University of Lund
School of Economics and Management
Department of Economics
Bachelors Thesis

Price Dispersion and Price Level: Online vs. Offline- A Quantitative Analysis

Erik Händén Svensson 2006

Abstract:

With the use of basic theory the effect of the Internet's reduction in search cost is analysed. Two samples of prices of CD-albums, one from the offline and one from the online market, are used to test the hypotheses that prices are lower and less dispersed on the Internet. The results are supportive of lower prices online but inconclusive on the issue of price dispersion. Some pointers for new research are also included.

Keywords: imperfect information, price dispersion, Internet, CD

Author: Erik Händén Svensson

Supervisor: Jerker Holm

Table of contents

1. Introduction

1.1 Background

When e-commerce first started to realise some potential in the late 90's many predicted it a revolution. Not only would competition increase but costs would also fall and thus leave the consumers with lower prices. Anyone able to design a website could start competing on the Internet and middle hands that were no longer adding value would be cut out of the supply chain. Perhaps more interesting was the Internet marketplace's drastic reduction in search cost - an obstacle for the *Law of the Single Price* to hold.

The *Law of the Single Price* is based on the perfectly competitive market and simply states that, in equilibrium, a good can have only one price. The relevant price is the opportunity cost. Since information is perfect, any deviations in stated prices would have to reflect differences in transaction costs. For example, if purchasing from a particular seller involves more travelling, the stated price should be lower so that the opportunity cost is the same.

Now, on the Internet the location of the seller, as long as the cost of delivery from all locations is the same, becomes irrelevant for the opportunity cost. Furthermore, irrespective of the delivery cost's magnitude in relation to the conventional market, as long as it remains constant, is not a cause of price differences on the Internet.

In the conventional market, price information is not perfect. Price differences could thus reflect the cost of searching for a lower price. The opportunity cost of the time and effort spent searching for a better price could be substantial. Furthermore, information could be asymmetrical meaning that some buyers are better informed than others.

The Internet shopper can within minutes search and compare prices and find the lowest price offered on the web. Hence the cost of searching for a lower price, within the Internet market, is very low.

1.2 Purpose

The aim of this paper is to investigate whether prices, of the representative good CD-albums, indeed are lower and less dispersed on the Internet, compared to the conventional market. The results could potentially cast some light on the importance of search cost as an obstacle for the *Law of the Single Price* to hold.

1.3 Outline

Section 2 of this paper outlines the theory behind how imperfect information causes price dispersion and why it is expected that prices are lower on the Internet. When search is costly, firms have more scope of charging different profit-maximising prices along a downward sloping demand curve. So, going from a state of imperfect information, where sellers are price makers, to one where price information is free and sellers are price takers at the competitive price, price dispersion and the price level falls. Consequently, the welfare implications of reduced search cost is outlined.

In order to test whether price dispersion is lower on the Internet empirically, several studies have been conducted. For the assumption of non-differentiation to be reasonable, homogenous goods such as CD-albums have frequently been used. Bailey (1998) pioneered with a study from 1997 on books and CDs as well as other goods. However, possibly thanks to immature online markets as argued by Pan *et al*. (2004), price dispersion was found to be higher online than offline. But later studies using CDs (Brynjolfsson and Smith 2000, Lee and Gosain 2002, Ancarani and Shankar 2004) have also fully or partly, depending on the measure of price dispersion used, rejected the hypothesis.

This papers secondary hypothesis that prices are lower on the Internet, has also been tested before. To mention a few, Lee and Gosain (2002) and Friberg *et al*. (2000) both found evidence in favour of lower prices online.

The above studies on price dispersion were all conducted in the US market. After the methodology and data is outlined in section 3, section 4 contributes with a quantitative analysis on data from the Swedish online and offline market.

The delivery cost within Sweden is constant and delivery charges from abroad are assumed to be prohibitively high. This fact is used to define the online market. The data also allows for a test of the market definition, which is supportive. Prices are found to be statistically significantly lower on the Internet. The evidence for less price dispersion is inconclusive-some in favour and some in favour of an alternative hypothesis that price dispersion is the same in both markets. However, no evidence of price dispersion being greater on the Internet is found.

The findings in Section 4 leads on to a discussion in Section 5 about possible explanations for the inconclusiveness on the issue of price dispersion. A few pointers for further research are thence generated before Section 6 concludes the paper.

The *conventional* or *offline market* consists of *stores.* On the *Internet*, or *online*, sellers are referred to as *sites.*

2. Theoretical Underpinnings

2.1 Braking the Law of the Single Price

In traditional theory the *Law of the Single Price* dictates that markets will make sure a good only trade at one price. A Walraisan auctioneer makes sure buyers and sellers find each other at this price. The traditional theory is, however, not completely blind to the fact that in reality almost all goods have a distribution of prices since it acknowledges that heterogeneity – differences in quality- and specific taxes that are passed on to consumers, cause price differences.

Why is it then that price dispersion exists even in the market for homogenous goods such as CD-albums? For one, even though the commodity itself cannot be differentiated, the sellers can differentiate themselves on brand and service reputation. For example Clay *et al*. (2001) found evidence of brand and service differentiation in the online book industry. Differentiation on selection is typical for conventional stores but not so on the Internet. The reason is that the sites do does not need to stock a good to display it on offer in the same way as the conventional stores do[1]. Seller differentiation is, however, as (or less) likely online as offline and since the main interest lies in comparing the two markets service differentiation is assumed a constant driver of price dispersion. Stigler (1961) makes the case that, bar heterogeneity, prices differ because of *ignorance* in the market - in more familiar terminology: buyer's (and seller's) information is *imperfect.*

Brand differentiation and imperfect information is, however, not the only explanations of price dispersion. Some examples of how *transaction costs* causes price dispersion, and also the assumptions about transaction cost as a driver of price dispersion, are presented below.

2.2 Transaction Cost

A purchase often involves costs that are not included in the price paid. The most obvious are the money and time spent on getting to the store. The cost can differ very much from person to person depending on the opportunity cost of time, where they live and how many purchases they make at the same time (the cost should be divide by the number of purchases made at one location to have a

[1]One retailer pointed out for me that he can order any CD a customer wants, without charging extra, and have it in store within two weeks, the same time (or even less) it takes sites to deliver a CD not currently stocked.

fair appreciation of the cost associated with each purchase). It should be clear that in the conventional market location is very important. Price differences can, for example, reflect the higher priced store's favourable location both in terms of distance from the buyer but also in terms of distance to other stores frequented by the buyer.

Sometimes a transaction incurs a transaction fee. Some stores charge buyers for using debit or credit cards, whereas others will not, and hence it might result in their prices to differ. It should be noted that transaction fees do not necessarily create price distribution. For example in the equity market brokerage fees is not causing the price of a share to show any distribution at a given point of time.

Associated with any purchase are also intrinsic costs. There are issues of uncertainty – whether the merchandise will live up to expectations, return warranties and quality guarantees. When buying on the Internet further intrinsic costs could be anxiety about leaving out ones credit cards number and uncertainty of delivery. Such costs can vary from site to site, depending on reputation and how the site is generally perceived. Those intrinsic costs are thus interlinked with service differentiation and advertising. Intrinsic costs are very individual- some people who like shopping might even experience negative costs.

The proximity between buyer and seller *per see* is not a variable on the Internet as all sites are just a ”click away”. Differing distances but also differing taxes and tariffs will instead show up in the shipping bill. Conveniently, within the Swedish online market, those bills vary, if at all, only a little. There are several additional costs associated with ordering from abroad. Often foreign site's language, or their shipping policy, prohibits purchase and delivery to Sweden. It will be assumed that such costs associated with buying the CDs in the sample from a site situated abroad are prohibitively high. Hence it is established that 1. the online market is defined as Swedish sites and 2. the location cannot be a driver of price dispersion in the online market.

2.3 Imperfect Information

When modelling economics it is often assumed that all agents have perfect information regarding demand, prices, costs, supply, available technology etc. Implicitly, information is free in the sense that it is readily available with no costs associated with its acquisition. The standard rationality assumption forces the agents to use the information to achieve their optimisation objectives.

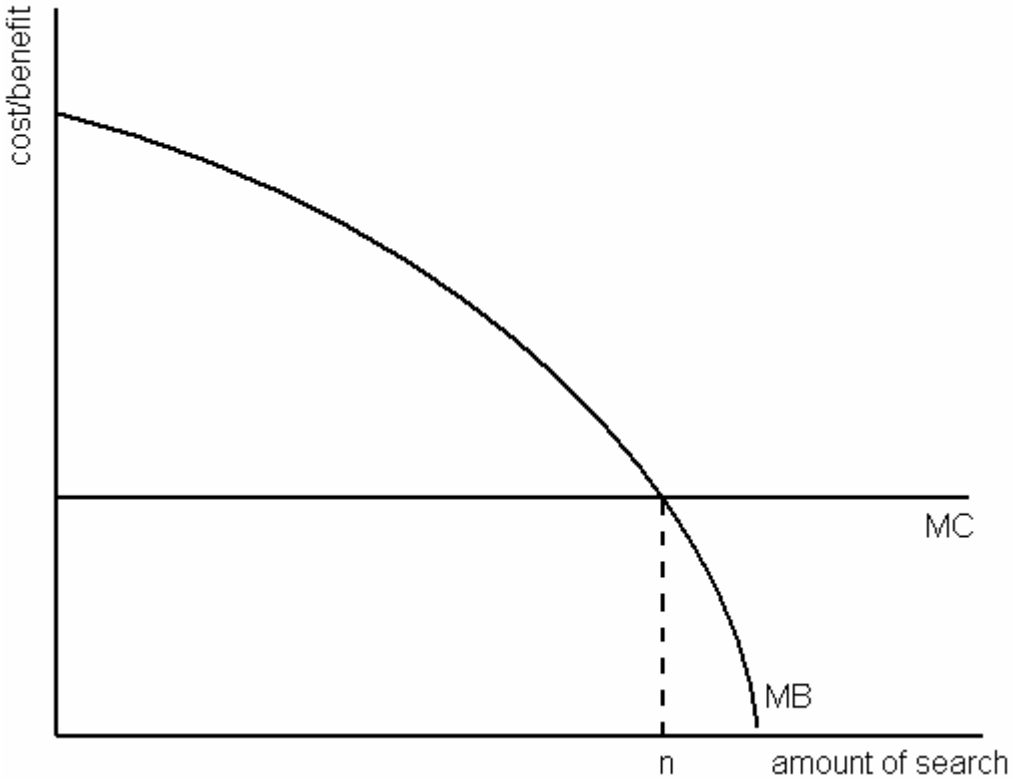
Information is, however, not free. The opportunity cost, defined as the *search cost*, of gathering information can be quite substantial. How many billions of dollars each year do firms not spend to assess the demand they are facing? It makes sense of course that such expenditure will be less or equal to the expected gain of having that information. A buyer searching for a lower price will keep on looking as long as he or she feels it is worth the time and effort. That is to say that the buyer will canvass an additional store if the cost of doing so is less than the expected reduction in price.

Search cost is relative and depends on the person conducting the search. Again it depends on the opportunity cost of time. Search cost can for some people be negative[2].

2.3.1 Buyer's Search

The basic rationale behind search is, as implied above, that search should be undertaken up to the point where the marginal cost of searching is equal to the marginal benefit of search. Figure 1 illustrates the decision on how much search to undertake is made.

Figure 1. The search decision



By setting the marginal cost (MC) equal to the marginal benefit (MB) of search, the optimal amount of search (n) is found. But how are these graphs derived?

[2] ”It is the thrill not the kill”

When searching for the lowest price, people use different search strategies (see for example McCall 1970, Mirman and Porter 1974, Rothschild 1974*a*, 1974*b* and Lippman and McCall 1976). First, think of a situation where the buyer has to decide how many stores to visit before visiting the first store. Even though the sequential search strategy is perhaps more relevant to our case, it is instrumental in showing how the decision whether to search or not, and how much, is made.

Let us assume that a risk neutral[3] buyer is facing a uniform distribution of prices between the lowest price p_L and the highest price p_H. The cost of search is c per additional search[4], meaning that marginal cost is constant at c, and n is the number of searches additional to the first. Knowing the distribution, the expected first price to be found is $p_E=(p_H + p_L)/2$. Now, since p_E is the expected first price to be offered, the expected reduction in price from searching one more time is $(p_E - p_L)/2$. The expected reduction in price is equivalent to the marginal benefit of search. Generally, the expected reduction in price resulting from the nth additional search is $(p_E - p_L)/2^n$. It follows that, *ex ante,* additional search will be undertaken as long as $(p_E - p_L)/2^n > c$. Putting $(p_E - p_L)/2^n = c$, we can solve for n as in Figure 1.

To exemplify the search decision, let us assume that c=1, p_L =5 and p_H =10 then p_E=7,5. The expected gain from an additional search is (7,5-5)/2=1,25 which is worthwhile since 1,25 > c. A second additional search will, *ex ante,* not be undertaken since ([7-(7,5-5)/2]-5)/2=(7,5-5)/2^2 =0,625 < c.

The main difference when search is sequential is that the consumer now knows the prices in the stores already visited. After each new price is observed, the expected reduction in price from another search is put against the cost of doing so. Implicitly, the cost of the searches already undertaken is sunk. Hence, even with identical preliminaries, the consumer can potentially search more than the one additional time predicted by the first search strategy.

The following can also be established:
As c reaches zero the buyer will search infinitely and will find the lowest price. If c is zero for all buyers there could be only one price-the lowest. When price dispersion, $p_H - p_L$, is greater more searches is undertaken. If p_H and p_L are both greater by the same proportion, i.e. so that,

[3]A risk averse individual would have a *lower* expected utility derived from the expected reduction in price (or in more familiar terminology, the expected increase in wealth) and is hence *less* inclined to search. The opposite applies to a risk loving individual.
[4]The cost of searching can be seen as a function, and can take many different forms. See Stiglitz(1989)

proportionately, $p_H - p_L$ remains the same but is absolutely greater, more additional search is done given that *c* remains constant5. More expensive goods will thus be researched more thoroughly, which is logical.

As noted above search cost is different from person to person. If the costs are prohibitively high no additional search can be expected. The situation is thus one where some buyers have more information than others-information is *asymmetrical*.

2.3.2 Demand Facing Sellers

An individual either buys a particular CD or not. Even if the CD is very cheap she does not buy two or three of this particular CD. Demand is therefore said to be *unitary*.

The demand upon which the *Law of The Single Price* is predicated is infinitely elastic. There is thus no scope for any one seller to offer a price that deviates from the no-profit equilibrium price. When buyers have different search cost the demand that a seller faces becomes less elastic.

Let us, with the help of an example, see how differing search costs cause the demand curve to become steeper. Assume that there are three kinds of consumers: those with prohibitively high search cost who visits only one seller, some with what can be called medium search cost who undertakes some additional search and a third type that searches infinitely[6]. We are interested in the consumers who enter the market, and therefore it is assumed that all consumers who search at least once buy. As will be seen, the price that they each pay is depending on their respective search cost. The consumer who searches infinitely will find the lowest price and will thus not buy from a store offering a higher price than the lowest price. The lowest price is called the *reservation price*, p_L, of the zero search cost consumer. Correspondingly, the medium search cost consumer's reservation price is p_M. Remember that preferences are identical. Therefore, *all* the high search consumer either buys or does not buy at all. If they do not buy they are not interesting to the case at hand. If we want, a fourth category, of non-buyers, can be thought of as those whose search cost is so high that they do not even visit one store. If the assumption of identical preferences is relaxed we could have a case where some buy and some do not, even if they have the same search cost. Later it is shown

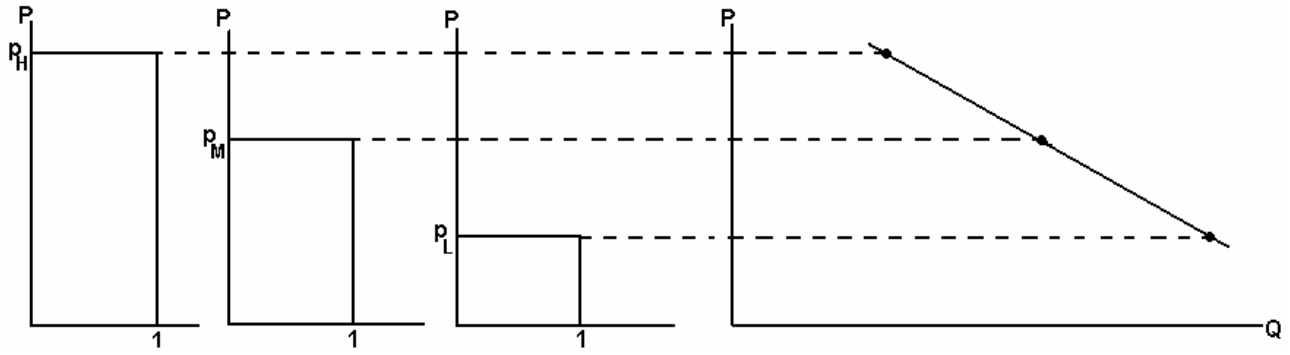
[5]Again this is only true for the risk neutral individual. It is not clear cut what the results would be for the risk averse but the risk loving indivual would undertake more search.

[6]This can be generalised to a continuum of buyers between the two extremes.

that non-buyers can also exist when a different search strategy is assumed. It follows that the individuals with prohibitively high search cost have a reservation price p_H such that they always buy.

Even though the argument, since we want to establish the demand that such price decisions are taken upon, becomes somewhat circular, assume that there are three different kinds of sellers-one offering a low price p_L, one setting a medium price p_M and one selling at the high price p_H[7]. The consumers search sequentially and do not canvass the same seller twice. All sellers are as likely to be visited. A seller offering p_L will catch all customers that enters no matter what their search costs are and no matter which number in the order of canvasses the seller is subject to. If p_M is stated all high search cost buyers will purchase, but only those which this is the first seller to canvass since they only canvass one. This price will also be accepted by the medium search cost consumers having this as their first or any subsequent seller to visit as long as no low-price seller is found first. A seller offering p_H will only sell to the high search cost individuals who happen to visit this seller as their first. It should be obvious that the lower the price offered, the higher the sales. In other words, each and every seller is facing a downward sloping demand curve. The situation is illustrated in Figure 2.

Figure 2. The demand curve is downward sloping when search cost differs



Each consumer in the three categories, with reservation prices p_H, p_M and p_L, demands one unit if the price is less or equal to their respective reservation prices. Since all sellers are as likely to be visited, based on the expected prices and reservation prices, the derived demand curve is downward sloping. The exact slope depends on the relative number of consumers in each reservation price category. Note that all consumers have identical preferences, the differences in reservation price is solely due to differing search cost.

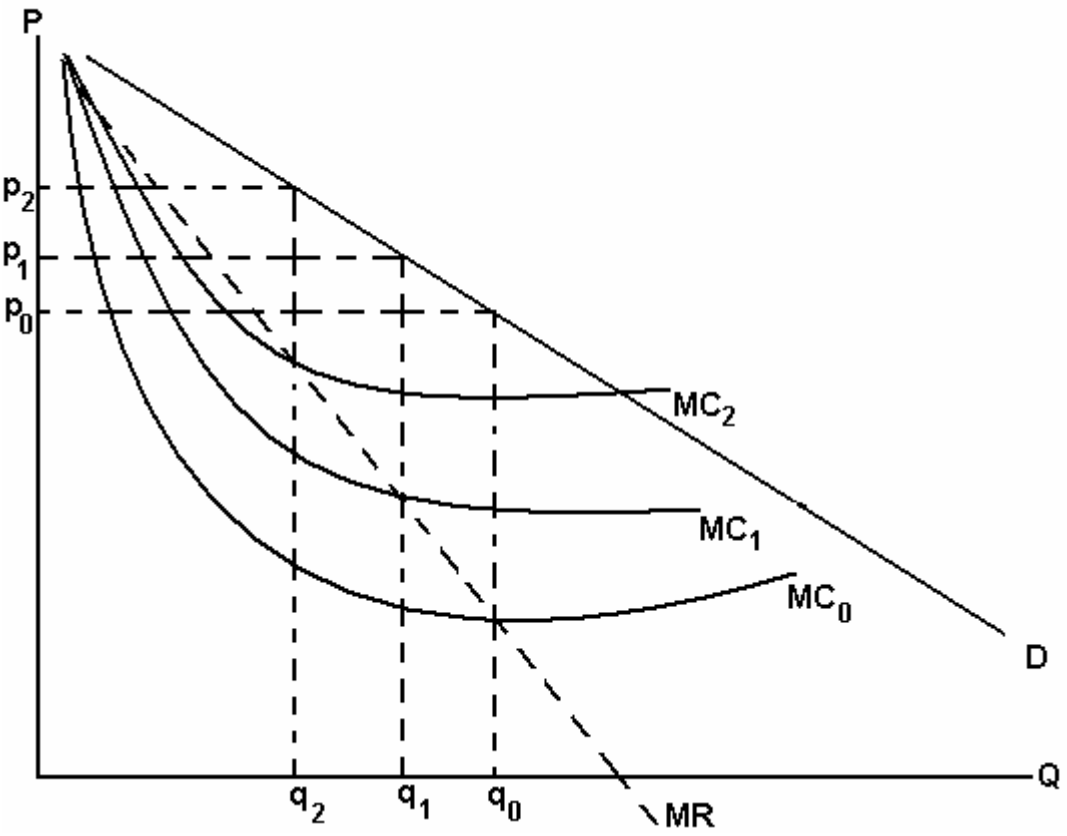
[7]This can also be generalised, to a continuum of sellers.

2.4 Price Dispersion

Even though the primary purpose of this demonstration is to show the intuition behind search cost causing price dispersion and what happens when search cost changes, and not an exercise of market equilibrium, it should be noted, as Reinganum (1979) argues, that imperfect information on its own is not sufficient to explain equilibrium price dispersion.

Firms are profit maximisers. In the traditional perfectly competitive equilibrium, all firms must have the same marginal cost. With the competitive price equalling marginal cost, an individual firm with higher marginal cost cannot survive in the long run. When each firm is facing a downward sloping demand curve, marginal costs can differ in the long run. Indeed, as Figure 3 shows, this causes price dispersion.

Figure 3. The profit-maximising price differs when the marginal cost differs



D is the demand facing an individual firm when information is imperfect. The profit-maximising price and quantity depends on marginal cost. With MC_0, the firm maximises profits by setting p_0 where $MC_0 = MR$ and the same for the other marginal cost curves.

Price dispersion can exist even if firms have the same technology. Stiglitz (1989) introduces search

cost in a perfectly competitive setting and outlines a few paradoxes. It is shown that the equilibrium, if it exists, is one with monopoly pricing. To avoid falling into these paradoxes it is crucial that the sellers in our context take into account the prices of other firms when setting their prices.

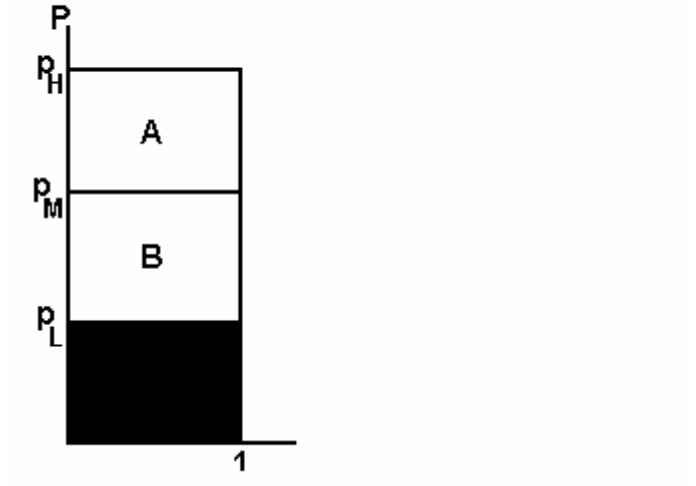
When there are a smaller number of firms the profit-maximising price for each firm depends much on the prices of the other firms. This is tantamount to a multiple peak profit function-the profit is maximised at several different prices, depending on the other firms' prices. Coming back to our example, the high price stores have few sales but a high margin whereas the low priced store fetch many sales at lower profits per sale and the medium priced somewhere between these two. For these three to charge different prices in equilibrium, causing price dispersion, all the peaks of the profit function must be at the same level of profits. If one firm would change its price, or say there is entry to the market, all others would have to adjust their prices to maximise profits. In (the Nash) equilibrium, all firms are in a position such that they would not gain from changing their prices.

Having covered perfect competition and oligopoly, even though not relevant to the CD-market, the case of monopoly should be mentioned. The monopolist of Salop (1977) takes advantage of the fact that buyers have different search costs by price discriminating, and thus creating price dispersion even when there is a single seller.

2.5 Welfare Implications

The consumer who searches until he finds the lowest price will have maximised the surplus available given the lowest price offered. In terms of Figure 4 the surplus is A+B. Again, this is independent of the number of visits and when, in the number of order, the consumer visits the store. It was assumed that the type of consumers with prohibitively high search cost have the reservation price p_H. If the first, and only, store that this consumer visits is a high priced store all the consumer surplus is appropriated by the store. It follows that if the first store is the medium (lowest) priced store the consumer gets a 'medium' (maximum) amount of surplus, B (A+B). The medium search cost type of consumer can of course end up with either the medium B or the maximum A+B surplus.

Figure 4. The surplus for an indivdual buyer when demand is unitary



Since preferences are identical and demand unitary, this surplus diagram is for the individual and makes the comparison between individual surpluses easy. The standard surplus diagram is more appropriate in the more general case where consumers have a continuum of different search cost. Given the price p_L, A+B is the maximum consumer surplus available.

Now, contrast the two search strategies. When the amount of search is determined a priori, the consumer with medium search undertakes one additional search. The consumer could thus be unlucky and canvass two high priced stores in a row. Given that the reservation price is p_M, the consumer does not buy at all. When the search is sequential, the decision on the amount of search is made as the search is undertaken. The cost of previous search is sunk, so no matter how many stores that have been canvassed before, when the best price found so far is p_H, the consumer always stand to gain from searching more.

Independent of search strategy, in comparison to the single price equilibrium associated with perfect competition or, perhaps more relevant to this case, Bertrand oligopoly where all consumers' surpluses are maximised, consumers lose out as a result of search cost and price dispersion.

2.6 Reducing Buyer Search Cost

Clearly, there is a welfare argument for reducing buyer search cost. Fortunately, sellers of course also see an advantage of being easily found. Borrowing an example from Stigler (1961), already in medieval times were merchants willing to pay tolls to get access to the market place. Today's multi-

billion advertising business is hard evidence of sellers willingness to be found[8]. Price promises such as 'we will lower our price if you can find the same good cheaper elsewhere' serve a purpose as to reducing search cost (although such promises can have negative effects on welfare by sustaining inefficiently high prices-see Salop 1986, Belton 1987, Edlin and Emch 1999 and Ingemarsson 2002)

Furthermore there are other agents in the market that have an incentive to facilitate buyers' search, such as consumer magazines. Not always do these other agents, however, work to increase information. Anecdotal evidence of such would for example be the traders wanting to stop transparency in the Eurobond market because they stood to gain from arbitrage (The Economist 1987).

The Internet is, one could say, a virtual reality parallel to the real life. People create identities, communicate and shop just as in real life. The largest difference is that none of those activities demands proximity between the people talking or between the seller and the buyer. Since the amount of time and effort, spent on going from one store to the next when undertaking price search, is determined much by proximity, the cost of searching is drastically reduced on the Internet. Also, as Bakos (1997) notes, ”[the Internet] may also reduce the ability of sellers to obscure their quoted prices (e.g., by including or excluding transportation costs, incentives, special promotions, financing costs, etc.).” As if searching oneself would not be easy enough, there is also a flora of consumer sites and shopping comparison engines[9].

Figure 5. When search cost is reduced the demand curve becomes more elastic

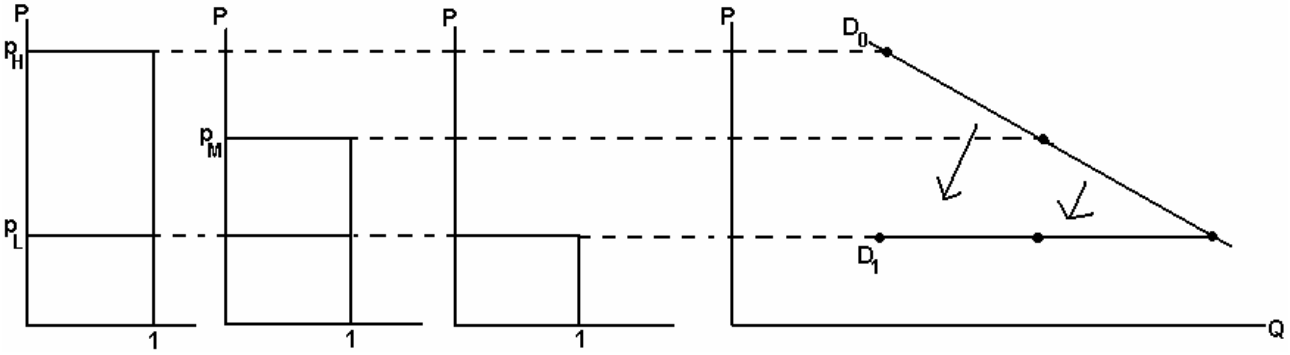


Figure 5 is Figure 2 superimposed with an illustration of what happens when search cost is eliminated. The earlier assumptions regarding preferences, store differentiation and transaction

[8]Advertising is used for many other purposes as well.
[9]Although their reliability have been questioned lately (see Metro January 3 2006 Martin Johnsen, Anders Heinkenfeldt, Filip Ekvall LUND)

costs are retained[10]. As all consumers now have perfect information about prices, since search is free, their common reservation price is p_L. There is no longer any scope for prices to deviate, and the dispersion of prices is thus reduced.

A second effect of reduced buyer search cost is that the average price is lower. In the extreme case illustrated in Figure 5, where there is only one price, p_L is of course also the average price. There are, however, two other mechanisms pushing prices downward on the Internet compared to the conventional market.

Firstly, competition is increased (Milgrom and Roberts 1982). There are no legal barriers to entry and nor are there prohibitively large sunk costs associated with CD e-commerce. Secondly, cost can be expected to be lower. There is no real need to actually store and display goods physically. The actual service of selling can be performed in a low-cost place and environment. That together with more specialisation of tasks, staff costs can be kept down. Because average cost is lower for sites, the competitive price is lower than in the conventional market.

In summary, it is hypothesised that

Prices of CD-albums are lower on the Internet than in the conventional market

and that

Prices of CD-albums are less dispersed in the Internet market than in the conventional market

[10]That demand would become infinitely elastic is perhaps not credible but, with these (perhaps also incredible) assumptions retained, it should be the case.

3. Methodology and Data

3.1 Methodology

On the 24th of November 2005, in rainy Malmö, prices of 12 CD-albums were collected from eight stores mainly situated in the city centre. A full day's work resulted in 81 prices. Three days earlier the same collection, of 75 prices, warm and comfortably placed in front of the computer had taken less than an hour.

Before the material was gathered several issues had to be addressed. In general, what good should be used, and in particular, which CD-albums should be included and how many? How is the market defined and what kind of sample could be representative of the population? How many, and which, stores and web-sites to include?

Previous studies have used a wide variety of goods. Bailey (1998), Brynjolfsson and Smith (2000), Clay *et al*.(2001), Ancarani and Shankar (2004) all used books. Its popularity is thanks to two factors. One, similar to the CD-album, each title is unique and a book is a book is a (...). Two, each book is labelled with a unique ISBN number which facilitates the collection. Arguably, CD-albums are more homogenous since books can come in different sizes and formats[11]. To mention a few other goods used: vitamins (Erevelles *et al*. 2001), insurance services (Brown and Goolsbee 2002), cars-within dealership (Morton *et al*. 2001), grocery products and cameras (Scholten and Smith 2002) and software (Bailey 1998).

Earlier studies made on CD-albums[12] have one thing in common that this one does not share-very large samples. This study's two samples totalling 156 observations are dwarfed by several thousands. As always in statistical sampling, the budget (in terms of time and money), is a major constraint. My manual sampling on the Internet can be compared to the bots-software that is programmed to automatically sample sites-used by for example Clay *et al*. (2001). This limitation applied to the number of different CD-albums to include in this study but not to the number of stores or websites to visit. All the same, restricted samples have been an issue for earlier studies too (Pan *et al*. 2004).

[11]Mp3's is a CD-album in a different format but the differences between, say, a hard back and a paper back are smaller than between a set of mp3's and a CD-album.
[12]See the Introduction

Following Lee and Gosain (2002) the sample consists of current hit-albums and old hit-albums, but with an extra category that can be called alternative music. The current hit-albums were at the time of collection on the top-ten of Sveriges Radio P3's list of hit-albums. To solidify the justification of the above given definition of the Swedish online market half of the albums have Swedish lyrics. Since such CDs are rarely sold to consumers outside Sweden, and are rarely sold by foreign websites to Swedes, the market is limited to Sweden. So, why not just include Swedish albums? The samples would be less representative of the population *CD-albums*, and more representative of the population *Swedish CD-albums*. Furthermore, the inclusion of foreign albums allows for comparison with Swedish albums. For example, if the Swedish online industry for foreign albums, on contrary to what have been assumed, includes foreign websites too, competition in that industry would be greater than in the Swedish album industry (which is still unaffected by foreign sites). Support for a lower price amongst the foreign albums in the online sample would thus discredit the given online market definition.

In order to test the hypotheses, two samples were required: one from the online market, and one from the conventional market. It was not strictly necessary that the same number of stores as websites were sampled, but a bias is introduced in the price dispersion if they differ. A larger sample is likely to include a few more extreme values, making the dispersion greater. As will be seen later, the severity of this depends on the measurement of dispersion. To be able to compare like with like, the albums in the two samples must obviously be the same.

As indicated, this study is purely cross-sectional. Even though the budget constrained this study to being of such nature, it does not necessarily mean that it suffers. Clay *et al*. (2001) did an intertemporal study (on books) and found that prices do not change much over time. By looking at stores' changes in rank in terms of price, over time, Baylis and Perloff (2002) investigated sites' usage of random pricing strategies. They found no support for such pricing strategies as the stores did not change much in rank over time.

When deciding which albums and which websites to include, availability was a limiting factor. To find the titles with good availability the following procedure was taken. The prices for a larger number of suitable titles were looked up from all identified Swedish websites selling new and unused CDs. A few of the sites were soon discarded for being either to niched or having a small selection. From the remaining nine, out of which two are niched on alternative music but still sells a few of the mainstream titles, the four most available titles in each category were chosen.

It remained to decide where the offline sampling would take place. To begin with, the stores had to be within an area small enough for them to be actual competitors. For example, because of the distance, a general record store in Malmö can hardly be seen as a competitor to one in Helsingborg. Furthermore, there had to be sufficiently many stores in the area in order to get a reasonably-sized sample. An Eniro-search[13] for stores selling CDs in Malmö resulted in nearly 20 hits. The suspicion was that not all would be good sampling spots. Some had names which implied niches or addresses very far out in the outskirts near the city border. *Ex post*, the actual number of stores that proved to have a good enough selection was eight, out of which one only had a few of the titles and one which had none of the alternative recordings.

Sampling from the Internet was relatively easy. All the sites had user-friendly search engines and clearly displayed prices. Collection from the conventional retailers was more open to problems. While contemplating writing letters, e-mails or making phone calls a few considerations arose. Firstly, perhaps most obvious, would the retailers actually be so helpful? With limitations already to the potential sample, non-respondence would be detrimental for the actual sample. Secondly, retailers would have to be convinced that they had no motive to understate their prices. Finally, the observations should be made at the same point in time. Of course a few days might not make any difference. But a time limit imposes a greater demand on respondents and could thus lead to reluctance to reply at all. With that in mind, collection was made by visiting the stores.

Having anticipated stressed and impatient shop-keepers, I first found the prices of the available albums browsing the store. The prices of the albums that were not on display I queried for. Before doing so I presented my reasons for sampling to convince the staff that I was not interested of which store offered the lowest price. To get fair observations of the actual price charged to the customer, I first asked if the store had a policy of ordering CDs not stocked, and whether they charged extra for it. None of the stores charge extra for orders, and thus they could give straight forward information on the price.

3.2 The data

The offline sample consists of 81 observations. Each album in the offline sample has between five and eight observations each, with the current hits being the most available. In contrast, the online

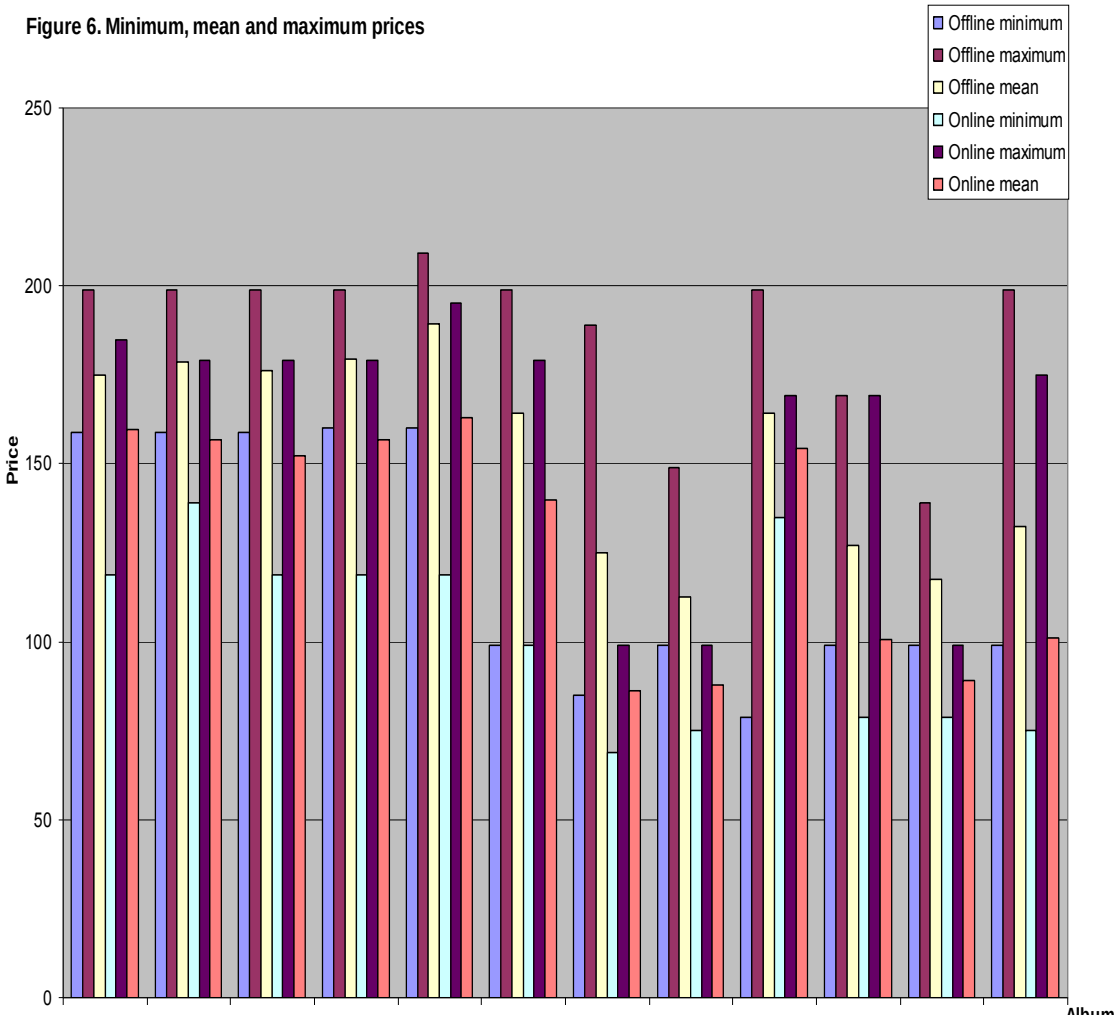
[13]www.eniro.se is the Swedish Yellow Pages online

sample of 75 observations varies between four and eight observations per album and with no specific pattern in the availability amongst the categories.

Common for both samples is a pricing pattern within most stores and websites. For instance, a seller typically has the same or very similar prices of the albums in one category. Unsurprisingly, the current hit albums are more expensive than the older ones. The price level in the alternative category varies more, both within the category and in relation to the current and old hits.

The highest price of each and every album, apart from one which tied, is found in the offline sample and the lowest, bar one, is found online. Mean prices were all lower in the Internet-sample.
Figure 6 provides an overview of the data.

Figure 6. Minimum, mean and maximum prices



4. Quantitative Analysis

4.1 Statistical Testing

The quantitative analysis, with the aim of drawing inferences about the population, includes a number of statistical tests. It could be argued that one test, whichever is the most appropriate one, should be enough. It is often said that one can prove any point with a sample as long as one uses the 'right' test-right in the sense that it proves the point. Partly to avoid including only those tests that, for instance, support the hypotheses of this paper, but also since the issue of which test is the most appropriate one is to some extent a subjective matter, a number of different test are presented.

As will be seen, some of the tests on price level and dispersion are using very small samples. This is when each album is treated separately, and thus using sample sizes of 12 to 16 observations. The reason for treating each album separately is that it has an intuitive appeal. After argued that CD-albums are homogenous goods-each title being unique and not close substitute to one another-why should all the sudden all observations be aggregated?

Some studies have done tests on the aggregated data, so that all the information in each of the two samples has been used in one single test. The approach comes with the benefit of overcoming some of the obstacles associated with the small samples-thanks to the Central Limit Theorem-but it also inherits more precise decision-rules. For example, if half of the albums, treated separately, show significantly lower dispersion on the Internet and half does not, what is the conclusion? The test on the aggregated data leads either to a rejection of the null, or not.

4.2 Price level

When testing the market definition the mean was used as a measure of the price level. This is of course the average price observed. But when hypothesising that prices are lower on the Internet, perhaps it is more relevant to look at the lowest prices observed. Although it will not be pursued further, note how the use of the lowest prices and mean prices corresponds to theory. The reduction in search cost would push the mean price down, but not necessarily the lowest price. Increased competition and lower costs would push both the mean and the lowest price down.

Brynjolfsson and Smith (2000) reports, in their longitudal study, that 84.6% of the time the lowest price of each album can be found online (p.570). No such comparison over time can be made here, but from Figure 6 everything points towards that the lowest price for each album is stated online. If half of the time the lowest price can be found online and half offline, the population proportion would be 0,5. A test on proportions, with the sample proportion 0.9, could be run. However, the test on proportions when the sample is small (12) is involved and beyond the scope of this paper.

Instead, tests were run on mean prices. Using all data, an equality of means t-test was run with the null hypothesis that the mean prices in the online and offline samples were equal. The alternative was that mean prices are lower on the Internet. The test relies on the variances being equal. The f-test for equal variances is included in the test summary of Table 1.
The variances were equal and the null of mean prices being equal was firmly rejected.

Table 1. Equality of means test, aggragated data

H_0: mean online = mean offline, H_1:mean online < mean offline	
t-stat: 4,3	p-value: 0,000
H_0: variance online = variance offline, H_1:variance online > variance offline	
f-stat: 1,03	p-value: 0,44

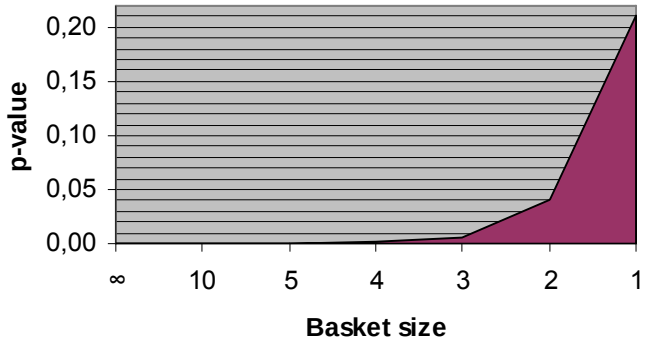
So far, to keep it nice and simple, the freight cost has been excluded from the online prices. More realistic, when making the price level comparisons, is of course if the freight is included in the price[14]. Note that since the market definition was tested using the deflated means of the mean differences, inclusion of freight would have made no difference.

The typical price for delivery is 29SEK. Since this is independent of the basket, the delivery charge must be divided by the number of CDs purchased at once, before added on to each album. Based on statistics from Forbes, Brynjolfsson and Smith (2000) used three as the typical number of CDs purchased at once. Whether that applies to the Swedish market too can be discussed, but, as will be seen, the result holds even for a smaller basket.

The equality of means test was run several times again, with different assumptions on the basket size. The resulting sensitivity analysis is presented in Figure 7.

[14]Some websites offer free delivery if the purchase amounts to over for example 900SEK. The average online price is 133. Therefore delivery is free if ordering seven or more CDs at once.

Figure 7. Equality of means test on aggragated data. P-value depending on basket size.



The null of equal means, in favour of the alternative of lower price on the Internet, was rejected at 1% (5%) significance level for a basket size of three (two) or more.

A third way of testing is with non-parametric tests. For example, Lee and Gosain (2002) used Wilcoxon rank test. However, the test makes adjustments for ties by excluding them. Unfortunately, the samples at hand do include many ties (much thanks to the sellers even pricing schemes) and as Körner, S (2003) argues, the test loses power when there are many ties.

When treating the albums separately, no patterns in the categories arise. Table 2 summarises the equality of means test for the individual albums.

Table 2. Equality of means test, separate albums
H_0: mean online = mean offline, H_1: mean online < mean offline
Number of rejected nulls, out of 12, at 10% and 5% significance level

Basket size		10%	5%
No delivery cost	(∞)	10	7
	5	7	5
	3	6	3
	2	4	2
	1	0	0

It would be fair to draw the inference that prices of CDs are lower on the internet, but looking at the entire market CDs in themselves are just one observation. It could therefore not be said that *prices* on the Internet are lower.

4.3 Price Dispersion

The spread of prices can be measured in different ways. The measurement to use depends partly on the data and partly on what the purpose of the description is.

Perhaps the most obvious one is the sample variance or standard deviation. Despite its statistical appeal, as a measure of price dispersion it suffers from two things. Firstly, all observations carry equal weight. This means that a distribution of ten prices could have a relatively low variance if, say, six of the observations are close to the mean. But if two out of the remaining four are relatively high and two are relatively low, this is arguably a dispersed price picture. Conversely, a distribution where five observations, a bit above the mean, are almost identical and the other five are identical but at a bit below the mean, could have relatively high variance- this despite the fact that the spread between the two camps of observations is not large relative to, for instance, the earlier example. The upside, on the other hand, of the observations carrying equal weight is that the variance is insensitive to extreme values and outliers. Secondly, it lacks intuitive appeal. Given the value of a particular variance, it is hard to describe in words what the value actually means.

Sorensen (2000) suggests two alternative measurements that overcome the outlined problems with using the variance. The range, defined as the difference between the highest and the lowest observation, clearly indicates the width of the spread. On the other hand, this measure is completely ignorant to the rest of the distribution and is also sensitive to extreme values and outliers. To reduce that sensitiveness the trimmed range, being the difference between the second highest and the second lowest observations, can be used.

All three measures become biased when the number of observations differs in the two samples. Because the chances of including the population's upper and lower values increase as the sample gets bigger, there is an inherit upward bias in the measurement of spread for the larger sample. Arguably, on the same basis as for extreme values, the bias in the variance is smaller compared to the other two.

As earlier suggested, price dispersion is proportional to the mean. Therefore, for the comparison of the price dispersion in the two industries to be unbiased, all three measurements need to be deflated by the mean price. For example, if now prices are higher in the conventional market, the absolute spread of prices would also be greater without necessarily indicating higher relative price dispersion.

Before turning to the results, let us briefly look at two other ways, not included here, to measure price dispersion. Clay *et al*. (2001) reports the dispersion as a deviation from a benchmark. In their case it is books with Amazon's prices as the benchmark. Of course, whose price to dub the benchmark is not so easy to decide upon, especially since one in each market is needed. In fact, Clay *et al*.'s study did not directly compare dispersion in the offline and online markets. Another method, that produces an indication of average price dispersion, is regression analysis. Lee and Gosain (2002) modelled the dependent variable *price dispersion* on each seller's deviations from the industry price. With, for an instance, 43 observations (albums) for each of the nine sites, the limitations for such regression analysis with this study's samples should be apparent.

As has already established, support was found for the equal variance F-test on the aggregate. Tested individually, where the variances were also deflated by the means, there was very little to suggest that price dispersion is lower on the Internet. When the values are deflated by the means, they again become dependent on the basket size. For example, if ordering two CDs at once from the Internet the mean online price increases by 14.5SEK for each album, and the deflated value becomes lower compared to when the freight is ignored. Invariably of the number of CDs purchased at once, three CDs had significantly lower variance online and the rest were not significantly different.

Ignoring freight, the deflated range proved to be lower in the offline market for seven of the albums. Eight of the twelve CDs had a deflated trimmed range that was lower on the Internet. Table 3 provides a summary of the ranges and trimmed ranges in percentages of respective mean (i.e. the deflated measures).

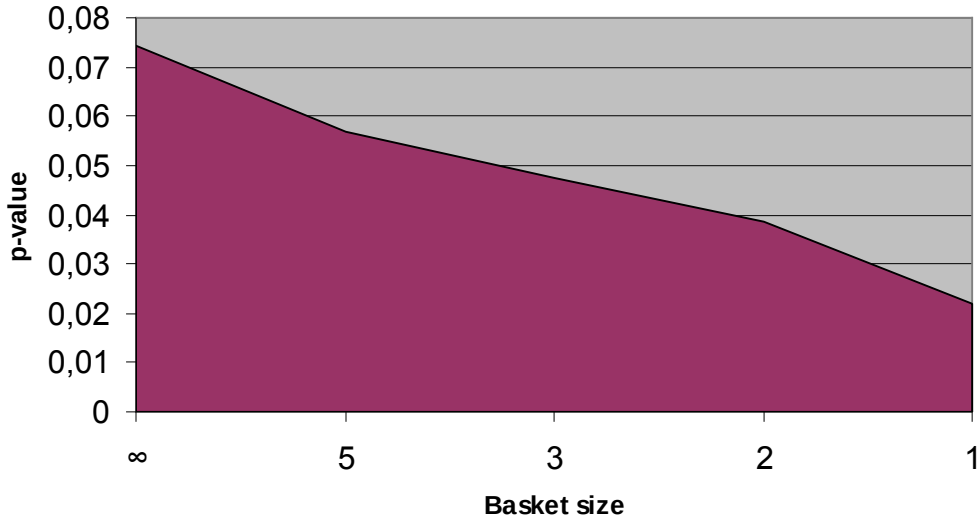
Table 3. Summary statistics of deflated range and trimmed range

Range (% of means)	Online	Offline
Minimum	0,22	0,22
Maxmimum	0,99	0,83
Mean	0,45	0,45
Trimmed range (% of means)		
Minimum	0,00	0,14
Maxmimum	0,46	0,69
Mean	0,20	0,28

The deflated range proved rather insensitive to the basket size. Only if ordering just a single album does it make a difference, when all albums have lower range online. The same thing applied to the deflated trimmed range, with the only difference that when ordering one CD, ten albums had a lower deflated trimmed range on the Internet.

To get a more decisive result, using all the information available, the online and offline range and trimmed range for each album were treated as matched samples. The normal t-test with the null hypothesis of no difference, i.e. no difference in range between the online and offline sample, was run with different assumptions of the basket size. The mean of (range offline-range online) was positive no matter assumptions or measurement, why both alternative hypotheses were that respective measure is lower on the Internet. When freight was excluded, neither of the measures were significantly lower on the Internet at the 5 % significance level, even though the trimmed range had a p-value as low as 0,07. When freight was included the picture was different, not so much for the range but for the trimmed range. Figure 8 shows how the p-value changes with basket size for the matched samples test on trimmed range.

Figure 8. Trimmed Range: Matched sample test with a null of equal means. Sensitivity analysis of p-value depending on basket size.



4.4 Market Definition

Lastly, the definition of the market was tested. The reasoning goes as follows. The conventional market is rather well defined, with an even competitive pressure on prices across Swedish and foreign albums. If the Swedish online market includes also foreign websites, the competitive pressure should be higher on foreign albums and thus prices lower for foreign albums. If the industry definition is indeed good, such differences in prices should not be significant.

For two reasons, the comparison between Swedish and foreign album has to be made on the relative

price-difference online and offline, and not the absolute online prices. One, both the retailers and the sites could be expected to pay higher prices for foreign albums, since it involves more freight and taxes. This would immediately counter-weigh any, compared to the Swedish albums, additional competitive downward pressure on the absolute online prices for foreign album. Since relative prices are affected by freight and taxes in an even way, these are controlled for when the relative prices are compared. Second, the samples split into two, Swedish and foreign albums, are asymmetrical what regards mean prices and distribution of albums in the categories. For example, the Swedish albums consist of one current hit, two old hits and three alternative albums. Again, such differences are controlled for when relative prices are compared.

The differences between the online and offline means for each album were calculated. The albums were then split into two groups-Swedish and Foreign. The mean of the differences in means and its variance, deflated by each group's absolute mean, for each group were used to conduct an equality of means t-test. The null-hypothesis that the deflated differences in means were equal could not be rejected with a p-value of 1. Note that when failing to reject the null, the risk of committing a type II error is uncontrolled for. So, used in this way the equality of means test is weaker, compared to when used for the price level. Then, in the event of a rejected null, the corresponding type I error is indeed controlled for by the significance level.

5. Discussion and Further Research

Let it be given that the hypothesis of lower prices online was supported and that the hypothesis of lower price dispersion was not supported. One possible conclusion, in correspondence to theory, is then that the cost and competition argument for lower prices are valid, making the demand curve shift downward. But since price dispersion is not lower, the demand curve has not flattened (compare with Figure 5). A downward shift in the demand curve would also be in line with the finding that the lowest prices are found on the Internet but the average price is not always statistically significantly lower.

The immediate question is what can have caused price dispersion on the Internet. This section will look at some possible explanations. A few directly questions the various assumptions about information, transaction costs, preferences and consumer search, whereas one proposes a different model.

The delivery charge of online-purchases 'locks in' consumers. This is thanks to the high *switching cost* that is created. Imagine the following scenario. Someone is searching to buy three CDs. After a complete search, one site sells two of the CDs to the lowest price for respective CD. The third CD can be found cheaper at a second site, where the two first CDs are more expensive. Was the consumer to buy all three CDs at their lowest price, the marginal price of the third CD from the cheaper site is rather greater than the listed price, and most probably higher than the price that the first site is offering. The reasons is of course because the freight has to be paid twice, and the second time it is added to the marginal price of the third CD[15]. The solution for the consumer is to buy from one site, the site that offers the lowest total price. For the consumer to be induced to buy from both sites (i.e. do a sort of *switch*), the price differential of the single-purchase total price and the twice-purchase total price would have to be higher than the freight cost.

What this means is that firms compete on the total price of a typical purchase. This gives scope for individual CDs to vary in price as long as the total price is the same. If a proxy for the typical shopping basket of CDs could be identified, the comparison between online and offline total-prices dispersion could be made. Implicitly, scope for particular pricing schemes, such as loss-leader pricing to attract customers, is also widened by high switching costs. Comparisons could be made

[15]One can reason the other way around too, that the marginal prices of the first two CDs are greater than the listed price.

directly to the conventional market.

Although it has been assumed that the consumers have different search costs, it has implicitly been assumed that there is no difference between the online and the offline consumers. In reality, they need not be the same. Brynjolfsson and Smith (2000) points out that ”it may be that the types of consumer who shop on the Internet have higher wages or are busier and therefore have systematically higher cost of time than those who shop in conventional stores.(p.569) The search cost induced price dispersion could thus still be relevant on the Internet.

The credibility of all consumers having knowledge about the distribution of prices, which underpins the search strategies upon which the hypothesis is derived, could be questioned. Stigler (1961) makes the example of how tourists, who have little knowledge about price distributions, routinely are charged higher prices in economies where haggling is still the norm. It could be that an individual who has got a good perception of the price distribution in the conventional market, enters the, to her, unfamiliar Internet market. As a 'tourist', since haggling is not allowed, chances are high that the first site offers a price that is low, relative to what can be expected from the conventional market. With the conventional markets distribution in mind, it could be that this price induces a purchase. In this scenario, sites' advertising, on for example television, could be of great relevance since the first store gets the sale. A possible equilibrium could be one where sites who advertise their presence sell mainly to 'tourists', at higher prices to make up for the advertising cost, and others who rely on the consumers to search until they find their, lower, prices. Such equilibrium is consistent with a model by Salop and Stiglitz (1977).

In a model by Pereira (2004), the presence of higher price dispersion on the web finds a possible explanation. The model is relatively simple, there are low search cost consumers with low reservation prices and high search consumers with high reservation prices. The firms differ in that one has lower marginal cost and can sell at its monopoly price at a low price, and one has higher marginal cost. Depending on the proportion of low cost consumers in the market, or their reservation price, two different equilibria can exist. The low marginal cost firm always charge the low price, since it is their monopoly price. If the proportion of low search cost consumers is high, or their reservation prices are high, the high cost firm will want to compete for them and will thus charge the low price too. This because, as Pereira describes it, ”the volume of sales effect dominates the per consumer profit effect” (p.67). The dispersion is thus very small in this competing equilibrium. But if the proportion of low search cost consumers is small or their reservation price is very low, it will pay the high marginal cost firm to charge only the high price and hence the prices

are distributed.

The model thus predicts that if the Internet market is characterised by a low proportion of low search consumers or their reservation price is very low and the opposite applies to the conventional market, price dispersion would be larger on the Internet despite lower search cost.

6. Conclusion

Central to this paper is the effect of imperfect information, caused by search cost, on prices. The intuition suggests that as search cost is reduced, competition should increase and prices fall. With the help of basic theory, it is also shown that firms' ability to deviate from a single price, causing price dispersion, is reduced.

A methodological sampling procedure produces an adequate, but only just, material that suffers somewhat from its limited size. A series of statistical test, all with different benefits and disadvantages, draws up a somewhat inconclusive picture.

As promised, it will be left for the reader to ponder the evidence and make up his or her mind, as to whether the hypothesis of lower price dispersion has been supported by the somewhat inconclusive evidence. In summary though, the market seems well defined and there is little or no evidence against the price level being lower on the Internet. The minimum prices are more often found online, whereas the evidence for the average prices is not altogether clear cut. For what price dispersion is concerned, there is some support of the hypothesis, some evidence of price dispersion being the same in the two markets and no evidence of it being greater online.

The *Law of the Single* Price does not hold even when search cost is practically eliminated. On offer is a row of explanations, all with their own merits. Basic assumptions regarding transaction costs and preferences could be unrealistic. The material is sampled in a way that leaves many alternative hypotheses to be tested on it and the previous section mentioned a few.

7. References

Ancarani, F., and Shankar, V. (2004). Price Levels and Price Dispersion Within and Across Multiple Retailer Types: Further Evidence and Extension, *Journal of Academy of Marketing Science, volume 32, no. 2, pp. 176–187*

Bailey, J. (1998) Intermediation and Electronic Markets: Aggregation and Pricing in Internet Commerce.
Unpublished doctoral dissertation, Technology, Management and Policy, Massachusetts Institute of Technology, Cambridge, MA.

Bakos, Y. (1997) Reducing Buyer Search Costs: Implications for Electronic Marketplaces, *Management Science, volume 43, no. 12, pp. 1676-1692*

Baylis, K., and Perloff, J.M. (2002). Price Dispersion on the Internet: Good Firms and Bad Firms, *Review of Industry*
Organization, volume 21, pp. 305–324

Belton, T.M. (1987) A Model of Duopoly and Meeting or Beating Competition, *International Journal of Industrial Organization volume 5, pp. 399-417*

Brown, J.R., and Goolsbee, A. (2002) Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry, *Journal of Political Economy, volume 110, no.3, pp. 481–507*

Brynjolfsson, E., and Smith, M. (2000). Frictionless Commerce? A Comparison of Internet and Conventional Retailers. *Management Science, volume 46, no.4, pp. 563–585*

Clay, K., Krishnan, R. and Wolff, E. (2001) Price Dispersion on the Web: Evidence From the Online Book Industry, *The Journal of Industrial Economics, volume XLIX, no.4, December*

Economist (1987) *No Computer for Euro-Bonds*, 30 May

Edlin, A.S. and Emch E.R. (1999) The Welfare Losses from Price-Matching Policies, *Journal of Industrial Economics, volume XLVII, pp. 145-167*

Erevelles, S., Rolland, E., and Srinivasan, S. (2001) Are Prices Really Lower on the Internet?: An Analysis of the Vitamin Industry (Working Paper). Riverside: University of California

Friberg, R., Ganslandt, M and Sandström, M.(2000) E-commerce and Prices-Theory and Evidence, *SSE/EFI working paper series in Economics and Finance no. 389*

Ingemarsson, M. (2002) Prisanpassning via Internet - en Studie av Konsumentprodukter pp.43-60 in *Om Internet och Marknader i Förändring*, Konkurrensverket Mars 2002

Körner, S. and Wahlgren, L. (2002) *Praktisk Statistik*, 3rd edition, Lund

Lee, Z. and Gosain, S. (2002) A Longitudinal Price Comparison for Music CDs in Electronic and Brick-and-Mortar Markets: Pricing Strategies in Emergent Electronic Commerce, *Journal of Business Strategies, volume 19, no. 1,*
pp. 55-72

Lippman, S.A, and McCall, J.J. (1976) The Economics of Job Search:A Survey, *Economic Inquiry, volume 12, pp. 347-68*

McCall, J.J. (1970) Economics of Information and Job Search, *Quarterly Journal of Economics, volume 84, pp.113-26*

Milgrom, P. and Roberts, J. (1982) Limit Pricing and Entry Under Incomplete Information, *Econometrica, volume 50, pp. 443-460*

Mirman, L.J., and Porter, W..R (1974) A Microeconomic Model of the Labor Market under Uncertainty, *Economic Inquiry, volume 12, pp. 135-45*

Morton, F.S., Zettelmeyer, F., and Silva-Risso, J. (2001) Internet Car Retailing, *Journal of Industrial Economics, volume 49, no. 4, pp. 501–519.*

Pan, X., Ratchford, B.T., and Shankar, V. (2004) Price Dispersion on the Internet: A Review and Directions for Future
Research, *Journal of Interactive Marketing, volume 18, no. 4, Autumn*

Pereira, P. (2004) Do Lower Search Costs Reduce Prices and Price Dispersion?, *Information Economics and Policy, volume 17, pp. 61–72*

Salop, S. (1977) The Noisy Monopolist: Imperfect Information, Price Dispersion and Price Discrimination, *Review of Economic Studies, volume 44, pp. 393-406*

Reinganum, J.F. (1979) A Simple Model of Equilibrium Price Dispersion, *Journal of Political Economy, volume 87, no. 41, pp. 851-858*

Rotschild, M. (1974*a*) A Two-Armed Bandit Theory of Market Pricing, *Journal of Economic Theory, volume 9, pp. 185-202*

----------(1974*b*) Searching for the Lowest Price When the Distribution of Prices is Unknown, *Journal of Political Economy, volume 82, no. 4, pp. 689-711*

Salop, S. (1986) Practices that (Credibly) Facilitate Oligopoly Coordination, in *New Developments in the Analysis of Market Structure*: Editors: Stiglitz, J. and Mathewson, F. MIT Press

Salop, S. and Stiglitz, J.E (1977) Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion, *Review of Economic Studies, volume 44, pp. 493-510*

Scholten, P.A., and Smith, A. (2002) Price Dispersion Then and Now: Evidence from Retail and E-Tail Markets, *Advances in Microeconomics: Economics of the Internet and e-Commerce, volume 11, pp. 63–88*

Sorensen, A.T. (2000) Equilibrium prices dispersion in retail market for prescription drugs, *Journal of Political Economy, volume 108, no. 4, pp. 833-850*

Stigler, G.J. (1961) The Economics of Information, *Journal of Political Economy, volume 69, no. 3, pp. 213-25*

Stiglitz, J.E (1989) Imperfect Information in the Product Market, pp.767-847, chapter 13 in *Handbook of Industrial Organization volume 1*, Editors:
Schmalensee, R. and Willig, R., Elsevier, Amsterdam