Your submission should consist of:  
One-page Summary Sheet,  
Two-page memo,  
Your solution of no more than 20 pages, for a maximum of 23 pages with your summary  
and memo.  
Note: Reference list and any appendices do not count toward the 23-page limit and  
should appear after your completed solution.

2016 ICM Problem D: (Measuring the Evolution and Influence in Society’s Information Networks, Are we heading towards a thirsty planet, Modeling Refugee Immigration Policies)

<https://www.dropbox.com/s/fl83mw7molzdu6y/2016i_examiners_report.pdf?dl=0>

2015 Question D (Is it sustainable development)

<https://www.dropbox.com/s/h2xr7za9np993gv/2015_C_examiners_report.pdf?dl=0>

2009 ICM (Creating Food Systems: Re-Balancing Human-Influenced Ecosystems)

<https://www.dropbox.com/s/gezrr6de0b400tn/2009i_examiners_report.pdf?dl=0>

2007 ICM (Organ Transplant: The Kidney Exchange Problem)

<https://www.dropbox.com/s/tyuq1lqfz96prcm/2007i_examiners_report.pdf?dl=0>

2006 ICM (Trade-offs in the fight against HIV/AIDS)

<https://www.dropbox.com/s/96orqazm0xhvc7u/2006i_examiners_report.pdf?dl=0>

2014 ICM Triage guidelines

<https://www.dropbox.com/s/6awysoxj9ecu3wa/mcm2014C_Triage_Guidelines.pdf?dl=0>

Triage Guidelines rewritten for this question:

1. Data extraction (applicable to this question?)
2. Do they develop influence measures to determine the price of data
3. Is there some for of visualisation of their work?
4. Do they at least discuss (something more advanced but necessary)
5. Do they answer the most important questions:
   1. WHAT ARE THE IMPORTANT QUESTIONS
6. Do they discuss the measurement of (something) of (factors) on price?
7. “Algorithm” (some difficult questions)
8. Do they discuss the science, understand and utility of the price mechanism?
9. Do they use their models to help decide on the policies?
10. Page limit
11. Do they discuss strengths, weaknesses, and sensitivity?

From the MCM guidelines:

1. **What is meant by “should” - what “should” the price of data be?**
2. Price of data shouldn’t be dependent on only one metric (duh)
3. Must make clear of the assumption in building of your model
4. Must consider the three different domains of data
5. Addressing issues over time (for higher level)
6. Price of data should be defensible
7. Your memo should be readable by the layman
8. References have to be well presented, figures and text cited
9. Need real data somewhere
10. Different types of data, individual

# 0 Summary

(Make the distinction between privacy and data)

We have modelled the individual valuation of personal data. We have confirmed the discrepancy between the valuations of personal data and the individual valuation of privacy. We have some data points to confirm the relationship of the curve. From this we determine what is the fair price.

Assumption we take in the building of the model include. (Describe how we handle quality of information.)

We illustrated the market failure due to imperfect information and inequity. We have also grounded our models in reality - there points are made up of real data points.

(Describe the strength, weaknesses and the results of the sensitivity analysis.)

We suggest three different policy decisions, and reflect its impact on our model. It includes strengthening existing legislation on data protection, increasing the sense of awareness through education campaigns, as well as the promotion and protection of a cryptographic data framework. With these, it is possible that our society can migrate that operates on a framework that is resistant to privacy violations, as individuals can take advantage of their data without disclosing the data to any other party. We discuss the impact of our proposals with respect to our model.

Some concluding sentences.

The Woke Framework

Team 87374

Table of Contents

[**0 Summary**](#_ww3nqmf17akq) **3**

[**1 Introduction**](#_9evcxm5nuv4h) **7**

[1.1 Should There Even Be a Price on Private Information?](#_n5zj3n78whth) 7

[1.1.1 Current Situation](#_93iy1hd8zosx) 7

[1.1.2 Risks of Sharing Data](#_vhzgb2m9rgwi) 7

[1.1.3 Why People Still Share Their Data (Rewards of Sharing Data)](#_k14g6mpifnsl) 7

[1.2 Private Information as an Economic Good](#_uxya9krej07k) 8

[1.3 Distinguishing between Private Information and Privacy](#_28nthifj8ulk) 9

[1.4 Similarities and Differences Between PI, PP and IP](#_8en8n43u2oie) 9

[**2 Assumptions**](#_89v290h3okhl) **10**

[2.1 Scope of Analysis](#_x3czp0bujzax) 10

[2.2 Quality of Information](#_8wk6mfarnpqs) 10

[**3 Base Model**](#_z8sapq2fmn03) **11**

[3.1 The naive-valuation market](#_kltnk2ufsu59) 12

[3.2 Market Failure Due To Imperfect Information](#_mswvp46u3b1v) 13

[3.3 Market Failure Due To Inequity](#_2sdkneywgint) 14

[3.4 The PI Equation](#_68a7d42daukt) 15

[3.4.1 The Parameters](#_mz61i79n1sbe) 15

[3.4.2 The Boundary Conditions](#_pd3a6f81xdpc) 15

[3.5 Grounding the values](#_qn2mpmsuvenl) 15

[3.5.1 Obtaining the data](#_z21ho5ovju2f) 15

[3.5.2 Comments on data methods](#_r7e746zhjh9s) 15

[3.5.3 Getting the numbers](#_ronvt2s53e85) 15

[**4 Modified Models**](#_e4rxb0kzki0w) **16**

[4.1 Individuals Value Their Private Information Differently](#_7cofg11ilixu) 16

[4.2 Individuals Value Services Differently](#_3poze8yr8gek) 16

[**5 Analysis**](#_cvqrzu10dnv4) **16**

[5.1 Strengths and Weaknesses](#_tr1qrva3b3bx) 16

[5.2 Sensitivity Analysis](#_b6vq7pdotpmh) 17

[5.3 Network Effects of Data](#_2m3kqh6ql63a) 17

[5.3.1 Impact of Network Effects on Our Model and Policy](#_1hyxmt98plwh) 17

[5.4 Applicability In The Real World](#_loo54kgxwynz) 17

[5.5 Past and Future Analysis](#_g1lxaemw48vn) 18

[**6 Policies**](#_2dpwg192w7wg) **18**

[6.1 Short Term: Legislation](#_ppahvy16e1h4) 18

[6.1.1 Pitfalls of Legislation Only](#_4noaras6lws) 18

[6.1.2 Illustration On Our Model](#_qfxceninb437) 18

[6.2 Awareness Campaign](#_o50isckmvce3) 19

[6.2.1 Illustration On Our Private Information Model](#_8g1xav7iew1m) 19

[6.3 Long Term Execution](#_5e9vmc7fceos) 20

[6.3.1 Decentralised Cryptographic Data Framework](#_5lo2iw3js8k1) 20

[6.3.2 Capabilities of a Crypto-Data Framework](#_53pqtl3zb4hl) 20

[6.3.3 Effect of the Crypto-Data Framework on our Model](#_ev0eub3m9s50) 20

[6.3.4 Issues with the Crypto-Data Framework](#_v1bn8cfrz58u) 21

[6.3.5 Support Required](#_ogiliupfaqsn) 21

[**7 Conclusion**](#_ezzcvsu126b0) **21**

[**8 Policy Memo**](#_nzvtm8gy51x9) **22**

[**9 References**](#_rng333jp373n) **24**

[**10 Appendix**](#_l5591dkuhreu) **25**

# 1 Introduction

## 1.1 Should There Even Be a Price on Private Information?

### 1.1.1 Current Situation

Before we start looking at the cost of private information, we need to first consider if there should even be a price on private information and what benefits such a price system could provide.

Today, people readily give up their information in exchange for “free” goods and services. For example, by creating an account on a certain platform (and revealing your private information to do so) to get free items like electronic books, games, newsletters, courses, etc., or leaving your private information for a chance to win a lucky draw.

The issue with the above situation is that people seem to believe that they are benefitting from the exchange as they get “free” goods and services. However, they fail to take into account that their private information has value too. In fact, the value of their private information may be worth more than the few dollars you save from getting the “free” goods and services. Furthermore, there are other factors to take into account as well - for example what the company does to your data afterwards. However, since most people assume entities like data brokers do not have malicious intentions like doxing and that reputable companies comply with privacy laws and protect consumers’ information, they usually do not consider the consequences of freely sharing their information in such situations.

~~(questioning assumptions → ethics)~~

~~Ideally, we should not put a price on safety (as safety is also hard to quantify). (what does this mean?) However, since we are unable to fully prevent people from freely publicising their data, realistically speaking, it is unavoidable to put a price on safety to mitigate such actions.~~ (We should not be restricting the voluntary action of people sharing their data for whatever purpose, neither we could.)

### 1.1.2 Risks of Sharing Data

(unfairness towards individuals) Once companies obtain private information from individuals, there is no way for the individuals to find out how their data is being handled and even if their private information is kept securely. There is also little enforcement that prevents companies from trading the obtained data amongst themselves to obtain even more benefits. This is especially so for larger companies as they have a monopoly over our data and can use it to rake up high profits at our expense.

~~There are little physical protections that prevent companies from sharing the data. We don’t know what is going on in the companies. They have monopoly over our data and use it to rake up high profits at our expense. (rewrite this more formal)~~

(In event of a data breach) There are unexpected risks that also come about with sharing your data. Data breaches are increasingly common and have affected large corporations like Uber (2016,

<https://www.bloomberg.com/news/articles/2017-11-21/uber-concealed-cyberattack-that-exposed-57-million-people-s-data>), Yahoo (2013,

<https://www.thenational.ae/business/technology/all-of-yahoo-s-accounts-were-hacked-in-2013-data-theft-1.663902>) and even companies that hold sensitive data like Anthem Inc. (2015, <https://www.nbcnews.com/news/us-news/anthem-pay-record-115m-settle-lawsuits-over-data-breach-n776246>).

Considering the frequency and scale of recent data breaches (mentioned above), it has become necessary for companies to remedy the situation by compensating users in the event of such breaches. This happens when users trust the companies to handle their data well, but they fail to do so, resulting in the potential for disaster. Moreover, there are people with malicious intents out there trying to obtain your data and compromising your safety and security.

Thus, we feel the need to develop a pricing model for data that gives a fair price to different kinds of private information. At the very least, this helps to guide people, companies and regulators to what the price of data should be.

### 1.1.3 Why People Still Share Their Data (Rewards of Sharing Data)

(convenience) One reason why people still provide their data is convenience. They do not know how to put a price on their data. They are also ignorant of the true risks and also of the true value that their data holds to companies. For example, a study (<http://www.pewinternet.org/2016/01/14/privacy-and-information-sharing/>) done by Pew Research Centre on Privacy and Information Sharing assessed participants on their willingness to trade specific private information for certain benefits without vague abstractions on what private data they are actually sharing.

(imperfect information) Another reason why people readily share their data is because that they are unaware of the dangers of doing so. As analysed by Ferenstein Wire, a syndicated news service, (<https://readwrite.com/2016/01/18/pew-privacy-study/>), the above study conducted by Pew Research Centre revealed that common daily scenarios when phrased differently, i.e. explicitly stating the private information extracted in these scenarios, actually resulted in people being less willing to trade in their private information, showing that people did not know the full extent of the private information they were sharing.

~~An example taken from the Pew study would be that instead of mentioning placing a GPS device in your vehicle, the situation was phrased as a device that “allows monitoring of your driving speed and location”, i.e. your driving habits. 45% of the research participants indicated that the benefits they were offered in exchange (discounts) were not acceptable, which is vastly different as compared to 9 in 10 smartphone users using location-based services on their phones (~~[~~http://www.pewresearch.org/fact-tank/2016/01/29/us-smartphone-use/~~](http://www.pewresearch.org/fact-tank/2016/01/29/us-smartphone-use/)~~).~~

(necessity) They want the service, but there are no alternatives.

(social benefits) Individuals share data also because they believe in the welfare generated from learning from their data. Technology companies have built a good reputation with their corporate social responsibility initiatives.. ~~There are also other unexpected benefits that arise when people share their data, but these benefits were not taken into account by the people whose data is shared, with or without their knowledge.~~ For example, Facebook’s social platform has been experimenting on tools that benefit the society, such as early detection of health issues (<http://www.bbc.com/future/story/20140220-can-giving-up-privacy-help-us>), awareness of suicidal thoughts

(<https://techcrunch.com/2017/11/27/facebook-ai-suicide-prevention/>) and terrorist tendencies. (<http://www.bbc.com/news/technology-42158045>) The potential social benefits outweighs off individual concerns the loss of privacy.

## 1.2 Commodifying Private Information

~~(Why are there not solutions, why is there not a price on data)~~

~~(Economics concepts - nature of data as difficult good to price.)~~

There is no standard to base the price of something as intangible as private information. Comparing to the price of a consumer product e.g. Coke, where the market price of Coke is charged based on the cost price of producing the Coke (which can be approximated to a general universal cost of producing the Coke). However, since the cost to the consumer is in the risks involved in revealing such information rather than the cost, it becomes a question of how much an individual foresees the risk as and how much the individual values his or her safety. Hence, it is impossible to come up with a general universal cost for an individual’s private information, making it unmarketable.

Mention the results of the OECD paper? Mention other major reasons, such as it depends too much on individual preferences.

Non-rivalrous good, easily transferable.

## 1.3 Similarities and Differences Between PI, PP and IP

Private property is easily enforceable with physical barrier to intrusion, its visibility and social condemnation of intrusion of private property. Intellectual property is still largely enforceable.

Even though music and film piracy once obtained, can be shared with everyone at no cost, it is difficult for an individual to make monetary profit. Despite the rampant piracy of media on the Internet, production companies are still able to survive and earn profit through legal means. Moreover, infringement of intellectual property could not be done by a company, because it will be condemned with fines and lawsuits.

However, in the current world, no similar physical protection is possible for personal information. We have to recognise that once data is disclosed, there is nothing physical preventing companies from selling data among themselves. It is especially hard to regulate, especially with the capabilities of technology today.

~~Even though Private Information (PI), Private Property (PP) and and Intellectual Property (IP) are related, PI is quite different from PP and IP. PP and IP can be defined in terms of tangible losses. The loss of PI might lead to the loss of PP and IP, like for instance a data breach might lead to a leak of your bank account password, which leads to a loss in PP. Also, the loss of PI in the form of email account passwords might lead to a loss of IP such as patents, blueprints, and unreleased episodes of shows (such as Game of Thrones).~~

~~In other words, even though a loss of PI leads to a chance of a loss of tangible wealth, the value of the loss of PI itself is hard to quantify. Any more?~~

## 1.4 Distinguishing between Private Information and Privacy

We find the need to distinguish between \textbf{*private information/data}* and \textbf{*privacy}*. People are often willing to exchange their private data for benefits such as discounts and convenience, but are willing to pay a price to protect their privacy.

We believe that the difference lies in the method of valuation - privacy is the valuable and meaningful personal space that is violated when a third party gets hold of your data, while private data is just a good that can be traded for benefits between any entities (be it individual or companies).

People value their privacy, more than they value their private information. (quote the paper or other resources) We will differentiate them in our models.

# 2 Assumptions

## 2.1 Scope of Analysis

The graph represents an average individual in a single transaction.

This shows the aggregated transactions between the individual and the company?

The notion that no other cost is involved in the transaction. Time spent is absent from the graph (not very likely users have time for every internet services out there).

## 2.2 Quality of Information (Data Valuation Framework)

Before attempting to arrive at a price for information, we first attempt to quantify the notion of the quality of information transacted, because different amounts and kinds of information are valued extremely differently. We make a huge assumption here by coming up with an idea of a totality of all information, which is an idea that all the information of an individual can be represented by a number, say 1, and that the value of each piece of data is then a fraction of that totality.

We first identify several categories of data: <http://www.visualcapitalist.com/much-personal-data-worth/>

To the companies, they value the quality of data based on how much economic benefits they can get out of it, be it extra revenue, market share, or customer loyalty. In other words, they **value the Private Information objectively as economic goods**. We listed out the information that could help such companies: Full general profile, Full Social Network profile, Current location, Marital Status, Web Browsing History, Email, Health condition

We are not considering the information sold by hackers, which is a breach of security and not a breach of privacy. We are also not considering the analytic work that data companies might have done, to value-add their data product. Here, we are simply considering the value of such raw PI to trustworthy companies and entities.

An average person values the quality of data transacted based on how much meaning or how intrusive the data is to him. In other words, they **value their data not based on how useful it is, but how much privacy it infringed upon**. We listed out the different types of data: Name, Gender, Marital Status, Photos and Videos, Address, Purchase history, Current Location, Credit card info, Social Security Number, Passwords, and medical info.

**Personal Data (some changes to the values to make them look better) - two levels**

**(adding explanatory information / constituent factors on the left)**

|  |  |
| --- | --- |
| Current location: | 0.0017 |
| Web Browsing history | 0.0048 |
| Email: | 0.0564 |
| Marital Status: | 0.0665 |
| Health condition | 0.1174 |
| Full social network profile: | 0.7533 |

**Privacy**

|  |  |
| --- | --- |
| Gender and name: | 0.0081 |
| Marital status: | 0.0306 |
| photos and videos: | 0.0452 |
| Home address: | 0.0626 |
| Purchase history | 0.0781 |
| location: | 0.1170 |
| credit card: | 0.1309 |
| SSN: | 0.1508 |
| passwords: | 0.1832 |
| medical: | 0.1934 |

This is how an average composition of information quality might look like. As we have no way of finding a universal average of price of different data types, these values have been obtained by reverse-engineering our model. These are also the values which we will use as examples throughout the report.

# 3 Base Model

The central question we are tackling in this model is: What should be the price of private information and privacy? There are three ambiguous terms here - “cost of private information”, “cost of privacy” and “should”, and all of which must be clearly defined in our base model.

What is meant by “cost of private information”? Since we take “Private information” to be the objective economic value of the information, it is simply the price that personal data is traded between companies. Or if individuals wish to sell their data voluntarily, it is also considered to be the sale of PI and should be judged by this cost.

.

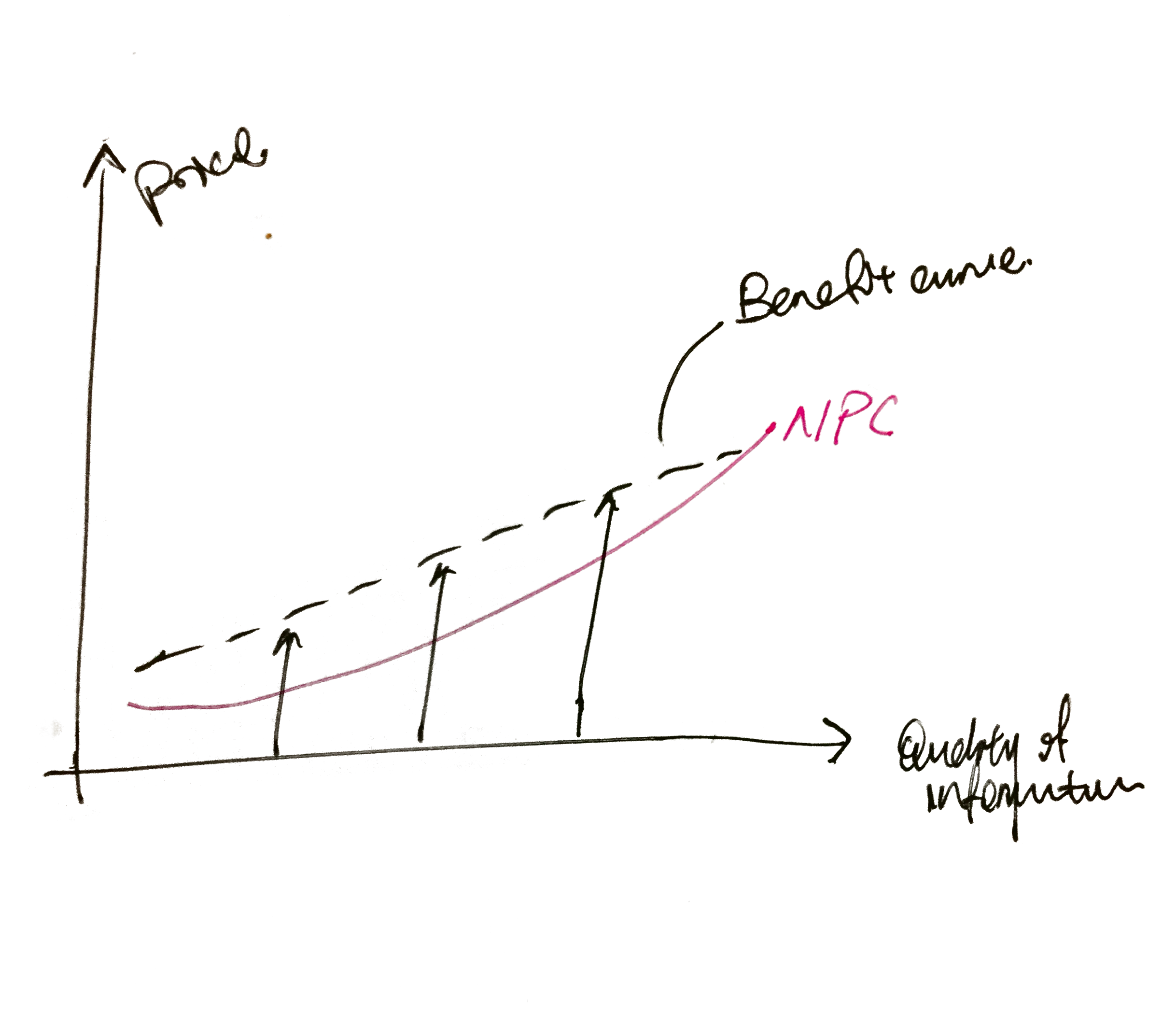
What is meant by “cost of privacy”? Since we take “privacy” to be protection of personal and meaningful information to the individual against unknown entities, the ”cost of privacy” is the cost that individuals think third parties should pay to access certain information.

What is meant by “should”? It means that the model sets an ideal standard we should follow, which is a price that is fair to both the data-provider and the data-purchaser. This also begs the question - why are we not already at the ideal?

~~What is meant by “price”? Currently people give up their data in exchange for free services - can we quantify such an exchange? What is the issue with such an exchange, and why are we not at the ideal? The same person prices their private information differently in different scenarios - whether be it keying into their Facebook account, or sending it to an unknown individual.~~

## 3.1 The initial idea: the naive provider of PI

The graph here plots the values of different types of PI per transaction for an average and naive individual.



Using Facebook as an example. Mark out the points and label them, and use these points.

(Make the following paragraphs more understandable)

(Quality of information) The horizontal Qty-axis refers to the completeness of data in terms of value. The domain of this axis is 0-1, corresponding to the percentage completeness computed using the Data Valuation Framework.

(Value) The vertical Value-axis refers to the value obtained by the data-provider or data-purchaser from a certain transaction. The value-axis is a generalization of price, to include intangible and non-measurable benefits that applications and services provide.

(Naive Valuation Curve) The curve in orange is the NVC. Every point on the curve refers to the amount of value naive individuals currently expect when they disclose their information to a trusted company. The naive individuals are unaware of how much their data is really worth, and the true risks involved in providing others with your data.

(Delta offers) The vertical lines indicating the deltas indicate an offer. An example of such an offer could be a $5 discount voucher if you disclose your payment information. If the payment information is worth 0.2 on the quality scale, this corresponds to a delta offer with a height (value) of $5 at 0.2 on the quality scale. If the delta offers intersect with the NVC,

(Benefit curve)

Next, we depict a graduated offer system, referring to services which offer you increasingly better services with more information you provide. This is similar to aggregating many separate offers (delta peaks) over 0-1 quality scale. This results in a benefit curve.

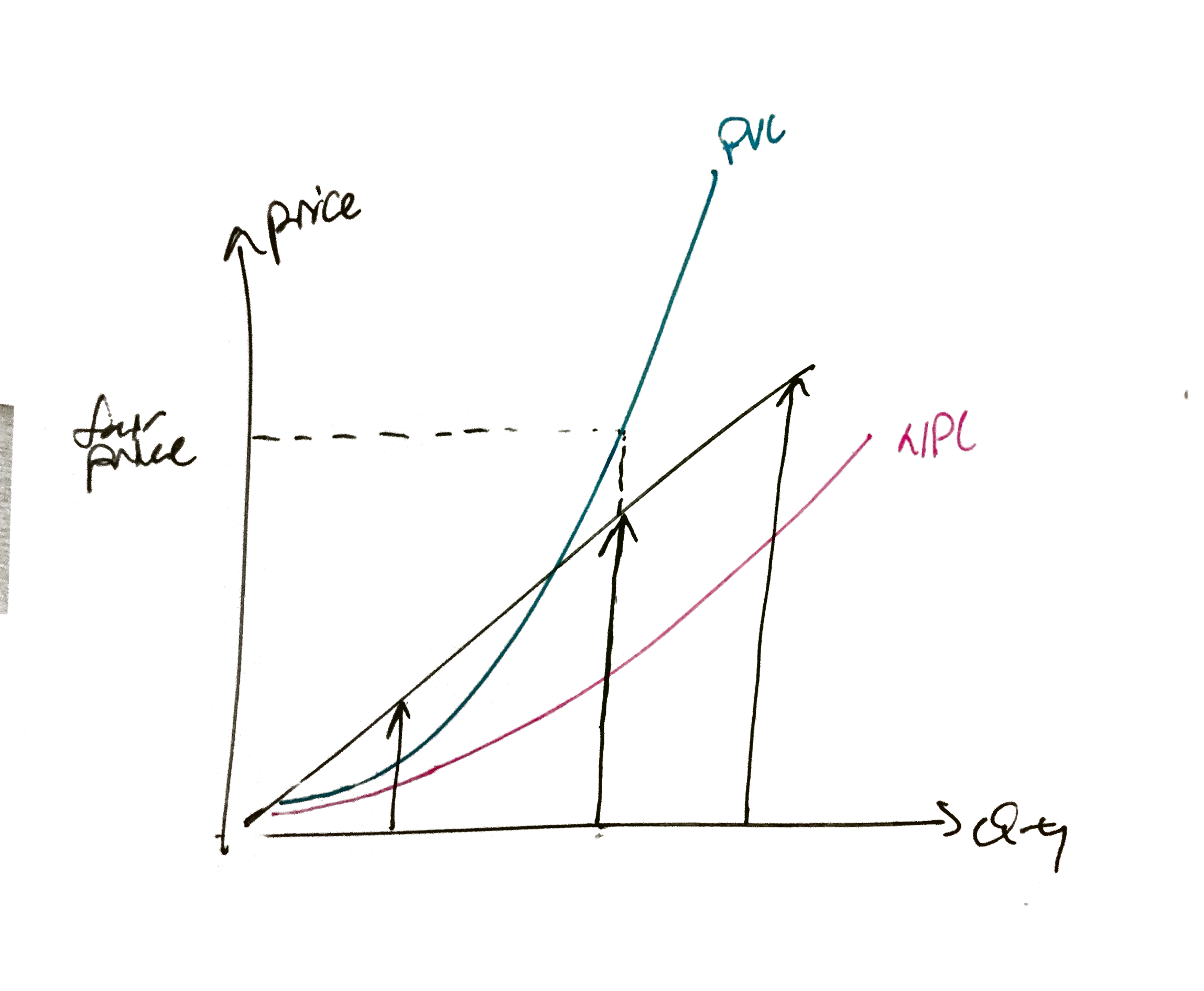
Let’s take Facebook as an example. Providing your basic credentials (0.02 quality) gives you access to post your information. Then, you have the option to provide data of your workplace and past schooling experience (extra 0.03 quality), in exchange for the convenience of adding your past colleagues and classmates as friends, shown as benefit p1. You can additionally provide your credit card information (extra 0.10 quality) in exchange for gaming services, and it gives you a total benefit of p2.

## 3.2 The enlightened provider of PI

So how do we find an ideal price for \textbf{privacy} in our model? We can borrow the Economics concept of Market failure. We understand that, if the individual becomes more informed about 1) the true value of his data, and 2) the true risks of sharing data, he will want higher prices and value for the PI that he provides.

(Privacy Valuation Curve) The PVC represents true valuations of not revealing sensitive personal information by an enlightened individual.

(Ideal Price) The ideal price would then be the intersection should be between the PVC and the Benefit Curve.. This is the fairest price for the individual who provided the data, and the companies who buy the data.

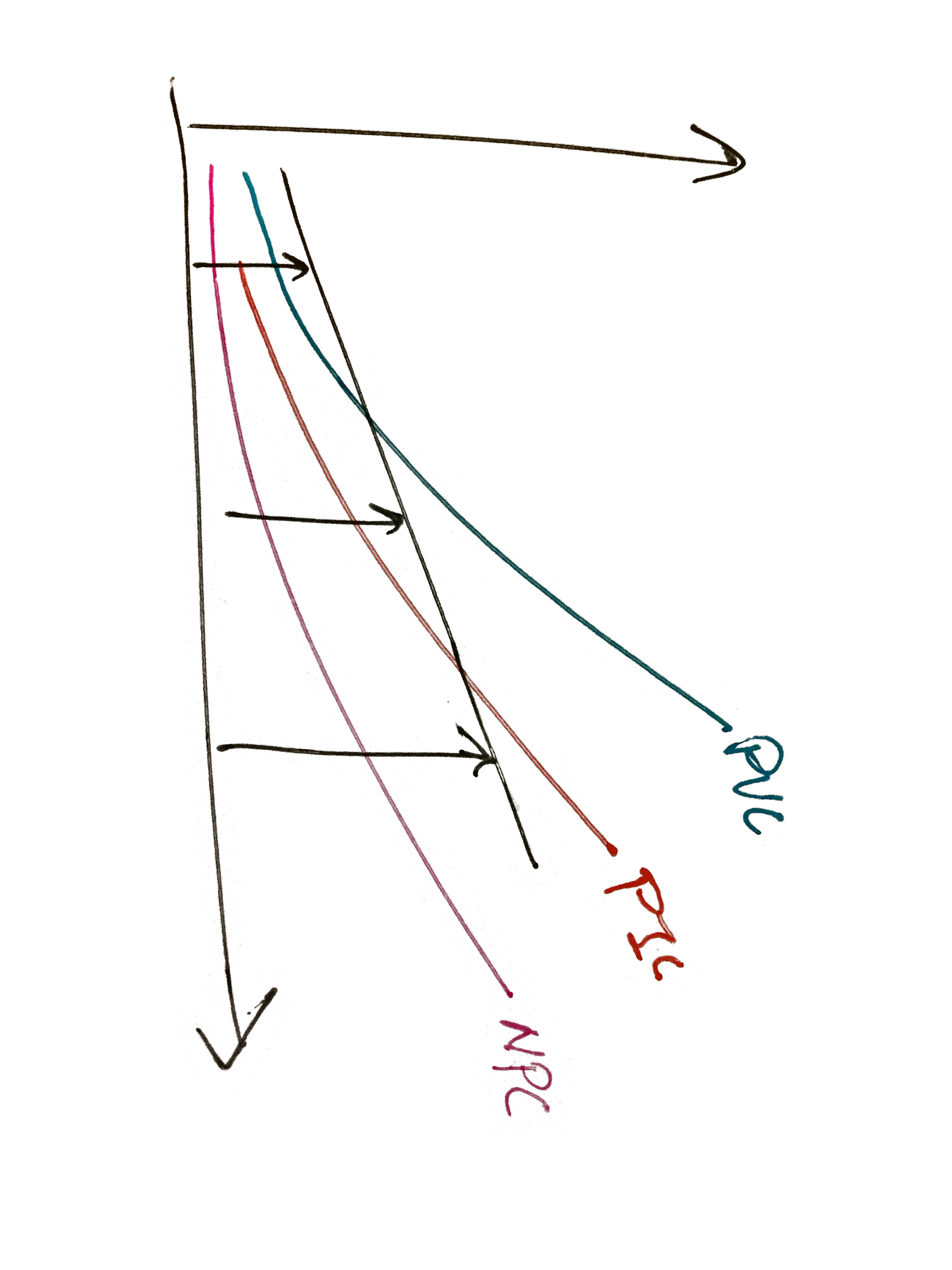


## 3.3 The risk-ignoring Companies

How do we then find an ideal price for \textbf{Private Information}?

(Private Information Curve)

Again, we borrow the concept of market failure and observe that while companies have close to perfect information of the value of PI, they choose to ignore the risks it brings to the individuals when they sell their information. [Why?] Thus, they demand a higher price when they sell the PI as compared to the naive data providers. [missing link] This price is also the objectively accurate price of the PI.



While we recognize that network effects play a role in determining the fair price of a batch of determined data, our model currently doesn’t consider that we we take the price of the PIC to be the fair price of PI as we are dealing with an individual transaction, independent of other transactions.

## 3.4 The PI Equation

Here, we find a generalized equation that works for the PVC, the PIC and the NVC.

The equation: AB(Me^(x^(1/(b+c)))-1)

### 3.4.1 Explaining The Parameters

M = maximum price for an average person’s PI or data. (at complete info of quality 1). M = 3500 for privacy and M= 100 for PI.

b = amount of knowledge of value of data in arbitrary units. b is low (around 0.05) for NVC.

c = amount of risk considered in arbitrary units. c is 0 for the PIC.

A = risk factor (net worth). Is a specific factor given to each person. If taken to be a financial risk, it can be interpreted as net worth/ average net worth of individuals. If an individual wants to consider a reputation that is not financial, it is also a multiplier here. The risk factor is roughly linear because if we consider the risks involved a

B = discount factor (for the benefit of communal efforts)

### 3.4.2 The Boundary Conditions

Maximum price (limits)

0 price at 0 data.

Exponential form, giving more value to a combination of a few information at once.

## 3.5 Grounding the values

### 3.5.1 Obtaining the data

Describe how the data was obtained

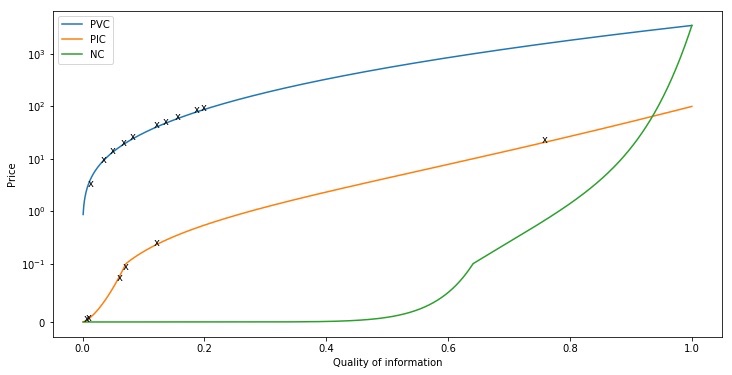
|  |  |  |  |
| --- | --- | --- | --- |
| <https://www.thinksmallcell.com/Opinion/how-much-do-you-think-your-personal-data-is-worth.html> | | | |
| <http://www.totallymoney.com/personal-data/infographic/> | | | |
| <https://thegreatdissonance.wordpress.com/2017/06/03/how-much-is-data-worth/> | | | |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| <https://www.trendmicro.com/vinfo/us/security/news/internet-of-things/how-much-is-your-personal-data-worth-survey-says> | | | | |
| <http://www.totallymoney.com/personal-data/infographic/> | | |  |  |
| <https://thegreatdissonance.wordpress.com/2017/06/03/how-much-is-data-worth/> | | | | |

### 3.5.2 Comments on data methods

Advantages/disadvantages. Kope from the paper.

### 3.5.3 Getting the numbers



(THE GRAPH OF THE COMMERCIAL VALUES OF PRIVATE INFORMATION)

(INSERT TABLE)

Insert graph within the value model

# 4 Modified Models

## 4.1 Individuals Value Their Private Information Differently

While Chapter 3 provides the case for a general individual, the difference in individuals could be accounted by how people will rate their Naive Valuation Curve and Woke Supply Curve.

As data is more valuable for someone’s net worth. The Naive Supply Curve and the Woke Supply Curve will be larger. The benefits from the curve may increase as well.

The generational difference can also be accounted in the supply curves. A more socially reclusive individual would not want to share too much information. In his case the NSC and WSC will also be greater. This results in the equilibrium quantity to be lower, they are willing to share less data. (ARE THE SURVEY RESULTS ACCOUNTING THE DIFFERENT GROUPS)

## 4.2 Individuals Value Services Differently

A top business executive could benefit more from Facebook by sharing. However, he is also more reluctant to disclose information to a trusted company. Yet, he is still likely to disclose information because he values the service more than ordinary people. This is illustrated by both shifts in NPC and PVC. (SOME QUANTIFICATION?)

# 5 Analysis

## 5.1 Strengths and Weaknesses

Give an awareness of the value of their data.

Comment on the assumptions?

Comment on the veracity of the price points?

Is our model qualitative only?

Can our model be good illustration to the proletariats?

Does not capture the exclusiveness of data provision? If something is widely available does it still have value?

Consider the extreme points.

This mainly work for the average person.

The justification of the quality of information?

## 5.2 Sensitivity Analysis

Test changing the different parameters, see how much the price change.

## 5.3 Network Effects of Data

Other factors external of the individual can also affect the value of the individual’s data. Individuals are part of a network, and when more and more individuals join the same network, network effects come into play, increasing the value and the desirability of each individual’s private information. One example that shows the significance of network effects is a social network like Facebook. Each time an individual creates a Facebook account and joins the network, the value of the network increases for each existing user, for example through more knowledge being shared, without the intention of the newly joined individual (the individual joined the network for his or her own personal reasons, and not for the purpose of growing the network and giving benefits to existing users).

Similarly, we can apply the network effects to the private information of individuals. As individuals form more networks and connections, for example through meeting new people and making friends, their personal information are highly likely to contain information about other people as well. This makes the individual’s private information much more valuable as attaining the individual’s private information now gives a company clues about the other people in his or her network, from whom companies can then attain their private information. This is with the assumption that groups have similar interests and hence the people in the individual’s network are more likely to be potential consumers for the company as well, allowing the company to benefit greatly from the initial individual’s data. As a result, companies have the incentive to buy data from individuals at a lower cost because after aggregating it, they can derive more value from the data.

### 5.3.1 Impact of Network Effects on Our Model and Policy

Network effects have a significant impact on our model and policy as they heavily influence the value of the individual’s private information. Looking at our model, by purchasing private information from the naive data providers at a low value (NVC in the graph), companies can then gather all the data from many individuals and sell the data at a much higher value due to the network effects. This could have been the main contribution to the higher PVC curve in our model, rather than the previously mentioned perfectly-informed individual, but we did not model this as it is outside the scope of our solution (looking only at the effect on the private information of a single individual).

On the other hand, by properly utilizing network effects, the value of an individual’s private information will rise significantly. This is especially beneficial for the individual if he or she knows the true value of his or her private information, and trade it for an equivalent amount of benefits if necessary. As a result, fairness will be enforced as the producer, i.e. the individual with the private information, gets compensated for the amount of data revealed while the consumer, i.e. the company receiving the private information, pays the individual with the true value of the data.

## 5.4 Applicability In The Real World

Compare with more price points.

massive data breach - compare with the fines given

## 5.5 Past and Future Analysis

Data is getting more and more expensive?   
As people are made more aware of the dangers of sharing their information?

# 6 Policies

## 6.1 Short Term: Legislation

Legislate that companies protect users’ data. Use the pricing structure.

Make the transaction of data illegal across borders? Transaction of data can still be legal if companies protect individual’s data?  
Enforce the ability for users to download their data. One example is downloading Facebook data.

~~(Market Equilibrium) The individual want to maximise benefit. They will share information until the positive difference between the benefit curve and the cost curve is maximised. This is when marginal benefit outweigh the cost. While the individual may want to reveal the juicy details about her past relationship experience for more benefits, perhaps more following, from the platform, she may want to withhold it due because the risk to her is too high. So the individuals would give up that amount of information Q, as such value V.~~

(<https://www.privacylaws.com/Publications/enews/UK-E-news/Dates/2013/3/ICO-to-evaluate-effectiveness-of-fines/>)

([https://pagefair.com/blog/2017/gdpr\_risk\_to\_the\_duopoly](https://pagefair.com/blog/2017/gdpr_risk_to_the_duopoly/))

(Consider the similar legislation out there already. References needed)

### 6.1.1 Pitfalls of Legislation Only

Cannot prevent companies from leaking information without the knowledge of users. We do not know if companies still sell the data, since communications within a company should be encrypted.

Despite such laws, data breach still happen. We often hear companies getting fined despite regulations in place. (refer to the examples in introduction).

The root cause it that these companies store data on their servers. This opens up the vulnerability to intrusion is thus always present. While companies that are more focused on security and privacy like banks would invest much money, firms less focused on security (like gaming) does not have the capacity to make the their data framework fully secure.

### 6.1.2 Illustration On Our Model

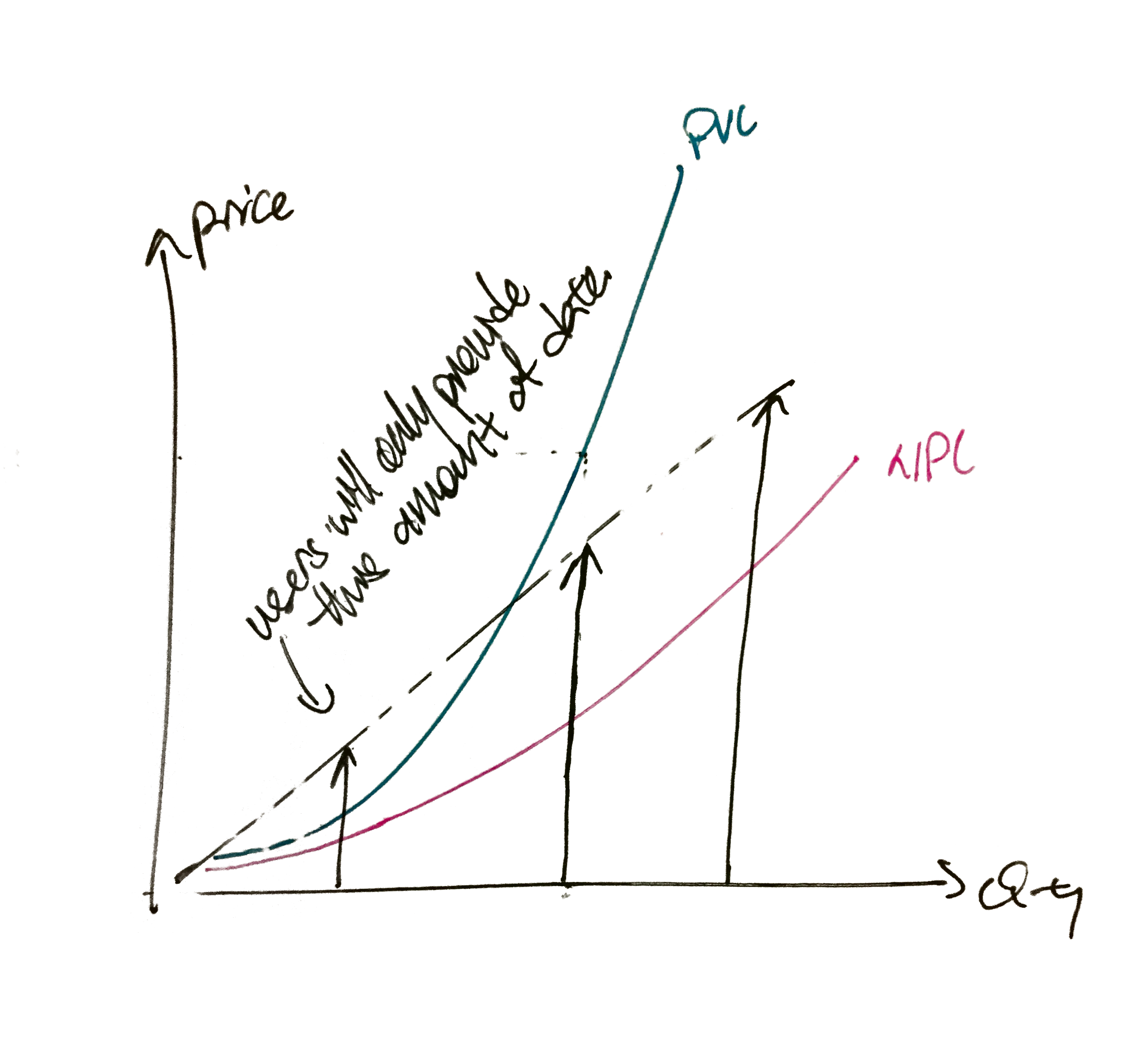
No effect?

## 6.2 Awareness Campaign

We recommend the decision maker to recognise the importance of privacy and begin an information campaign. We need to let the people know, though press release and published information, on the dangers of disclosing your information to the public. We should also encourage publicity of data breaches to illustrate how real the problem is.

### 6.2.1 Illustration On Our Private Information Model

We assume now, that everyone has a clear understanding that their privacy, and safeguard their personal information as much as their privacy. We illustrate this on our information curve. People will maximise the PVC w.r.t. their benefit curve. In the example of Facebook, people may not longer want to use Facebook’s messaging function each other anymore, because they can be held liable for anything they transmit.



Making the entire population paranoid about their data is neither ideal. If all people from now on strongly believe that any information will eventually public, the economy will not function. As the society is made up of communications between one another, that channels that are not fully privacy respecting, society today can longer function without electronic communication.

However, realistically, the world is not going to behave this way. The population will never be paranoid about their data overnight - despite Snowden’s highly publicised revelations of NSA’s privacy overreach, most of the Americans are apathetic to the affair. People with high computer literacy also have not changed much other their habits either, due to the lack of alternative technologies. Currently available fully private solutions such as Tor web browser is insecure. Therefore we need to nurture alternative technologies that fully respect the privacy of the users.

## 6.3 Long Term Execution

We propose supporting and protecting a crypto-data framework for public use. (<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7163223>

openmined.org)

[CITATIONS!!!]

### 6.3.1 Decentralised Cryptographic Data Framework

A decentralised cryptographic data framework allows users to sell the benefits from the usage of their data without even disclosing their data. It is possible with a few recently developed methods, as well as a few assumptions. One such framework is OpenMined (openmined.org). It aims to provide such a framework with deep learning, federated learning, homomorphic encryption as well as blockchain smart contracts. It assumes a fully rigorous framework and a trusted and verified third party known as the Oracle.

This is how an arbitrary company can train their data on this framework. The data scientist provides the initialised model to the oracle, together with the bounty. The oracle will negotiate with nodes that are fully controlled by individuals to request to train on their data data. The oracle then passes the model to the individual’s node for training on the data. The model passed is homomorphically encrypted to protect the learnt information. Then the model is then returned to the oracle, and if the information they provided is evaluated as correct and useful, the user is compensated through the contract. The oracle is independent and autonomous, hence no one have access to the model until it is complete. Then the oracle homomorphically encrypt the model again and seek another users to train the model.

### 6.3.2 Capabilities of a Crypto-Data Framework

Users can now allow their private data, such as their chat or search history, to be used for training a model without divulging the data at all. The user can be compensated for every time the data is being used. Companies can make use of people’s data from the public without access to their data. This is a step to fulfilling the original dream of the Internet to be decentralised and equal.

### 6.3.3 Effect of the Crypto-Data Framework on our Model

As people are guaranteed that their data will not be shared, more people are willing to share the benefits of their data for a low price. This will decrease the PVC curve to low levels. People “paranoid” about their own data will the opt for such decentralised services because they rightly know that these services will never require anyone to have access to the data. As more people realise the benefit of not having their data being stolen. They move on the decentralised data framework.

(insert picture here) Curve goes down. PVC facebook on top, PVC crypto model below.

People who are very savvy and conservative about their data now move on to data that protect the users’ privacy.

In recent years we have seen the rise of apps that market themselves with privacy features, such as Telegram and Snapchat. While these apps may not be rigorously secure, it illustrates the current desire for access to privacy of many people. Moreover, the society can make use of the vast amount of data to provide more value to the economy, and also for altruistic purposes.

(<https://research.googleblog.com/2017/04/federated-learning-collaborative.html>)

### 6.3.4 Issues with the Crypto-Data Framework

An important thing to consider is the question why there is such a lack of services in the first place. While the methods have only been recently developed, the state-of-the-art encryption algorithms were still the same ones we used since the start of the millenium. We could probably attributed to the lack of awareness of data privacy from individuals.

Moreover, it is difficult to start a decentralised application given the pervasiveness of the major players in the market. Diaspora is proposed to be a decentralised social media framework, but it size pales in comparison to Facebook, and the benefits of privacy is insufficient to convince the users to move away from the comfort and wide network on major social networks. There is no competing alternative version of platform that are decentralised and focuses on privacy.

### 6.3.5 Support Required

Therefore we call for the support the growth of decentralised frameworks that that only the user full control over their data. It is best that we start supporting and protecting decentralisation of the Internet for the better of the people.

# 7 Conclusion

We quantified personal information in terms of normalised quality of information etc. We have provided a pricing structure for PI. We recommended some policies.

# 

# 8 Policy Memo

Write a two-page policy memo to the decision maker on the utility, results, and recommendations based your policy modeling on this issue. Be sure to specify what types of PI are included in your recommendations.

Dear Decision Makers (I don’t think we will be showing any graphs here)

(Introduction to privacy) Privacy is an inherent human right, and a requirement for maintaining the human condition with dignity and respect. There should be a price to privacy. ~~Citizens’ privacy is a national issue. Entities can interfere with our national decisions by targeting the masses with customised misinformation campaigns.~~

(Individuals value privacy differently from private information) Research after research has shown that individuals tend to misvalue their privacy and private information, in fact, this holds more true for privacy than their private information. While individuals are willing to exchange some of their personal information for certain benefits, they are willing to pay much more to perfectly reverse a data breach i.e. completely cover up their private information after it has been revealed.

(Our model) The normalised quantity of information. We have illustrated the discrepancy with the two curves. We consider private information to \_\_?. While the individual may be comfortable in giving pieces of information to trusted sources at low value, they are vigilant against, and the rapidly increasing curve illustrates the relationship. The bottom curve refers to the individual valuation of their own personal information, while the top curve refers the individual valuation of their privacy.

(The resulting market failure due to imperfect information - explain only)

(The resulting market failure due to inequity - companies monopolising data?)

(where did we get our information)

(Analysis of our model)

(Network effects)

(Policies - legislation) We need to strengthen our data protection.

However, the end goal is not to let companies get in the way monopolising data control.

(Policies - awareness)

(Policies - crypto-data framework)

The original dream for the web, was for everyone to participate in the common neural equally for the betterment of the humanity. The web first pioneers had the ethos that data should be freely available for all that wanted. However, we are far from the goal.

Briefly explain what is it and its potential benefits, and the state the issue why this is not implemented.

I look forward to your resolute decision to protect citizen privacy. Please fund and protect decentralisation of the Internet for the better of the people.

Regards

Team 87374

~~We recommend a few actions to take. However, recent developments allow this to be possible. We should move on to a crypto-data framework. It is possible that.~~

~~(please rephrase following)~~

~~Bring the power back to the people. The original dream for the web, was for everyone to participate in the web equally. The web first pioneers had the ethos that data should be freely available for all that wanted. Everyone hosted their data. This altruism began to fade. Entrepreneur realised that , Centralised services collected data monetised it. You freely service exchange your own data to access the free service. The service now own the world data. The web has become centralised.~~

~~This also has serious implication on user privacy. People lose control of their data and are not compensated for it.~~

~~People do still have the free choice. While users maintain ownership.~~

# 

# 9 References

# 

# 10 Appendix

Likely none