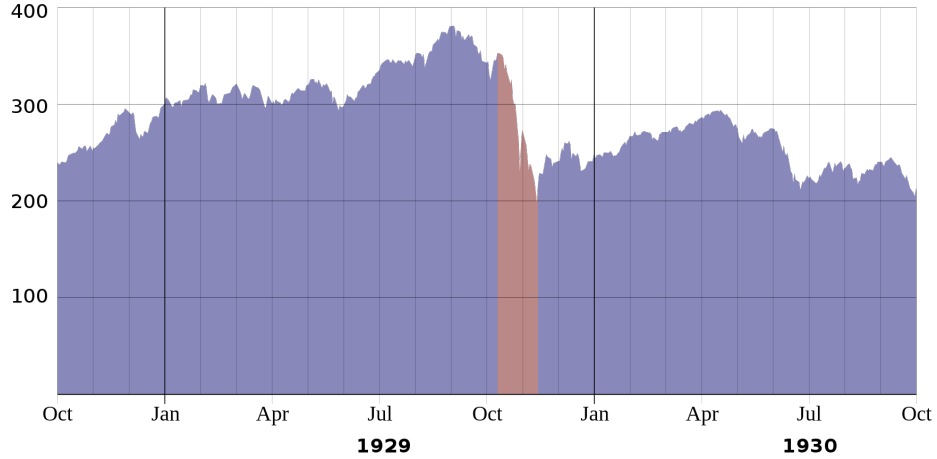
Name	Price	YTD %
☆ ₿ **Bitcoin** 1 BTC	$20,509.80	▼56.98%
☆ ◆ **Ethereum** 2 ETH	$1,141.88	▼69.69%
☆ **BNB** 3 BNB	$231.99	▼55.98%
☆ ✕ **XRP** 4 XRP	$0.3239	▼61.84%
☆ **Cardano** 5 ADA	$0.4468	▼67.52%
☆ **Solana** 6 SOL	$35.35	▼80.18%
☆ **Dogecoin** 7 DOGE	$0.06403	▼62.99%
☆ **Polkadot** 8 DOT	$6.82	▼76.13%
☆ **TRON** 9 TRX	$0.06566	▼14.27%
☆ **Avalanche** 10 AVAX	$18.34	▼83.92%

Terra Classic to USD Chart

| Price | Market Cap | Candle Chart | | 1D | 7D | 1M | 3M | 1Y | YTD | ALL | | LOG |

2 Measuring Risk (i): Single assets

Now that we're comfortable that risk and return ought to be related, how do we measure risk? One way is by the **risk premium**. Another is to look at **past volatility**. Let's consider both.

2.1 Risk Premium

Take two assets:

- US Treasury bill, or *T-bill*.

- S&P 500 Exchange-Traded-Fund (ETF).

The former is an asset we discussed in an early class. Here's a reminder of some of its features:

- **Zero coupon** bond —pays no coupon payments.

- Maturity of one year or less.

- Pays back face value, usually in denominations of $1,000.

- Sells for *discount* on face value.

Note that the T-bill is backed by the US government, *who can always print more money to ensure the bond is paid.* You also know *exactly* what cash flows the bond generates. Whilst there is no such thing as an *entirely* risk free asset, the T-bill is as close as it gets. What about the second asset, the S&P 500 ETF? What are its features?

- An 'ETF' is a financial asset that tracks a specific index —in this case, the S&P 500

- The S&P 500 is an index containing the 500 largest US firms by market capitalisation[1]

---

[1]For those of you who haven't come across this term before, it is the value you get if you multiply the number of shares that exist for a firm by its share price. So, if you had a firm with 1,000 shares outstanding, and a share price of $1,000, the market capitalisation (AKA market 'cap') would be $1,000 \times \$1,000 = \$1,000,000$.

- By buying an 'ETF' you are buying a *share* in this index —stops you from having to buy one of every stock.

- Pays *dividends* from holding the shares—**unknown!!**.

Holding shares in an ETF is a *risky* venture. You don't know how much you will be paid in dividends, and hence you do not know what the cash flows associated with owning this asset will be. Given what we just learned about the relationship between risk and reward, therefore, we should expect that the *return* to holding the S&P 500 ETF *should be greater* than holding the US T-bill. Is that the case?

Suppose we look at the *average annual return* of the S&P 500 and a 3-month T-bill from 1928 to 2021[2]. What do we find?

$$\begin{array}{c|c} \textit{S\&P 500} & 9.98\% \\ \textit{T-bill} & 3.28\% \end{array}$$

So yes, the S&P 500, as a risky venture, *did* pay higher returns than the T-bill. Given that the T-bill is *risk-free*, then the *difference* between these two average annual returns is a measure of the *risk premium* associated with investing in the S&P 500 ETF *over* the risk-free T-bill. So, we can say the *risk premium* for the S&P 500 is roughly **3.70%**.

Definition 1. The **risk premium** of a financial asset is the *difference in average annual returns* of that financial asset and the *risk-free* US T-bill.

Note how the tied relationship of risk and return is *integral* to this definition. In effect, the claim is that where there are returns, there is risk. To show that the S&P 500 was indeed more *risky* than the T-bill, consider the following figure, mapping the average annual return of both assets from 1928 to 2021:

As we can see, the returns are *considerably more volatile*. The riskiness associated with this volatility requires that the S&P 500 have greater returns, otherwise it is not attractive as an investment opportunity.

2.2 Past Volatility

One shortcoming of the 'risk premium' is that it is only an *indirect* measure of risk: we require that risks and returns are related in order to trust its conclusions. If we

---

[2]**Source**: Aswath Damodaran datafiles, Professor at NYU Stern in Corporate Finance.
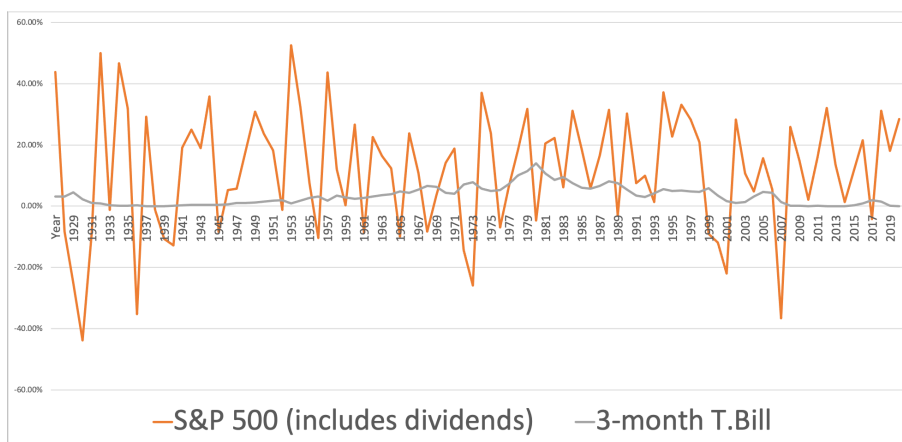
Figure 2: Annual return of S&P 500 and the US T-bill

want to measure the risk of a financial asset more directly, a natural choice is to look at its *volatility* over time—that is to say, *it's variance*. Once we've done that, we can take that direct measure, and see if it does indeed correlate with the returns!

2.2.1 A warning

Okay, full disclosure: **this is a bit challenging mathematically**. I envisage this subsection to be the most difficult so far for those of you less comfortable with math. I'm sorry that it has to be this way, but there really isn't a 'math-lite' alternative that captures the important features of this measure of risk.

2.2.2 Variance: A reminder

What is the variance of a random variable? Conceptually, the variance is a measure of how much a random variable varies. When it varies a lot, the variance is large, and when it varies little, the variance is small. If an object is a *fixed constant* then it doesn't vary at all: its variance is 0.

We can estimate the variance of a random variable by looking at *how far* our observations of that variable are *from the mean, on average*. Putting this notion into mathematics, the variance of random variable X is given by:

$$\sigma^2 = \mathbb{E}[(X_i - \bar{X})^2]$$

Note that taking the square of the deviations means that we are just measuring

distance from the mean; we don't care about whether its above or below. Let's apply this tool to two example financial assets, to show that this measure is a good indicator of the *riskiness* of an asset.

Suppose we have 2 assets, Assets A and B. Suppose that we have 10 observations of the annual returns of those assets:

t	A	B
1	6.7%	2.3%
2	16.6%	2.1%
3	-14.7%	1.9%
4	26.3%	2.4%
5	-7.4%	0.1%
6	-2.3%	2.0%
7	11.3%	-0.1%
8	21.9%	1.1%
9	7.5%	2.2%
1	-3.8%	1.6%
Average	5.61%	1.56%

Which one seems riskier? It seems as though the returns of asset A were far more *variable* than the returns of asset B. Thus, intuitively, asset A is *riskier* than asset B. Can we formalise that? First, let's look at how much each observation deviates from the average:

t	A	Deviation(A)	B	Deviation(B)
1	6.7%	-4.91%	2.3%	0.74%
2	16.6%	10.99%	2.1%	0.54%
3	-14.7%	-20.31%	1.9%	0.34%
4	26.3%	20.69%	2.4%	0.84%
5	-7.4%	-13.01%	0.1%	-1.46%
6	-2.3%	-7.91%	2.0%	0.44%
7	11.3%	5.69%	-0.1%	-1.66%
8	21.9%	16.29%	1.1%	-0.46%
9	7.5%	1.89%	2.2%	0.64%
1	-3.8%	-9.41%	1.6%	0.04%
Average	5.61%	0%	1.56%	0%

Now of course, the *average* deviation from the mean must be equal to 0 by construction. Whilst we can *see* that the deviations are greater in absolute size for

11

A, taking the average of *just* the deviations won't capture that difference, *because the positive and negative deviations cancel each other out.* Now let's square those differences: this way *every deviation measure will be positive*, and so gives a measure of the *distance* from the mean, rather than the *difference* from the mean. This term is called the *squared deviation*:

t	A	Deviation$^2(A)$	B	Deviation$^2(B)$
1	6.7%	24.11%	2.3%	0.55%
2	16.6%	120.78%	2.1%	0.29%
3	-14.7%	412.50%	1.9%	0.12%
4	26.3%	428.08%	2.4%	0.71%
5	-7.4%	169.26%	0.1%	2.13%
6	-2.3%	62.57%	2.0%	0.19%
7	11.3%	32.38%	-0.1%	2.76%
8	21.9%	265.36%	1.1%	0.21%
9	7.5%	3.57%	2.2%	0.41%
1	-3.8%	88.55%	1.6%	0.00%
Average	5.61%	160.7%	1.56%	0.7%

The averages in the columns for Deviation^2 *are the variance of assets A and B.* If we take the *square root* of the variance, we get the *standard deviation*: this is a measure of how much the annual return deviates from the mean *on average*. For asset A, $\sigma^A = 12.68\%$; for asset B, $\sigma^B = 0.86\%$. Both the variance *and* the standard deviation are measures of the *past volatility* of the two assets. Note that if the variance of A is greater than the variance of B then necessarily the standard deviation of A is also greater. This brings us to our 'past volatility' rule:

Definition 2. The **past volatility** rule says that an asset A is *riskier* than an asset B if the *past volatility* of asset A (the *variance/standard deviation* of past returns) is *greater* than the past volatility of asset B.

2.2.3 A brief diversion: linear regression

Okay, so now we have a *direct* measure of riskiness: past volatility. We've mentioned repeatedly that risk and return ought to be directly related. So, if we construct measures of volatility, do they correlate with returns?

To test this possibility, we're going to run a *regression*. A regression is a statistical technique that helps us establish whether or not two things are related. All that

a regression does is *find the* <u>*straight*</u> *line of* **best fit**. So, for example, if we looked at data on rainfall and data on umbrella sales, and plotted them together, a linear regression would find the *line of best fit* between these data points, much as has been done in the following figure:
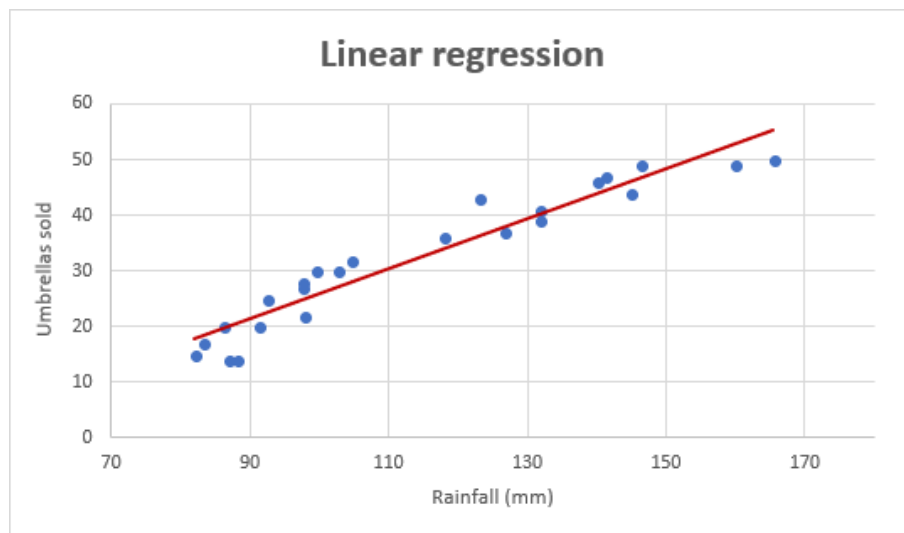
As we'd expect, there seems to be a *positive* relationship between these two variables; when it rains more, people buy more umbrellas. Usually we represent a *regression* not as an image, but as an equation. So in this example, let Y denote umbrellas sold, and X denote rainfall, and then we represent the linear regression as:

$$Y_t = \beta_0 + \beta_1 X_t + \epsilon_t$$

When we *estimate* this expression, we want to estimate, $\{\beta_0, \beta_1\}$: note that this is just the *intercept* and the *slope* of the line of best fit! The key parameter we care about is β_1, as this tells us: are these two things *related*? Note that we allow for some error: very few things in life are deterministic. In a sense, this error is accounting for all the *other* factors that might drive umbrella sales *other* than rainfall. This is why our line of best fit in the umbrella example does not touch every observation —the gap between the line and the observation *is the error*.

Two things matter for any estimate that we construct: its size, and how confident we can be that our estimate of the size is *correct*. The former is referred to as the *coefficient*, and the latter is called the *standard error*. The standard error gives us a *confidence range* that we can reasonably expect the true estimate to be in. Perhaps we estimate a positive coefficient, but the data is so noisy that we can't really be

confident that our estimate is accurate?

To show how this concern works in practice, I'm going to *generate* some data, where I *know* what the relationship between two variables that I generate is. I'll then run a regression on the relationship between those variables, *but vary how noisy the data is*. What we'll find is that when the data is noisy, we might not be able to confidently ascertain what the relationship between the two variables is. The whole process will work something like this:

1. I start by randomly generating 1,000 values between 0 and 1, which I will label X

2. I then randomly generate two sets of *errors*, both of size 1,000. These errors are both drawn from a normal distribution, but the *variance* or *noise* of those errors is different.

   (a) For the first set of errors, ϵ^1, the variance is 0.1.

   (b) For the second set, ϵ^2, the variance is 10.

3. I then *construct* two new variables, Y^1 and Y^2, in the following way:

   (a) $Y_i^1 = 0.5 \times X_i^1 + \epsilon_i^1$

   (b) $Y_i^2 = 0.5 \times X_i^1 + \epsilon_i^2$

4. I then *estimate* two linear regressions, and see whether the computer tells me that there is *statistical significance*. I don't include the actual errors in my regression, as we assume these to be unknown.

   (a) $Y_i^1 = \beta_0^1 + \beta_1^1 X_i^1 + \nu_i^1$

   (b) $Y_i^2 = \beta_0^2 + \beta_1^2 X_i^1 + \nu_i^2$

Before we go ahead and estimate these, let's take a look at what these two different setups look like. On the left hand panel of the Figure below, we have Y^1 plotted alongside X, and on the right hand panel, we have Y^2 plotted alongside X. We can see that the data with less noisy errors gives a very clear, visible positive relationship. By contrast, with highly variable errors as in the right panel, we really can't make out much at all.
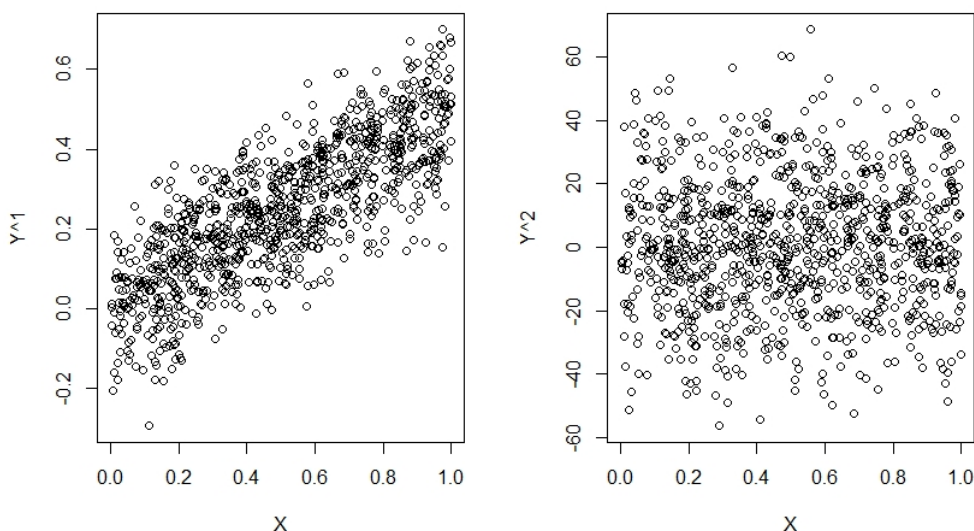
Maybe actually estimating the data will change things? What *should* we find. Well we *know* the true relationship between X and Y, because we constructed it! We should find that $\beta_0 = 0$, and $\beta_1 = 0.5$. What *do* we find?

First, let's look at the coefficients in the regression table below: those are the values to the right of (Intercept) (β_0), and X (β_1) that are *not* in brackets. When the variance of the shocks is low, i.e. in the Y^1 case, we pick up the *true* coefficients pretty closely! We estimate that $\beta_0^1 = 0.0026$, and $\beta_1^1 = 0.4895$. When the variance is high, i.e. in the Y^2 case, we are way off: $\beta_0^2 = -0.1862$ and $\beta_1^2 = 2.164$.

What about the *standard errors*? The standard errors are the values in brackets immediately below our coefficients. **The rule of thumb is that you can be 95% confident that the true value is between 1.96 standard errors of your estimated coefficient**. So, for β_0^1, the range is $[-0.0107, 0.0159]$, and for β_1^1, the range is $[0.4664, 0.5126]$. That is extremely tight!! It also contains the true value!

For β_0^2, the range is $[-2.8126, 2.4402]$, and for β_1^2, the range is $[-2.4087, 6.7367]$. These ranges are so large that we really can't be very confident in our estimate at all. So in the case with smaller errors, *we can be considerably more precise* about what seems to be a reasonable range for our estimates. Note that there is *at least* a 95% chance in the first case that the *true* β_1 is positive. In the case with larger errors, we can't even say that!

The final comment to make on this regression table is about the asterisks. As-

15

Dependent Variables:	Y^1	Y^2
Model:	(1)	(2)
Variables		
(Intercept)	0.0026	-0.1862
	(0.0068)	(1.340)
b	0.4895***	2.164
	(0.0118)	(2.333)
Fit statistics		
Observations	1,000	1,000
R^2	0.63304	0.00086
Adjusted R^2	0.63267	-0.00014

IID standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

terisks denote *degrees of statistical significance.* All that an asterisk tells you is that, with some probability, the estimate you've found *is not zero*: i.e. the relationship you've identified is *meaningful.* For this table, one star denotes that there is at most a 10% chance that the true relationship is actually 0, despite your estimated coefficient. Two stars denotes that there is at most a 5% chance, and three stars denotes that there is at most only a 1% chance that the true relationship is zero. Note that we have three stars for our estimate of β_1^1: what is this telling us. Simply that there is a less than 1% chance that the *positive* relationship we've estimated is in fact just noise.

How do we do this in practice? One way is to do it via Excel. A downside of using Excel is that it's difficult to pull out the *standard errors* from excel, which makes it hard to assess how confident we should feel in the estimated slope. This is where the more sophisticated software packages/languages like Stata, R, Matlab, Python, or Julia come into their own.

If you want to learn more about how these tools and techniques work, then you should consider taking an econometrics or statistics class. If you're going to work in finance, you will no doubt come across regressions at some point, so it may be worth your while to learn how they work. However, **this is not a statistics class**, so I will not be requiring you to perform complicated regressions. At most, you'll have to fit a line of best fit using Excel, which is really easy.

2.2.4   Does volatility correlate with returns?

Okay, back to our original question: does our measure of risk correlate with returns? To answer this question, I've downloaded data on the Annual Returns of 25,249 stocks, from 2008 to 2016, using a database called the Center for Research in Security Prices or CRSP. For each stock, I estimate its volatility using its past returns. I then run a regression of annual returns (RET) on volatility (VOL):

$$RET_{i,t} = \beta_0 + \beta_1 VOL_{i,t} + \epsilon_{i,t}$$

Note that our theory tells us that β_1 *should be greater than 0*: i.e. the greater the volatility, the greater the return. What do we find?

Dependent Variable:	*RET*
Model:	(1)
Variables	
(Intercept)	-0.0338***
	(0.0024)
VOL	0.3127***
	(0.0028)
Fit statistics	
Observations	223,659
R^2	0.05414
Adjusted R^2	0.05414

IID standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Well now, isn't that a relief! The coefficient we find is *positive* and equal to 0.3127. The standard error is *very* small: we can say with 95% confidence that the *true* value of β_1 is somewhere between $[0.3072, 0.3182]$. That is very precise!!! Similarly, we can be almost certain that the true relationship is positive —the estimated probability that the true relationship is actually zero is numerically *equal* to 0.

3 Measuring Risk: (ii) From stocks to portfolios

In the previous section, we only considered how risky a *single* asset was. It is very rare, *and strongly discouraged* to only invest in a single asset. Most of the time we hold a *portfolio* of stocks. In this subsection, we'll talk about how we measure the riskiness of a *portfolio* of stocks.

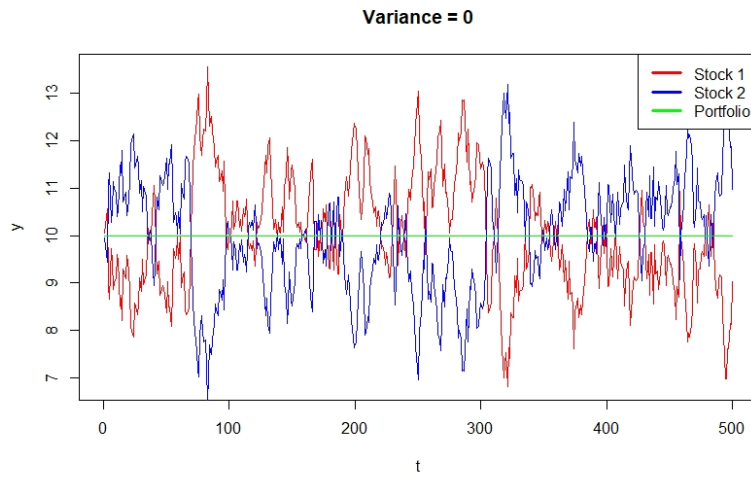## 3.1 The importance of covariance

Can we just add up the variances of all the stocks in the portfolio? Will that work? I'm going to show you three simulated examples of two-stock portfolios to demonstrate why this won't work. In the first example, the shocks that affect the two stocks are *perfectly inversely related*. In the second example, the shocks are *independent* across the two stocks. And in the third, the shocks are *identical* across the two stocks. Let the value of a given stock be given by y_t^i, and let $y_t^i = 1 + \rho y_{t-1}^i + \epsilon_t^i$, where $i \in \{1, 2\}$. Let $\{\epsilon^1, \epsilon^2\}$ denote the shocks for the two stocks in each of the three examples, *and suppose they have identical variance*. Then our three cases can be summed up as:

1. $\epsilon_1^1 = -\epsilon_1^2$

2. $\epsilon_1^1 \perp\!\!\!\perp \epsilon_1^2$

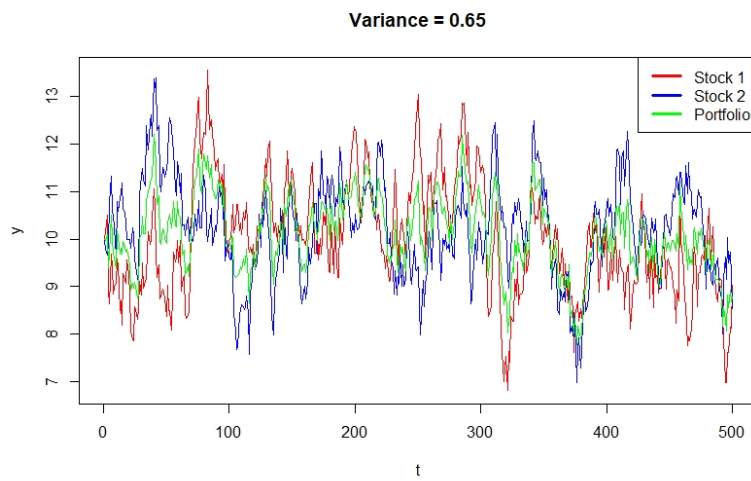3. $\epsilon_1^1 = \epsilon_1^2$

Let's simulate these now, supposing that that in each case, the portfolio is a 50/50 blend of the two stocks.

What do we find? When the shocks exactly offset one another, as in Case 1, *the variance of the portfolio is 0.* There is no risk *at all*: the portfolio will *always* be worth 10. In Case 2, the stocks occasionally offset each other, but not perfectly: the variance is 0.65. When the two stocks effectively move identically, the portfolio variance is *maximised* as 1.44.

What does this tell us? It tells us that the *covariance* between the stocks is *critical* in determining the overall variance of the portfolio! So we cannot simply the sum the variances of the two stocks and expect to come up with a sensible measure of the volatility of our portfolio: in all three of the cases above, the *sum* of the variances of the shocks is the **same**.

Variance = 0

(a) Case 1: $\epsilon_1^1 = -\epsilon_1^2$

Variance = 0.65

(b) Case 2: $\epsilon_1^1 \perp\!\!\!\perp \epsilon_1^2$

Variance = 1.44

(c) Case 3: $\epsilon_1^1 = \epsilon_1^2$

19
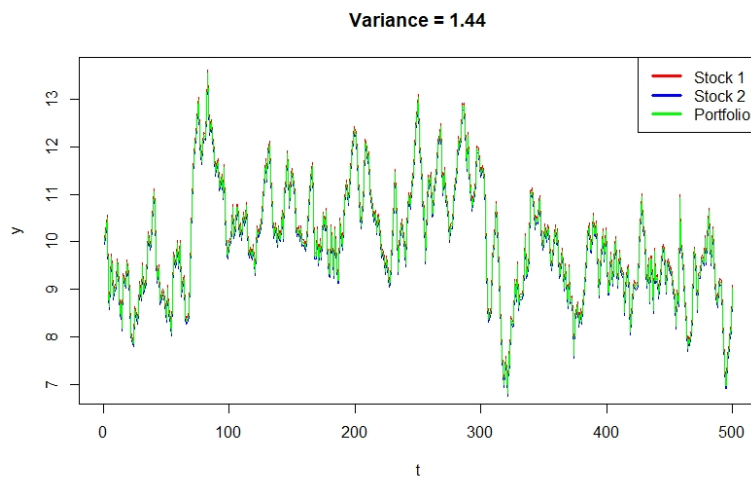
3.2 Formalising the argument —two-stock portfolios

Suppose we have a two-stock portfolio, P, consisting of two stocks, $\{x, y\}$, in proportions $\{\omega_x, \omega_y\}$. Then:

$$P = \omega_x x + \omega_y y$$

Then the returns of the portfolio will just be the weighted sum of the two returns:

$$r_P = \omega_x r_x + \omega_y r_y$$

What about the variance of the portfolio? As we've just seen:

$$\sigma_P^2 \neq \omega_x \sigma_x^2 + \omega_y \sigma_y^y$$

What is the formula for variance?

$$\sigma^2 = \mathbb{E}[(x - \mathbb{E}[x])^2]$$

Let's apply this formula to our portfolio:

$$
\begin{aligned}
\sigma_P^2 &= \mathbb{E}[(r_P - \mathbb{E}[r_P])^2] \\
&= \mathbb{E}[(\omega_x r_x + \omega_y r_y - \mathbb{E}[\omega_x r_x + \omega_y r_y])^2] \\
&= \mathbb{E}\left[\omega_x^2 (r_x - \mathbb{E}[r_x])^2 + \omega_y^2 \mathbb{E}[r_y - \mathbb{E}[r_y]]^2 + 2\omega_x \omega_y (r_x - \mathbb{E}[r_x])(r_y - \mathbb{E}[r_y])\right] \\
&= \omega_x^2 \sigma_x^2 + \omega_y^2 \sigma_y^2 + 2\omega_x \omega_y \sigma_{xy}
\end{aligned}
$$

So, *key* to this formula is the third term: the *covariance between x and y*! When x and y move in *opposite* directions, then σ_{xy} will be *negative*, and so the portfolio variance *falls*. In the extreme case where they are perfect mirrors, then the variance collapses to zero, as in our case 1 from before. If the two stocks are *unrelated* then $\sigma_{xy} = 0$, and so the *weighted sum of the variances* will give us the answer!

3.2.1 Application to simple example

What the formula above tells us is that it is not *just* the volatility of the *individual* stocks in our portfolio that matters, but *also* their *covariance*. To really hammer this point home, consider the following example. Suppose we are choosing between two portfolios:

- A:

  - Stocks: $\{x, y\}$
  - Weights: $\omega_x = \omega_y = \frac{1}{2}$
  - $\sigma_x^2 = 5$, $\sigma_y^2 = 2$, $\sigma_{xy} = -3$

- B:

  - Stocks: $\{c, d\}$
  - Weights: $\omega_c = \omega_d = \frac{1}{2}$
  - $\sigma_c^2 = 2$, $\sigma_d^2 = 1$, $\sigma_{cd} = 1$

Note that *both* of the stocks in Portfolio A are more volatile than *either* stock in Portfolio B. What is the expected variance of the two Portfolios? Well, we just apply the formula from above!

$$
\begin{aligned}
\sigma_A^2 &= \omega_x^2 \sigma_x^2 + \omega_y^2 \sigma_y^2 + 2\omega_x \omega_y \sigma_{xy} \\
&= (0.5)^2 5 + (0.5)^2 2 - 2(0.5)(0.5)3 \\
&= 0.25
\end{aligned}
$$

$$
\begin{aligned}
\sigma_B^2 &= \omega_c^2 \sigma_c^2 + \omega_d^2 \sigma_d^2 + 2\omega_c \omega_d \sigma_{cd} \\
&= (0.5)^2 2 + (0.5)^2 1 + 2(0.5)(0.5)1 \\
&= 1
\end{aligned}
$$

So Portfolio B has *four times* the variance of Portfolio A!! Extraordinary!

3.3 Generalisation to 'n'-size portfolios

This is advanced. Full warning. Suppose that instead of having just two stocks in your portfolio, we generalise to 'n' stocks. What is the variance of the portfolio in that case? Although this sounds like it would be super hard to calculate, there's actually a simple way of outlining and calculating this number. First, note that the 'n' dimension portfolio can be expressed in the following way, where there are 'n' weights that communicate how much of our portfolio is in each of the 'n' stocks:

$$P = \omega_1 x_1 + \omega_2 x_2 + ... + \omega_n x_n$$

The return on this portfolio is just:

$$r_P = \omega_1 r_1 + \omega_2 r_2 + ... + \omega_n r_n$$

What about the variance of the portfolios returns? These can be calculated as before... but the algebra is super messy!! I'll go through the stages step by step below, but don't worry if you feel intimidated by this derivation: it is not easy!!

$$
\begin{aligned}
\sigma_P^2 &= \mathbb{E}[(r_P - \mathbb{E}[r_P])^2] \\
&= \mathbb{E}\left[(\omega_1 r_1 + \omega_2 r_2 + ... + \omega_n r_n - \mathbb{E}[\omega_1 r_1 + \omega_2 r_2 + ... + \omega_n r_n])^2\right] \\
&= \mathbb{E}\left[(\omega_1(r_1 - \mathbb{E}[r_1]) + \omega_2(r_2 - \mathbb{E}[r_2]) + ... + \omega_n(r_n - \mathbb{E}[r_n]))^2\right] \\
&= \mathbb{E}[\underbrace{\left(\omega_1^2(r_1 - \mathbb{E}[r_1])^2 + \omega_2^2(r_2 - \mathbb{E}[r_2])^2 + ... + \omega_n^2(r_n - \mathbb{E}[r_n])^2\right)^2}_{\text{Own-product}} + \\
&\qquad \underbrace{2\omega_1\omega_2(r_1 - \mathbb{E}[r_1])(r_2 - \mathbb{E}[r_2]) + ... + 2\omega_n\omega_{n-1}(r_n - \mathbb{E}[r_n])(r_{n-1} - \mathbb{E}[r_{n-1}])]}_{\text{Cross-product}} \\
&= \sum_{i=1}^{n} \omega_i^2 \sigma_i^2 + \sum_{i=1}^{n}\sum_{j \neq i} \omega_i \omega_j \sigma_{ij} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \omega_i \omega_j \sigma_{ij}
\end{aligned}
$$

Okay... so what the hell does that mean? It is very hard to think of how to actually calculate this portfolio variance using a calculator. *However*, there is a very *neat* way of representing our data that will allow us to find out this number very easily!

3.3.1 Covariance and Weight Matrices

A matrix is both a classic action movie from 1999 and a very convenient mathematical representation of data. All a matrix is is a group of numbers, represented in a grid. Matrices are very useful, because they allow us to represent lots of data in a more compact fashion. There are also useful techniques we can apply to matrices to make our lives easier. We aren't going to worry much about the latter advantage (save that for your linear algebra classes!), but we absolutely will take advantage of the former.

We're going to construct a covariance matrix and a weighting matrix because it will make it *much* easier to calculate the *portfolio* variance. So let's take a look at how we construct both.

A covariance matrix is a *classic* matrix. It's a way of grouping together all the variances *and* covariances between some set of random variables. The idea is that each 'row-column' entry gives the corresponding covariance between the variable in the row and the variable in the column. Along the diagonal, we are comparing row 'n' with column 'n', so this diagonal just gives us the *variance*. A covariance matrix is often represented as Σ, which is the upper-case of σ. A covariance matrix for the random variables, $\{x_1, x_2, ..., x_n\}$ will look something like this:

$$\Sigma = \begin{array}{c} \\ x_1 \\ x_2 \\ \vdots \\ x_n \end{array} \begin{array}{ccccc} x_1 & x_2 & \dots & x_n \\ \sigma_1^2 & \sigma_1\sigma_2 & \dots & \sigma_1\sigma_n \\ \sigma_1\sigma_2 & \sigma_2^2 & \dots & \sigma_2\sigma_n \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_1\sigma_n & \sigma_2\sigma_n & \dots & \sigma_n^2 \end{array}$$

So we can see that all the *diagonals* correspond to the *variances* of the random variables, and all the *off-diagonals* correspond to the *covariances*. If all the variables are *independent* of one another, what should the off-diagonals look like?

$$\Sigma_{independent} = \begin{array}{c} \\ x_1 \\ x_2 \\ \vdots \\ x_n \end{array} \begin{array}{ccccc} x_1 & x_2 & \dots & x_n \\ \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{array}$$

Note that $\sigma_i \sigma_j \equiv \sigma_j \sigma_1$, so the upper 'triangle' is just the exact same as the lower 'triangle', just organised differently.

The *weighting* matrix is going to capture something very similar. We're going to lay out all the possible combinations of weighting combinations. Weight matrices are typically denoted by Ω, which is the upper-case of ω. If we do this for our weights, it will look something like this:

$$
\Omega = \begin{array}{c} \\ x_1 \\ x_2 \\ \vdots \\ x_n \end{array}
\begin{array}{ccccc}
x_1 & x_2 & \ldots & x_n \\
\omega_1^2 & \omega_1\omega_2 & \ldots & \omega_1\omega_n \\
\omega_1\omega_2 & \omega_2^2 & \ldots & \omega_2\omega_n \\
\vdots & \vdots & \ddots & \vdots \\
\omega_1\omega_n & \omega_2\omega_n & \ldots & \omega_n^2
\end{array}
$$

Okay, so we have two matrices now. If we take *each element* in *each matrix*, and multiply them together, *element-by-element*[3], we'll have a matrix we can call $\Sigma.\Omega$. It'll have a matrix that looks like this:

$$
\Sigma.\Omega = \begin{array}{c} \\ x_1 \\ x_2 \\ \vdots \\ x_n \end{array}
\begin{array}{cccc}
x_1 & x_2 & \ldots & x_n \\
\omega_1^2\sigma_1^2 & \omega_1\omega_2\sigma_1\sigma_2 & \ldots & \omega_1\omega_n\sigma_1\sigma_n \\
\omega_1\omega_2\sigma_1\sigma_2 & \omega_2^2\sigma_2^2 & \ldots & \omega_2\omega_n\sigma_2\sigma_n \\
\vdots & \vdots & \ddots & \vdots \\
\omega_1\omega_n\sigma_1\sigma_n & \omega_2\omega_n\sigma_2\sigma_n & \ldots & \omega_n^2\sigma_n^2
\end{array}
$$

Note that the definition of the portfolio variance was:

$$
\sigma_P^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \omega_i\omega_j\sigma_{ij}
$$

Note that this is just *the sum of all the entries in the matrix above!* So, if we want to calculate the variance of a portfolio, we just need to follow four steps:

1. Calculate the covariance matrix[4], Σ.

2. Calculate the weight matrix[5], Ω.

---

[3]This is called 'element-wise' multiplication. There are other ways you can multiply matrices together, but don't worry about that!

[4]If you're using a software package/program like R, or Stata, this is one line of code: super easy once you know how!

[5]This is also easy if you have a list of all the weights, but requires some knowledge of matrix

3. Multiply *element-wise* the matrices Σ and Ω.

4. Take the sum of all the entries.

If you gain some familiarity with matrix algebra, and you learn a little coding, this exercise is extremely quick to perform.

3.3.2   Application: Using CRSP Data

Now that we have the tools to calculate the variance of a portfolio for any number of stocks, 'n', a natural question emerges: what effect does increasing 'n' have on the portfolio variance?

To answer this question, I took the same sample of stocks from CRSP that I used to estimate the relationship between volatility and returns above. Now, I randomly select a certain number of stocks from that sample, and compute the volatility of my stock. I start by picking just 1 stock at random, and then compute the portfolio variance. I do this many times (10,000), then take the average portfolio variance across all 10,000 of these trials. *This gives me a measure of how volatile on average a stock portfolio containing just one stock is.* I then do the same for a portfolio with 2 stocks, then with 3, so on, all the way up to a stock portfolio containing 40 stocks. What do I find? So, the *average* variance of my portfolio is *strictly decreasing* in the number of stocks that I pick, *if I pick at random*. What this is telling us is that *the more stocks I have*, the lower my *portfolio risk is*.

---

algebra. If you let Ω denote the weighting matrix, and ω is the list or *vector* of weights, then $\Omega = \omega'\omega$. This will mean *nothing* to you if you're not familiar with matrix algebra, but if you are, is a very simple operation that you can easily plug into a computer
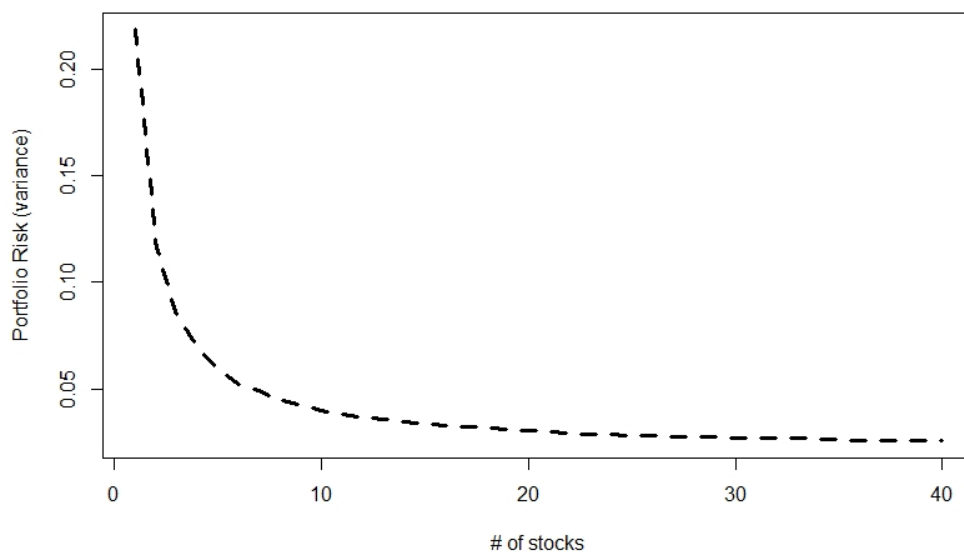
Figure 4: Portfolio Risk is falling in the size of the portfolio