

Supply in bits? Assessing cultural diversity in the record industry
in the Netherlands.

Master thesis
Cultural Economics & Cultural Entrepreneurship

J.W.F. Keizer
janwillemkeizer@student.eur.nl
295352

Dr. C.W. Handke
E. Dekker MSc

Erasmus University Rotterdam - *Erasmus School of History, Culture and
Communication*

Table of Contents

Abstract	6
Acknowledgments	6
Introduction	7
2. Theoretical framework	8
2.1 Review of the literature	8
<i>2.1.1. Basis principles literature review</i>	<i>8</i>
2.2. Cultural Diversity in world of politics	9
<i>2.2.1. Protecting cultural expressions</i>	<i>9</i>
<i>2.2.2. A smart move</i>	<i>10</i>
<i>2.2.3. "... attention must be paid to the diversity of the supply"</i>	<i>10</i>
2.3 Cultural diversity in the academic literature	11
<i>2.3.1. A very short history of diversity research</i>	<i>11</i>
<i>2.3.2. Cultural Diversity defined</i>	<i>11</i>
<i>2.3.3. Towards a multi-dimensional understanding: variety, balance, disparity</i>	<i>12</i>
<i>2.3.4. The second dimension.</i>	<i>13</i>
<i>2.3.5. The third dimension</i>	<i>13</i>
2.4. Previous research on diversity in the music industry.	14
<i>2.4.1. Theoretical models</i>	<i>14</i>
<i>2.4.2. Empirical studies</i>	<i>15</i>
<i>2.4.3. The relation between diversity and industry concentration</i>	<i>15</i>
2.5. How should cultural diversity be measured?	17
<i>2.5.1. The Shannon-Wiener index, Stirling's Hsw and Hbfp</i>	<i>17</i>
2.6. How is cultural diversity measured in the record industry?	18
2.7 The relation between Digitisation & diversity: Theory.	21

2.7.1. Nelson's product searches.	21
<i>2.7.2. Information goods: search & experience</i>	<i>21</i>
<i>2.7.3. Additions to Nelson's theory.</i>	<i>22</i>
<i>2.7.4. The impact of Internet on search costs</i>	<i>22</i>
<i>2.7.5. Are albums experience goods or search goods?</i>	<i>22</i>
<i>2.7.7. Sampling</i>	<i>23</i>
2.8. Frictionless markets	23
2.9. The Long Tail	24
2.10. Digitisation leads to more diversity.	24
3. Method	25
3.1. Data	25
3.2. Dependent variables: diversity in the charts.	25
<i>3.2.1. Why do I use the Album top 100 as data source?</i>	<i>25</i>
<i>3.2.2. The Short History of the Album Top 100.</i>	<i>26</i>
<i>3.2.3. A proxy for diversity</i>	<i>26</i>
<i>3.2.4. How I processed the charts data into a workable data set</i>	<i>26</i>
<i>3.2.5. Missing data chart data</i>	<i>27</i>
3.3. Independent variable: Making digitisation operational.	27
<i>3.4. Research strategy</i>	<i>28</i>
<i>3.5. Defining two periods</i>	<i>29</i>
3.6. Criteria in social research.	30
<i>3.6.1. Reliability and replicability.</i>	<i>30</i>
3.6.2. Validity	30
Results	32
4.1. Checking for trends	34
<i>4.1.1. Pre period: 1979 - 1993</i>	<i>34</i>

4.1.2. <i>The post period: 1994 - 2009</i>	34
4.3. Correlation analysis	35
4.3.1. <i>Checking normality</i>	36
4.3.2. <i>A first exploration of the data</i>	36
4.3.3. <i>Pearson's r.</i>	37
4.3.4. <i>Using R² for interpretation</i>	38
4.3.5. <i>Correlation Digitisation and Diversity in the period 2000 - 2009</i>	38
4.4. <i>Summary</i>	39
Conclusion	40
5.1 <i>Suggestions for further research.</i>	40
References	42
Appendix 1: SPSS output	45
Appendix 2: data set	51

Abstract

This paper investigated the relation between digitisation and diversity in the record industry in the Netherlands. The recent discussion on cultural diversity has its roots in the political sphere where members of the WTO have opposite interests regarding free trade for cultural goods. The concept of cultural diversity is equivocal and has attracted ample attention in the academic literature. I surveyed the literature discussion cultural diversity in the record industry that is mainly concerned with the relation of concentration and diversity. Drawing from theory of Nelson (1970), the “frictionless market hypothesis” by Brynjolfsson & Smith (2000) and the Long tail of Anderson (2004) I tested whether digitisation my hypothesis - *Digitisation leads to more diversity in the supply of records in the Netherlands* - and found that both variables indeed shared a strong relation.

Acknowledgments

I am intensely indebted to parents, friends and my supervisor. I would like to thank my parents for their unconditional love, trust and financial means to reach the end of my academic career. My roommate and fellow student Lisa for inspiring me. My roommate Patrick for continuously annoying me through the process. Also, Michel Admiral at GfK for providing the necessary data.

I greatly appreciated the help from my supervisor, Christian Handke. Especially his patience and the useful insightful comments were encouraging.

1. Introduction

Systematic studies of current technological change – sometimes referred to as ‘digitisation’ – are rare. It remains unclear, for example, on what level the diffusion of digital ICT has its most important effects. (Handke, 2010: 149)

This was one of the clues for me when I was looking for a topic for this paper. This I combined with a topic from one of the Cultural Industries seminars: cultural diversity. Maybe it was because I gave a presentation about cultural diversity that I choose to write my paper about the connection of digitisation and diversity.

Either way, this paper studies whether digitisation influences the cultural diversity in the record industry. Every student of the Cultural Economics & Cultural Entrepreneurship master has its own field of interest. Mine is the record industry.

There is a vast amount of literature which concentrates on the interrelation of industry concentration and diversity. There is not much written about my topic. With this paper I try to fill this gap in the literature a bit further.

The following research will be a deductive study into the development of diversity in the record industry. From theory I formulated this hypothesis: *“Digitisation leads to more diversity in the supply of records in the Netherlands.”* This statement is tested with my collected data.

As it turns out, digitisation shares a strong relationship with the cultural diversity in the market for sound recording in the Netherlands.

2. Theoretical framework

2.1 Review of the literature

The purpose of this review is to explore the literature on the topic of this paper: cultural diversity. I will identify relevant concepts and theories. This review is divided into different parts with its own emphasis.

The first part sets out how I did my literature research. The second part will deal with the political context of cultural diversity. The third part deals with studies conducted on cultural diversity in the academic literature and discusses definitions. The following part considers how to measure diversity. Part six is concerned with cultural diversity in the music industry. The last part deals with theory. In this part I will present relevant theories which are the foundations for my empirical research. I will discuss the product search theory of Nelson (1970), the “frictionless market hypothesis” by Brynjolfsson & Smith (2000) and the Long tail of Anderson (2004).

2.1.1. Basis principles literature review

There are some criteria I used to select the articles for this paper. First of all, the article had to be academic and preferably from well established journals with a good reputation. Furthermore, I preferred the articles which were quoted a reasonable amount of times. All the articles which I read for this study are found with several academic search engines. Among these are Google Scholar, sEURch and ISI web of knowledge. The books are tracked down through the search engine of the Universiteitsbibliotheek Erasmus University Rotterdam and again, sEURch.

2.2. Cultural Diversity in world of politics

The term “cultural diversity” denotes several meanings and it has changed through time. In the present debate about cultural diversity the term generally has two meanings. The adjective “cultural” refers to culture in the broadest sense of the word, as “the set of distinctive spiritual, material, intellectual, and emotional features of society or a social group. In addition to art and literature, it encompasses lifestyles, basic human rights, value systems, traditions, and beliefs.” (definition adopted the Mexico City MONDIACULT World Conference in 1982)¹ Yet cultural diversity also indicates “the manifold ways in which the cultures of groups and societies find expression” (Art. 4, No. 2, UNESCO Convention on Cultural Diversity.²) This meaning is a more “commoditized” approach - cultural diversity as a limited resource - rather than a definition which takes into account the differences between cultures - an anthropological approach. (Curtis, 2006)

“‘[C]ultural diversity’ in international cultural politics is the standard-bearer of a campaign to exclude cultural goods and services from global free trade rules.” The main strategic purpose of cultural diversity is to legitimize protective measures for nationally produced cultural goods. (Isar, 2006) This is a rather bold statement, but there is at least some truth in it. I will explain this claim in the following paragraphs.

In 1995 the World Trade Organization (WTO) was established. This was successor of the General Agreement on Tariffs and Trade (GATT). The main objective for the WTO is to “...open trade for the benefit of all.”³ The United States did not sign the GATT treaty, but they did become party to the WTO treaty. As the United States, “in which diversity refers to the free flow of ideas and expressions” (Curtis, 2006: 61) made an effort to apply the free trade rules to all cultural goods, this attempt was countered heavily by certain other member states. (Acheson & Maule, 2004; Isar, 2006).

Especially France and Canada, among others, tried to challenge the US commitment to free trade of cultural goods. The countries had successfully negotiated an “*exception culturelle*” for these goods under the GATT treaty during the Uruguay Round (1986-1994) and were determined to incorporate these agreements into the new statute of the WTO.

2.2.1. Protecting cultural expressions

From the various reasons why a country would not follow strict free trade rules, the protection of the domestic language is the most prevalent. Obvious examples here are French speaking Canadians in Canada and Catalans in Spain. Also many countries fear English-imperialism as Hollywood dictates the cinema and programs on the television. Yet another reason for protectionism is the prevalence of the so-called “diversity gaps”. (Acheson & Maule, 2004) This means that the actual state of supply in cultural products does not level the desired state.

1 <http://www.diversite-culturelle.qc.ca/fileadmin/documents/pdf/update0303.pdf>, accessed on August 1, 2011

2 <http://unesdoc.unesco.org/images/0014/001429/142919e.pdf>, accessed on May 17, 2011

3 http://www.wto.org/english/thewto_e/whatis_e/wto_dg_stat_e.htm, accessed on August 1, 2011

Although this requirement is very arbitrary, one can imagine what this would mean in practice. Exemptions on the trade rules are welcome in case of a skewed supply in cultural expressions of a country. There could be a shortage of plays in the domestic language or too less knowledge of traditional music. For a more comprehensive discussion of justifications to exclude cultural goods in free trade see Mas-Colell (1999) and Bala (2005).

2.2.2. A smart move

“[T]hrough a subtle process of semantic sleight of hand...” (Curtis, 2006: 374) the proponents of the *exception culturelle* tabled the concept of cultural diversity at the negotiations. A smart move, because they were now able to tap into a larger reservoir of cultural commitments. See for example what the United Nations Universal Declaration on Human Rights (Universal Declaration) states in article 27:

“Everyone has the right freely to participate in the cultural life of the community, to enjoy the arts, and to share in scientific advancement and its benefits. Everyone has the right to the protection of the moral and material interests resulting from any scientific, literary or artistic production of which he is the author.” These provisions are mirrored by the language used by the International Covenant on Economic, Social and Cultural Rights (CESCR), which observes “the right of everyone: (a) To take part in cultural life; (b) To enjoy the benefits of scientific progress and its applications; and (c) To benefit from the protection of the moral and material interests resulting from any scientific, literary, or artistic production of which he is the author.” The CESCR is an international binding legal instrument. (Curtis, 2006)

So, according to Isar (2006: 372), it seems that from this perspective cultural diversity is seen as “a strategic notion...” that “...tend[s] to favor ‘billiard ball’ representations of cultures as neatly bounded wholes - [...] – hence mainly to be ‘protected’ or ‘preserved’”. After all, every government in the world which works under the conditions of the Universal Declaration is bound to preserve these rights to culture for its citizens. So goes the argument. Cultural diversity has thus to be protected, according to the proponents of protectionist measures on cultural goods.

2.2.3. “... attention must be paid to the diversity of the supply”

To reach this goal two independent organisations were created: the International Network of Cultural Policy (INCP) and the International Network for Cultural Diversity (INCD). The INCP is a forum for the ministers responsible for culture to develop strategies to promote cultural diversity.⁴ The Netherlands is member as well. Their aim is to gear both globalism and the preservation of national cultural heritage to an optimal state. The INCP was founded by Canadian Heritage and acts as a “worldwide network working to counter the adverse affects of globalization on world cultures”⁵ It is an umbrella group for artists, cultural activists and cultural NGO’s from other countries. The French speaking countries also held conferences to draft the New International Instrument on Cultural Diversity (NICD). Among others, representatives of the cultural industries in Canada were active in the process of the creating of this NICD.

⁴ http://www.incp-ripc.org/about/index_e.shtml Accessed on May 19, 2011.

⁵ <http://www.incd.net/about.html> Accessed on May 20, 2011

Eventually, these events led to the creation of a convention on protecting cultural diversity in 2005. Originally, members wanted to write a legal binding document as opposed to much of the other treaties and agreements which were mere awareness-raising documents. (Acheson & Maule, 2004) This setup failed, but at least the UNESCO member states agreed to the following article: “In the face of present-day economic and technological change, opening up vast prospects for creation and innovation, particular attention must be paid to the diversity of the supply of creative work, to due recognition of the rights of authors and artists and to the specificity of cultural goods and services which, as vectors of identity, values and meaning, must not be treated as mere commodities or consumer goods.” (article 8, Universal Declaration on Cultural Diversity)⁶ This led Isar (2006) to conclude that cultural diversity has transformed from a state of the human condition into a normative meta-narrative. The Convention on the Protection and Promotion of the Diversity of Cultural Expressions made it necessary to further define cultural diversity and the diversity of cultural expressions in objective measures to evaluate implemented policies. The developments in the political sphere around cultural diversity attracted a reasonable amount of scholarly attention. I will now discuss this literature.

2.3 Cultural diversity in the academic literature

Let us not beat about the bush: cultural diversity is a multi dimensional, equivocal and strange concept. (Ranaivoson, 2007; 2010; Benhamou & Peltier, 2007) There is not a standard definition in the literature.

2.3.1. A very short history of diversity research

To get a thorough understanding of cultural diversity it is instructive to go back to where it all began. Practically most research done on diversity of any kind has its roots in the study on biodiversity. (Flôres, 2006). The issues of how and whether to study variety, balance and disparity were already prevalent in biology in the 1940s. Much biodiversity research is based on Simpson’s (1949) index that measures the different number of types (variety) and the relative proportions of these (balance). This index is also known as the Hirschman-Herfindahl index used in industrial economics. (Ranaivoson, 2007) Also, there is another branch in biodiversity research which puts more attention to disparity. (Weitzman, 1992)

Beside in biology, diversity has had ample attention in economics for many years. (Ranaivoson, 2007) She discussed in length the theoretical model of Hotelling (1929) which was the first to model product diversity in economics.

2.3.2. Cultural Diversity defined

Diversity is variation and therefore often defined as the sheer multiplicity of types. Researchers often only look at this one sole dimension (Benhamou & Peltier, 2007). In the book industry authors have looked at different number of titles offered (for example: Van der Ploeg, 2004) and in the music industry at different songs in the top of the charts (Peterson & Berger, 1975)

⁶ <http://unesdoc.unesco.org/images/0014/001429/142919e.pdf>, accessed on May 19, 2011

For our purposes this idea can be made more generic: diversity is the number of different products supplied in a particular industry. In the book industry the products are books or at a more abstract level genres. In the music industry the products are songs, albums. Products can also be understood as artists, since these are also the results of the efforts of the music industry as a whole. Moreover, television programs are products in the broadcast industry and even so program genres.

Interestingly, the discussion on the definition of cultural diversity seems to be intertwined with the operationalization of the concept. Scholars define the concept by explaining how they are going to measure it. I did not find any definition without a measurable component. Originally, cultural diversity was introduced in the field of cultural economics by authors investigating the connection of between market concentration and diversity. Peterson & Berger (1975) were the first to study the record industry. In the same tradition they were followed by Burnett (1992), Alexander (1994, 1996), Lopes (1992) and Dowd (2004). (Benhamou & Peltier, 2007)

2.3.3. Towards a multi-dimensional understanding: variety, balance, disparity

Benhamou & Peltier (2007) and Ranaivoson (2007) have developed a *general* multi-dimensional understanding of cultural diversity, based on research in the field of biology and in the economics of technological change. It seems they wrote their articles in the same time and therefore it is hard to say who was the first. However, Ranaivoson quotes in her article Moreau & Peltier (2004).

Inspired by Stirling (1999) they name three properties of diversity: variety, balance and disparity. The greater the variety, balance and disparity, the greater the diversity. (Ranaivoson, 2007)

“...[V]ariety refers to the number of categories into which a quantity can be partitioned. Balance refers to the pattern in the distribution of that quantity across the relevant categories. Disparity goes beyond these measurement schemes by accounting for the nature of the categorization scheme and adjusting for the degree to which the categories are different from each other” (Benhamou & Peltier, 2007). Each of the components are linked to each other. If variety increases this also effects balance. (Ranaivoson, 2007)

Of these three properties of diversity, disparity is the most tricky. As its measurement completely relies on arbitrary qualifications, this dimension is the most subjective. You need a classification system to define disparity in a sample. In my case I would need a universally accepted list of genres in music with a defined distance between the different kinds of music. Yet, who is to say that metal and classical music differ “6” from each other, while metal and pop differ only “5.3”?

Still, we are discussing diversity in general terms: these three dimensions can also easily be applied to other populations. If we want to go beyond this, what should we incorporate to discuss cultural diversity?

2.3.4. The second dimension.

A further addition to the definition of the concept is the notion of open diversity and reflective diversity (Van der Wurff and Van Cuilenburg, 2001). Benhamou & Peltier (2007) find this useful, but adjust these concepts to supplied diversity and consumed diversity and incorporate these in their definition. However, the normative intentions of Van der Wurf & Van Cuilenburg (2001) are ignored by them. (Ranaivoson, 2007) Reflective diversity implies the extent to which supply and demand are geared or not. In the cultural industries however, supply will never be geared to demand, due to what Caves (2000) calls the *nobody knows* property of cultural industries. It is rational to overproduce, since there is great uncertainty about consumers preferences and whether or not the product will sell. Ranaivoson (2007) also supports this line of thought, but distinguishes consumed diversity from demanded diversity. The last corresponds to consumer taste and to the level of what is desired, independent of the actual supply.

Ultimately, Benhamou & Peltier propose the following definition:

“The cultural diversity in a country means the quantitative and qualitative diversity of the production and consumption of cultural goods and services. It represents the possibilities open to consumers for gaining access to a large supply of a cultural product (in terms of quantity), including segments (in terms of genres and original languages) of relatively well-balanced sizes and diversified. It also represents the effective consumption of these numerous and diversified cultural product.” (Benhamou & Peltier, 2007)

2.3.5. The third dimension

Yet, Ranaivoson (2007) adds another dimension. This is the last layer and encompasses product diversity, producer diversity, and consumer diversity. The different characteristics of a product form product diversity. The number of different producers of cultural goods is the producer diversity. The variation in the public that consumes the supplied cultural diversity is the consumer diversity. This is not the same as the consumed diversity of Benhamou & Peltier (2007) or the demanded diversity of Ranaivoison herself. “...consumer diversity is linked to diversity of cultural identities, beliefs and habits. It reaches far beyond economic issues.” (Ranaivoson, 2007)

The following definition results from the discussion above:

“... I define cultural diversity as the variety, balance and disparity of products in the way they are made available and then consumed; of producers according to their potential market power and the way this power is expressed; and of the consumer as far as his tastes and different identities are concerned.” (Ranaivoson, 2007)

Schematically, this discussion of the definition of cultural diversity lead to this following picture:

Table 2.1: Dimensions of Cultural Diversity

Cultural Diversity			
3	Producer diversity	Product diversity	Consumer diversity
2	Supplied diversity		Consumed diversity
1	Variety	Balance	Disparity

source: Ranaivoson (2007), Benhamou & Peltier (2007)

2.4. Previous research on diversity in the music industry.

See Ranaivoson (2007: 12) for a comprehensive survey of literature on cultural diversity. I will focus my attention on the studies conducted on cultural diversity in the music industry since that is the scope of this paper. There are numerous studies done on diversity in the music industry. Scholars have produced theoretical models as well as empirical studies.

Following Ranaivoson (2007) I will first review the theoretical models and then discuss the empirical studies, since these outnumber the first category.

2.4.1. Theoretical models

Baker (1991) researched the interaction between record producers and royalty receivers that are important in determining the industry performance: the amount of choice customers have (supplied diversity) and the achieved sales in the market. Baker analyzed whether competition or monopoly would yield the best results. Under his assumptions, the first form of organization seems to be preferred for better performance of the industry. Diversity plays the role of risk reducer in his article. He discusses firms with a diversified portfolio of titles are better equipped dealing with demand uncertainty in the record industry.

In their working paper ‘Music Variety and Retail Concentration’, Alain & Waelbroeck (2006) investigate whether market structure influences product variety. They show that vertically integrated firms (those who control the total production flow) will produce more different products than a chain of monopolies will do, because they internalize the fixed costs of production better. Furthermore, competition will give more incentives to maintain a diversified product portfolio. A result that Baker (1991) found as well. Finally, Alain & Waelbroeck (2006: 5) report that competition between retailers can “lead to more product variety than a vertically integrated structure when the cost of launching the new product is mainly supported by the upstream firm.”

2.4.2. Empirical studies

This next part will discuss the empirical studies that focus cultural diversity in the music industry. Some of the studies are also concerned with innovation (e.g. Peterson & Berger (1975); Lopes (1992))

These studies have common ground in Schumpeter (1942) who discussed the relation between industry concentration and innovation. He asserts that large firms are more innovative, since these firms have resources to fund research and development (Peterson & Berger, 1975; Dowd, 2004) As Peterson & Berger (1996: 176) noted: "There must be diversity if there is innovation, but the opposite is not true, because there can be great diversity with no innovation."

2.4.3. The relation between diversity and industry concentration

The academic debate on the relation between innovation and diversity and industry concentration in the music industry started with Peterson & Berger (1975). They investigate their assertion that culture forms go through cycles. They examined the top 10 of the singles Billboard charts that includes 26 years from 1948 to 1973. They show the impact of market concentration, producer diversity - "the extent to which a few firms dominate a market" (Dowd, 2004) - on product diversity in the record industry in the US. Changes in structure precede changes in the music. The degree of vertical integration (see also Alain & Waelbroeck, 2006) and some other mechanisms explain why customers do not necessarily get what they want. Furthermore, they show that some cycles last for a long period characterized by homogeneity in supply, alternated with short periods of improved diversity and competition.

Anderson, Hesbacher, Eitzkorn & Denisoff (1980) discuss in their article the long term trend in the music industry in the U.S. They studied market concentration, genre, artist type and lyric content. Based on their research Anderson et al. pose the hypothesis that song type rather than market concentration is key to predict diversity in the industry.

Rothenbuhler & Dimmick (1982) extend the analysis of Peterson & Berger (1975) to the year 1980. Their objectives were to test the predictions of Peterson & Berger and to see whether they could find evidence for the presence of "unsated demand" which Peterson & Berger coined. This term represents the "non-buying" response of customers in response to low levels of diversity in the market and should result in lower sales. As a third goal they analyze the relation between the number and the success of music producers and industry concentration. Rothenbuhler & Dimmick find evidence supporting the claims of Peterson & Berger: the concentration in the market increases in the investigated period and the number of different top 10 hits decreases. The number of producers falls as well. Regarding the unsated demand they conclude that there are four years with lower sales than the previous year, which implies weak support.

Contrary to the results of Peterson & Berger (1975) Lopes (1992) puts forward that innovation and diversity depend on the development and production system of the record companies in the U.S. rather than on concentration. He argues that an oligopolistic market can sustain a notable level of diversity. In his study, he analyzed singles and album charts published by Billboard. However, singles are no longer the dominant form to buy music. Therefore analysis of the album charts are much more adequate. Lopes only finds partly evidence for the hypothesis that high market concentration reduces diversity.

The number of different artists in singles charts fell slightly over the years together with a tremendous fall in the number of different firms. In the album charts the number of artists however increased at the same time the industry got increasingly concentrated.

Peterson and Berger's argument (1975) about declining diversity seems to hold for the 1970s, but not for the years there after. The open system proves to stimulate diversity from the 1980s on. (see also Burnett, 1992) Another determinant of diversity can be found in the way music is distributed, according to Lopes. Lower diversity in the 1970s is also because so-called "rack jobbers" - intermediate distributors who service the racks in department, discount and other retail stores - controlled 80 per cent of the record sales. They did not have any inclination to promote new artists or styles. They just wanted to sell as much as possible, so their supply was mainly based on the Billboard charts (Lopes, 1992: 69). The increase in diversity reported by Lopes can also be explained by the introduction of the record stores we now know today during the late 1970s, begin 1980s.

Christianen challenged in his 1995 article the operationalization of the variables diversity and innovation in the studies mentioned above. He discusses two pitfalls: the gatekeeping function of the media and the development of diversity of taste in the market for records. So instead of using the charts, he analyzed the total supply of records in the Netherlands in relation to the concentration in the market for record companies in the Netherlands.

He finds a weak negative correlation between these two variables, although not significant. Yet, the result for the percentage of new titles - also negatively correlated - is moderately significant. Spreading over genres of the titles correlated with concentration results finally also in a negative correlation, but this relation has a strong significance.

Even after many years Peterson & Berger's text (1975) received attention from Peter J. Alexander (1996) in an article which contests earlier results from Peterson and Berger. Decreasing diversity, so argues Alexander, is connected with both high and low levels of concentration in the market. Maximum diversity will be realized with a moderately concentrated market. Furthermore, Alexander (1996) proposes a different measure of diversity: entropy. This reflects the degree of randomness in the charts, which he also studies. Instead of using a count method (the different number of artists or albums in the charts) Alexander analyzes sheet music of hits in the top 40 in the U.S. per year. He concludes that concentration and diversity have a nonlinear relation.

In a reply to Alexander (1996), Peterson & Berger (1996) state that with the current developments in mergers and acquisition activity in the music industry it is more instructive to check whether creative control rather than industry concentration influences diversity in the charts. After all, the majors now have competing divisions. Although Alexander (1996) presents his method of analyzing sheet music as robust, Peterson & Berger (1996) bring to the fore that sheet music of top 40 hits are not a good representation. Sheet music is produced after the completion of recording.

Finally, Dowd (2004) addresses gaps in the literature so far. These include the lack of multivariate analysis, failing to investigate the interaction between the open and closed system of production in the music industry and lacking analysis of new performers and recording firms. Dowd (2004) bases his research on these gaps. He argues in favor of the open systems account representing extensive evidence for his claim. Moreover, he extends the open systems account. He puts forward that majors seek to have a more diverse portfolio to capture changes in taste of the public. They do so by acquiring independents and not relying on their own back catalog.

2.5. How should cultural diversity be measured?

The political debate and its policy implications require a straightforward way of measuring diversity. Implementing policies could only succeed when the effects are measurable. This led to the search for numerical indicators or indexes of diversity. These indexes are always to be handled with care since they quantify quality and are therefore not undisputed. (Ranaivoson, 2007) For a thorough analysis of the different diversity indexes see Flôres (2006) and Benhamou, Flôres & Peltier (2009) I will give a short summary, since I will not be able to work with them. Moreover, there are not many studies in which these indexes are applied.

2.5.1. The Shannon-Wiener index, Stirling's H_{sw} and H_{bfp}

A common mistake, according to Flôres (2006) is that researchers associate diversity with the sheer multiplicity of types. He stresses that the relative frequencies of the types within a sample are important as well to determine the magnitude of diversity.

The Shannon-Wiener index measures both indicators - relative frequencies and multiplicity - of diversity. This index originates in biology to determine the diversity of species in a certain environment. It is the mother of the diversity indexes. Stirling (1999) proposed an improved index in which also the degree of similarity (disparity) is measured: H_{sw} . Yes, this is the index which introduced the properties variety, balance and disparity into the debate. Still, these indexes are inspired on work done in biology (Weitzman, 1992, 1993) since Stirling adds the distance function D_{ij} already seen in Weitzman's work.

Although the incorporation of similarity is appealing, Stirling's index leaves us with some serious considerations to be made. Flôres (2006) points to some limitations: in the original formulation, the index is dependent on the measurement unit for the distances. Comparison with other indexes is therefore difficult if these adopt other disparity measures. Furthermore, "the index is invariant to linear transformations on the set of distances", yet it is not invariant to other classes of transformations. Also, the maximum of Stirling's index is difficult to interpret, because this is dependent on the set of distances within the sample. (Benhamou, Flôres & Peltier, 2009) The maximum for Shannon-Wiener is the natural log for N and counts for the uniform distribution, "which has a strong intuitive appeal". Another drawback against Stirling's index comes to the fore when a researcher decides to use subsets of types in its sample.

Flôres (2006) names movies, which can be categorized in genre, then in a branch of this genre and then by the country of production for example. You would expect diversity not to decrease when undertaking this form of categorization, but Flôres (2006) shows that H_{sw} can display any behavior.

In an attempt to improve Stirling's H_{sw} Benhamou, Flôres & Peltier (2009) introduce a referent in their index H_{bf}, so the index will also reflect the openness of the index. This way they measure the distance between the types, weighted by their importance *and* the distance of each type to the referent.

The important research done on cultural diversity in the record industry does not employ any of these indexes discussed above. In the next section I will explore the ways in which cultural diversity has been made measurable in the record industry.

2.6. How is cultural diversity measured in the record industry?

See tables 2.2. and 2.3. for an outline of the various variables used in academic cultural diversity research. The tables mirror the different properties and the different dimensions set out by Ranaivoson (2007) and Benhamou & Peltier (2007). Striking is the almost complete absence of disparity. This property is not easy to measure and therefore only Dowd (2004) tried to measure it in the record industry. Also, consumer diversity has not been researched in the record industry. Main emphasis is on the products and producers, rather than the diversity of consumers. Since economics is not concerned with individual tastes it is understandable that this dimension is absent. Furthermore, all studies (except Christianen, 1995) research consumed diversity. In those studies bestseller data is used to analyze. This type of data reflects the top of the market and does not capture all the cultural goods produced in the market. Yet often, this is the only way to gather data. It is very rare that information is available on all supplied goods in a market.

Table 2.2: Consumed diversity: variables used in diversity research in the record industry.

3. Consumed diversity	
2. Product diversity	
1. Variance	
Content characteristics	Alexander (1996)
Lyrical content	Peterson and Berger (1975), Anderson et al. (1980)
Type of interpret	Anderson et al. (1980)
Number of products available	Baker (1991)
Number of new products	Allain and Waelbroeck (2006)
Number of different products/artists in top x	Peterson and Berger (1975), Rothenbuhler and Dimmick (1982), Lopes (1992), Alexander (1996)
1. Balance	
Share of producers in top x	Peterson and Berger (1975), Anderson et al. (1980), Rothenbuhler and Dimmick (1982), Lopes (1992), Alexander (1996), Dowd (2004)
Share of sales by genre/artist (of a given genre) in top x	Anderson et al. (1980)
1. Disparity	
Content characteristics	Dowd (2004)
2. Producer diversity	
1. Variance	
Number of different producers in top x	Peterson and Berger (1975), Rothenbuhler and Dimmick (1982); Lopes (1992)

source: Ranaivoson (2007)

Table 2.3: Supplied diversity: variables used in diversity research in the record industry

3. Supplied diversity	
2. Product diversity	
1. Variance	
titles of albums	Christianen (1995)
1. Balance	
record companies and labels	idem
genre	idem

2.7 The relation between Digitisation & diversity: Theory.

2.7.1. Nelson's product searches.

The theoretical basis for the hypothesis of this thesis is based on theory developed by Nelson (1970, 1974) on consumer information search. He introduced the concept of experience goods. Nelson was able to make and test some predictions which are still relevant for my topic. He stresses that limitations of consumer information about the quality have profound effects upon the market structure of consumer market goods. Since digitisation dramatically changed the way we get information this theory could be useful to make and test predictions about the music industry.

2.7.2. Information goods: search & experience

Consumers already know what they can obtain. Their problem is to evaluate the utility of each of the options open to them. They have an information problem. Discovering the prize of the item you want to buy is easy. Information about the quality of this product is however hard to obtain and costly.

In his seminal article, Nelson (1970) proposes a classification of information goods in search and experience goods. His focus is on strategies to discover the quality of a product in the market. Keep in mind that theories are usually a product of their time. The following discussion of Nelson's theory shows this tendency. Some statements were perfectly plausible in the 1970s, but may seem strange now.

According to Nelson (1970), there are two strategies to obtain information: search and experience. The interplay between the expected utility and the costs of search determine which strategy will be used to solve the information problem of a customer. Consumers do not have to start with no information whatever. Prior to search or experience, information can be obtained from magazines, advertisement and friend and relatives.

The first strategy is search and this means for Nelson two things: the consumer must inspect the option and this must be done prior the purchase of the good. All the information for these goods is thus available, it only has to be uncovered. Nelson (1970) builds on Stigler's theories of search. (Stigler, 1961; 1962) Stigler asserts that under some given conditions his model could predict optimum the number of searches for a product. If we assume every customer tries to maximize its utility, this very customer will employ search until the marginal costs of a search will exceed the marginal utility.

Yet, search will sometimes be more expensive than the good itself. Especially when there is not much knowledge available about the quality. For example, in the 1970s you had probably to go to a library to look for consumer reviews on a certain product. This requires first a subscription to a library and a lot of time and effort to decide which product to buy. In those instances the consumer will simply buy a product to evaluate it. Nelson calls this process "experience".

Consequently, Nelson (1970) assumes that experience goods are products which quality can only be determined after use (Huang, Lurie & Mitra, 2009) and that the most preferred brand will remain the top choice over time.

Other assumptions he makes are that consumers sample at random from all available products or from products in a certain price range. His last assumption is important: the only way to experience is to buy the brand. We will come back later to this premise as I discuss sampling.

As with search goods, consumers will buy experience goods until the expected marginal costs of experience is equal to the expected marginal return. The marginal costs in the experience case is the difference in utility from using the best brand minus the costs of consuming a brand at random. Furthermore, the number of experiments is related to the frequency of purchase. These two factors determine the number of experiments a consumer will undertake to evaluate a product. It is this number of experiments that affects the variation in market structure.

Nelson predicts lower monopoly power for firms in a market for search products. Consumers who sample products will evaluate less options than when they search for information. In other words, the sample size for search is larger than for the experience goods. The greater the sample size of consumers the smaller the monopoly power is for an individual firm, since the number of experiments sets an upper limit to the number of brands a market can support.

2.7.3. Additions to Nelson's theory.

Darby & Karni (1973) later proposed a third class of properties, called "credence" qualities. These qualities cannot be determined in normal use. An extreme example is the removal of an appendix: even after experience, information about the quality is difficult to obtain and very costly. These credence qualities occur when a good is used in combination with other products having uncertain qualities or in a production process in which the output is stochastic. Or both.

2.7.4. The impact of Internet on search costs

The Internet (as a part of the more coordinating term digitisation) has dramatically changed the way people search for information. Literally, trillions of gigabytes of information is now available for everyone with an Internet connection. Searching for information has become cheaper, since a lot of effort can be done behind a desk with a computer instead of going out in the physical world.

2.7.5. Are albums experience goods or search goods?

Have all the goods transformed to search goods with the advent of Internet? Not immediately. Music albums are information goods and are often labeled experience goods. (Gopal, Bhattacharjee & Sanders, 2006; Caves, 2000; Nelson, 1970) It is very difficult to predict the particular value and quality of an album beforehand for an individual. Even if there is unlimited information available. Only after a couple of rounds listening to a record you could determine whether its quality is just as you expected. If you compare a music album with computer parts, it is easy to see which product is more fitted to sell online. For the computer parts you just have to list some technical details and these characteristics speak for themselves.

However, perfect search or perfect experience goods do not exist in the real world. Most likely, all the goods will have properties of the three formulated categories. (Darby and Karni, 1973; Nelson, 1974; Klein, 1998; Handke, 2010) Although music albums have clear experience goods characteristics, they also attest to some search good properties. The opinion of others is in this respect helpful to make decisions.

Luckily, the Internet makes it very easy to find opinions of others to aid your decision. Sites like Last.fm⁷ provide consumers with extensive pre-purchase information. Eventually, the clear distinction disappears between search and experience goods. All the information⁸ is available before purchase, but you still have to experience music to pass your judgement on the quality of the product.

2.7.7. Sampling

“Online music technologies are fundamentally altering the landscape of the music business.” (Gopal, Bhattacharjee & Sanders, 2006)

The possibility for consumers to sample music through the Internet is what bridges the gap between music as search and/or experience good. Simply put, sampling music through the Internet is listening to a shortened song. Arguably, digitisation leads to a shift in search strategies: music transforms from a mainly experience good (paid sampling) to a search good, yet partly. It is now more feasible to judge the quality to a customer prior to purchase of music. However, “in practice searching will not overcome consumers’ quality uncertainty entirely due to extensive product differentiation and oversupply” (Handke, 2010: 117)

Nelson (1970) stresses that producers of experience goods have more monopoly power in the market. Monopoly power occurs when consumers sample from a small range of products in a market. With the advent of Internet and the possibility for consumers to sample from a very large pool of music this monopoly power undergoes a major shift. The market structure changes. Gopal, Bhattacharjee & Sanders (2006) stress the same argument when they state that the decreasing sampling costs for consumers for music results in an erosion of superstardom. In short, superstardom reflects that few artists cover a very large share of the market (Rosen, 1981) and thus have significant market power.

2.8. Frictionless markets

The massive surge in digitisation could eventually lead to a ‘frictionless’ economy: markets which come close to perfect competition. If fixed costs fall due to exceptional savings on the production, reproduction and distribution of cultural goods and if the ancient information problem of consumers has been solved a whole new reality could divulge. Brynjolfsson and Smith (2000) provided first evidence for the hypothesis that the Internet is “a more efficient channel in terms of price levels and menu costs” In other words, the “contestability” (Baumol, 1982) of the market will rise. There will be more competition. Contestable markets refer to markets with low barriers to entry where a small number of suppliers compete. Still, this small number of suppliers is fiercely competing, since there are always firms trying to capture market share not hindered by high entry and exit costs.

⁷ <http://www.last.fm>

⁸ The information what can possibly be provided. Information about how you will like an album if for the future to show.

Handke (2010) argues furthermore that due to lower information costs demand will be less concentrated on well established brands. Brands are important if goods have mainly experience goods characteristics, this requirement in the market fades if goods become increasingly search goods.

2.9. The Long Tail

In line with the two theories above Anderson (2004) argues that top hits will conquer less market share as the Internet use expands. The Pareto principle - a small portion of the products accounts for a larger part of the sales - seems not valid any more. Anderson coined the term "Long Tail" to describe the shift in the balance between hits and niche products: the latter sell more, even with stable sales. This would imply more diversity, since it will become more profitable to release niche products. Evidently, the Internet is able to create or open new previously undiscovered markets for obscure tastes among consumers. Explanations for the phenomenon include the virtually unlimited shelf space for retailers on the Internet on the supply side and the lowered search costs in combination with an increase in quality of pre-purchase information (including sampling). (Brynjolfsson, Hu, Simester, 2011) Brynjolfsson, Hu, Simester (2011: 4) find empirical evidence that the Internet channel "exhibits a significantly less concentrated sales distribution". They believe that lower search costs is the main driver behind this shift.

2.10. Digitisation leads to more diversity.

Digitisation may therefore lead to greater diversity in the music industry. Consumers are able to tap into a larger sample of available music. This could result in two dynamics related to diversity: the charts will reflect the larger sample size and show more variety in artists and albums. This is an effect of digitisation on consumed diversity. Furthermore, since established firms lose their acclaimed monopoly power, the supplied variety will also increase. As more firms are able to enter the market due to lower entry barriers, the overall supply of music will increase and consequently the diversity as well. I will test the first effect, since this is the only prediction which is testable at this moment. There is no data available for me on the overall supply of music in the Netherlands. Yet, there is for consumed diversity in the form of the charts.

All theories discussed attest to the following impact of digitisation on diversity: the latter will rise as digitisation continues. Therefore I formulated my hypothesis as follows:

"Digitisation leads to more diversity in the supply of records in the Netherlands."

Considering the discussion of diversity in political context, improving our knowledge about the relation of digitisation and diversity could aid the decision making process of new cultural policy.

In the next chapter I will set out my research strategy, discuss and test my hypothesis and report my results.

3. Method

3.1. Data

To find support for my hypothesis I will analyze secondary data with SPSS. (statistical package for the social sciences). This data is not collected by myself, but is generated by others for other purposes. I have two data sources and four variables in my study. My dependent variables are the different number of artists in the top 50 of the Album top 100 (UA) and the different number of albums in the top 50 of the Album top 100. (UAL) These two are proxies for diversity in the charts. The independent variable is digitisation and made operational by the number of Internet user per 100 people (IU). I will discuss these variables in greater detail in the next paragraphs.

3.2. Dependent variables: diversity in the charts.

The data source for the dependent variables is the Album top 100 issued by GfK Dutch Charts in the Netherlands. They publish this top 100 every week on their website www.dutchcharts.nl. I contacted Michel Admiraal - a senior chart manager at GfK, and he sent me almost every issue of the Album Top 100 from 1980 until 2010 in Microsoft Excel format. This was very convenient, since this data format allowed me to easily filter the data.

GfK publishes The Album top 100 every Friday. The rankings are based on *sales* - either physical and digital copies. The sales are measured from Friday until Thursday. GfK applies a lower limit on sale prize of EUR 8,01 from the record companies to the retailers. After an album has been in the chart for two years it gets relegated to the back catalogue Top 50 chart⁹.

3.2.1. Why do I use the Album top 100 as data source?

Since individual singles are not longer the dominant form of record music (Lopes, 1992), I will analyze the album charts rather than single charts. The NVPI market report for 2010 illustrates this claim for the Netherlands.¹⁰ According to their data, in 2010 177 million albums were sold against 1,2 million singles. Moreover, because of the paucity of market level data on performers and firms authors of diversity studies use chart data to track the mainstream market (Dowd, 2004). The charts are often the only data available. It is only Christianen (1995) who was in the fortunate position to have an other longitudinal data source to his disposition.

⁹ The information in this paragraph is from Michel Admiraal personally: he wrote this in an e-mail to me.

¹⁰ <http://www.nvpi.nl/sites/default/files/nvpi-marktinformatie-audio-2010.pdf>, accessed on August 7, 2011.

Beside that the charts fill the most elementary availability gap, one of the advantages is that they make it possible to look at developments over a long period of time. The Album top 100 is the only album chart which has been published uninterrupted to date and therefore the only candidate to use for this research design.

3.2.2. The Short History of the Album Top 100.

In this paper I refer consistently to the Album top 100 (“the charts”), but this hit parade has a long history. It is the oldest hit parade in the Netherlands. Through out its existence it has changed name and composition many times. The parade started of as the “Hilversum 3 LP top 10” at May 23, 1969. In 1974 the chart welcomed an additional 10 quotations. Four years later the Nationale Hitparade LP top 20 became the Nationale Hitparade LP top 30. In that year the list was extended to 50 quotations. In 1985 was the next change to 75 entries each week. The 100 quotations is still has today was first issued in 1989. Its current name Album top 100 was introduced on July 10, 1999.¹¹

3.2.3. A proxy for diversity

Earlier in the paper I wrote: “A common mistake, according to Flôres (2006) is that researchers associate diversity with the sheer multiplicity of types”. He might be right, but I have no choice to do exactly what he has forbidden me. With my charts data I am able to measure the product and producer variety of the supply in the record industry of the Netherlands on an interval level. It may be a crude measure of diversity (Handke, 2010: 114), but it will hopefully bring our understanding of diversity in the record industry a bit further. Other dimensions of cultural diversity proposed in the literature by Ranaivoson (2007) are unfortunately not measurable considering the nature of my data. There are no disparity taxonomies included in the data. Until 2002 my data has no values for the publishing record company. So although there possibilities to measure producer balance, the time span is short just now. I have decided to concentrate on the long term developments of diversity. I do can say something valuable about the development in the variety component of diversity.

In a way a chart is in itself a measure of balance. It reports on the share of total supply which passes a lower limit of sale to the greater public. If only I know how many different titles are published each year in the Netherlands. I do not have data that tells me.

3.2.4. How I processed the charts data into a workable data set

The first step was to remove all the entries beyond number 50 in my data set. This was simply done by using the filter options in Microsoft Excel. In the column which hold the positions of the different entries I filtered all values above 50. The corresponding rows were then deleted from the data set. Now I had left the top 50 of the 1979 - 2009 period of the Album top 100. The next step was the extraction of all the entries per year. Another column in the data set contained the date of issue of each edition. Again, I used the filter options in Excel and selected all cells for a particular year. These cells I copied to a new file. In the new file I selected first the entire column that included the artists of every quotation.

¹¹ Source: http://nl.wikipedia.org/wiki/Album_Top_100, accessed on August 4, 2011. This page is checked by Michel Admiraal.

This selection was then filtered with the 'Unique records only' advanced filter option of Excel: Data > Advanced Filter > Unique records only. This was a very easy way to get rid of all the many duplicates in the data. I pasted the selection into a blank work sheet and just read of the number of the last row containing data. This number was the number of different artists in the top 50 of the Album top 100 in that year. The same procedure I employed for the number of albums. Eventually, I did these two procedures 31 times to compile my data set which I could use with SPSS to carry out the statistical analysis.

3.2.5. Missing data chart data

Unfortunately, data from the years 1988 and 1989 is missing. Consequently, I do not use these years in my calculation, since this omission could harm the analysis. The missing values are indicated to SPSS by the missing values option in the Variable View. I used the "Discrete Missing Value" option and used the 9 as the indicator. SPSS gave this message in its output "User-defined missing values are treated as missing."

3.3. Independent variable: Making digitisation operational.

In this paper I test the relationship between digitisation and diversity. How can I make this first concept measurable so I could use digitisation as a variable in the correlation analysis? The concept of digitisation is rather abstract and means in ordinary speech "converting to numbers". Not very useful for me though. This paper handles about the influence of upcoming ICT on diversity. There is however not an index available which reflects the extent to which digitisation has developed in a country. So I have to use a proxy for my independent variable. This is a measure which serves as an indicator of digitisation. As shown in my discussion of the theory connected to digitisation we saw that the information dissemination of the Internet is a very important factor. This led me to search for data on Internet usage. If lowered search costs for consumers due to the Internet leads to shifts in markets ('frictionless' markets) and a larger product population to choose from (Nelson, 1970) Internet use is a good candidate to serve as a proxy of digitisation.

I found a data set from the World Bank containing information about Internet usage per country to serve as an operational definition (Bryman, 2009: 144) of digitisation. I selected the Netherlands. This set reflects Internet users (per 100 people) where Internet users are "people with access to the worldwide network."¹² Their sources include the International Telecommunication Union, World Telecommunication/ICT Development Report and database, and World Bank estimates. The data is measured at an interval level and ranges from 1990 until 2009 with 8 eight decimals accurately.

¹² <http://data.worldbank.org/indicator/IT.NET.USER.P2>

3.4. Research strategy

I will adopt in this paper a deductive research strategy. From the theory I formulated a directional hypothesis which I will test with the collected data.

Hypothesis: Digitisation leads to more diversity in the supply of records in the Netherlands.

Null Hypothesis: Digitisation has no effect on the diversity in the supply of records in the Netherlands.

I will carry out a bivariate correlation analysis between my dependent variables for diversity and the independent variable digitisation. I will fit a statistical model to the data that represents my hypothesis and I will see to what extent it explains the observed variance. (Field, 2009: 49) Why not a complete regression analysis? As I studied my data it occurred to me that my independent variable is almost at its maximum: in 2010 there are almost 90 out of 100 people who use Internet in the Netherlands. Main feature of regression is to predict future values, but this is not a useful exercise considering my data. Moreover, the correlation coefficient R^2 in this situation are both the same.

With a correlation analysis I cannot however prove any causality between the variables. As I employ this analysis, I will only be able to determine whether there is a relationship between the variables at all. Moreover, I can draw conclusions about the strength of the relationship. So I cannot prove my hypothesis, but I can support it with my data analysis.

To support my hypothesis I will have to find reason to confidentially, at least, reject the null hypothesis. We will look at “changes of obtaining the data that we’ve collected assuming that the null hypothesis is true” (Field, 2009: 27). This chance has to be as low as possible.

This study will be a longitudinal secondary analysis: it is a correlational research that involves repeated observations of the same variables over a long period of time (Bryman, 2009: 49). My data set include 31 years of observations. The data ranges from 1979 until 2009.

The first setup of the study included the analysis of the complete Album top 100. Yet, this proved to be impossible due to nature of the data. The Album top 100 has not always been a chart with 100 entries. Since it has had different compilations through its history, I will have to account for these changes. The validity of the results is best served if the sample in each year stays the same.

The choice to analyze data beginning in 1979 is based on this requirement. This year was the first year with 50 entries during the whole year. The compilation changed to 75 in 1985 and to 100 in 1989. By selecting the years 1979 to 2009 this study covers 77,5% of the total number of years in which the Album top 100 (and its predecessors) are published.

3.5. Defining two periods

The Internet hit its significance in 1994: this was the year of its breakthrough to the public. “Originally intended as a distributed network of computers it is increasingly viewed instead as a distributed space of information” (Schatz, 1994). It was the year in which the first browser, Mosaic (predecessor of Netscape), was introduced. The period before 1994 will be my pre period. The reason for the periodization is that to uncover a genuine relationship between digitisation and diversity we have to exclude the possibility of a trend in the diversity in the years before digitisation. I will test for trends in the pre period and the post period. This post period will consequently be 1994 - 2009. A simple linear regression test will be employed to test whether there is a significant slope in the data in both periods. To check for the trend I create a variable that signifies the years, where 1979 = 1. This enabled me to see whether through time diversity increased, regardless of digitisation.

3.6. Criteria in social research.

3.6.1. Reliability and replicability.

The result of this study is repeatable. The data I collected is for everyone available on the Internet and I tried to describe my research methods as detailed as possible. The measures I devised for the concepts Digitisation and Diversity will be applied in a consistent manner. Test-retest methods are not needed to establish the reliability of this study. I will always get the same results, since I measured diversity with objective indicators as the the number of artists in the charts. The *stability* of my study is good. Reliability is more at stake when you formulate multiple-indicator measures (Bryman, 2009: 150), but I did not do that.

3.6.2. Validity

Will my research generate trustworthy conclusions? This final section will reflect on relevant different notions of validity. I did not find any independent statistical tests to check validity objectively for my research. So I will discuss issues and point to problems and strengths.

3.6.2.1 Measurement validity.

Peterson & Berger (1975) used the charts and in the literature only few scholars had reservations to the use of charts *an sich* for research into diversity ever since. From this observation we could establish my research *face validity* (Bryman, 2009: 152) There are no alternatives for the The Album top 100 in the Netherlands. It is the only consistent album chart. Its composition has changed several times throughout the years, but these changes were often not major influences.

Whether of not my measure of diversity reflects the concept of cultural diversity is hard to say. The concept is in itself very equivocal. In the literature a range of solutions are offered to measure the concepts comprehensively, but no attempt has succeeded. (See Flôres, 2006; Ranaivoson, 2007; Benhamou, Flôres & Peltier, 2009) I stated clearly what and which dimension exactly I measured, devoid of all pretence. As for the measurement validity of digitisation, we can say that the data comes from a reliable source. I do not know whether the methods of gathering the data has changed over the years. However, measuring the amount of Internet users is straightforward and not subject to subjectivism.

As I already mentioned the maximum value for my proxy of digitisation has almost been reached. In 2010 there are 90 out of 100 people who have access to the world wide network. The predictive potential is therefore remarkably low and so is the *predictive validity*. Finally, do I measure exactly what I want to measure? Not completely. I am interested in the development of diversity in the total supply of records, but I do not measure this. I measure the development of diversity in the charts. These charts cover only the most popular fraction of the market for record in the Netherlands. Still, I will be able to uncover a trend in the development of diversity.

3.6.2.2. Internal validity

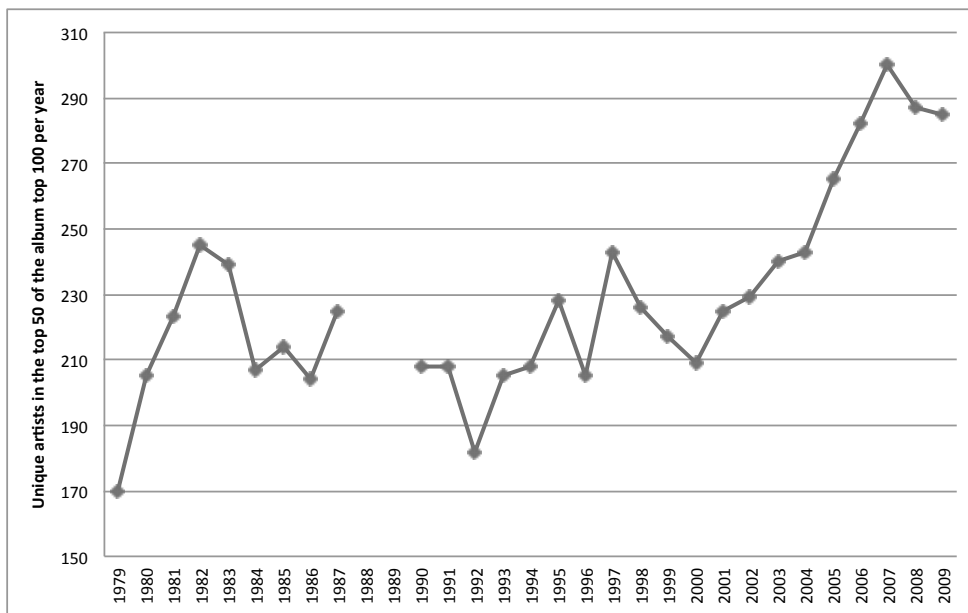
This refers mainly to the issue of causality. Internal validity for this study is typically weak, since I cannot draw any conclusions about the direction of causality. Correlation analysis will not yield such results as it only tells us about the shared variance between the dependent and independent variables in my study.

3.6.2.3. External validity

In this study I studied the development of the diversity in the top 50 of the Album top 100. The question is whether my results can be generalized beyond my results. We cannot assume that the amount of different artists doubles if I enlarge my sample from the top 50 to the top 100 of the charts. I do not sample randomly from my population of artists or albums in the charts. The fact that I selected the top 50 of the Album top 100 implies concessions the external validity. However, regarding external validity I was in a lose-lose situation anyway. If I studied the entire Album top 100 I could only study twenty years instead of the 31 I now examined. I chose to have more years to study instead of greater generalizability. External validity is therefore also weak.

4. Results

Figure 4.1: Development unique artists in the top 50 of the Album top 100



When we look at the data collected in figure 4.1. to measure diversity in the Album top 100 over the span of 31 years, it immediately strikes that for a long period the sheer number of different artists moves within a restricted bandwidth. From 1979 until 2000 diversity goes up or down, but there are ten years as well which show a the same number of artists - around 210 - per year. The peak in this first period is eye-catching. In 1982 245 different artists entered the charts. The amount would be reached again fifteen years later in 1997. What could have caused this spike in the graph? The answer is provided by Christianen (1995: 60). In this year new firms entered the market. These record companies published works from foreign labels new to the market in The Netherlands. A large amount of new titles became available. This effect lasted two years. Also, the missing two years 1988 and 1989 are clearly visible. We have reason to believe the values for these years will also lie between the bandwidth of 180-245. There are no major changes in composition of the Album top 100 in these years.¹³

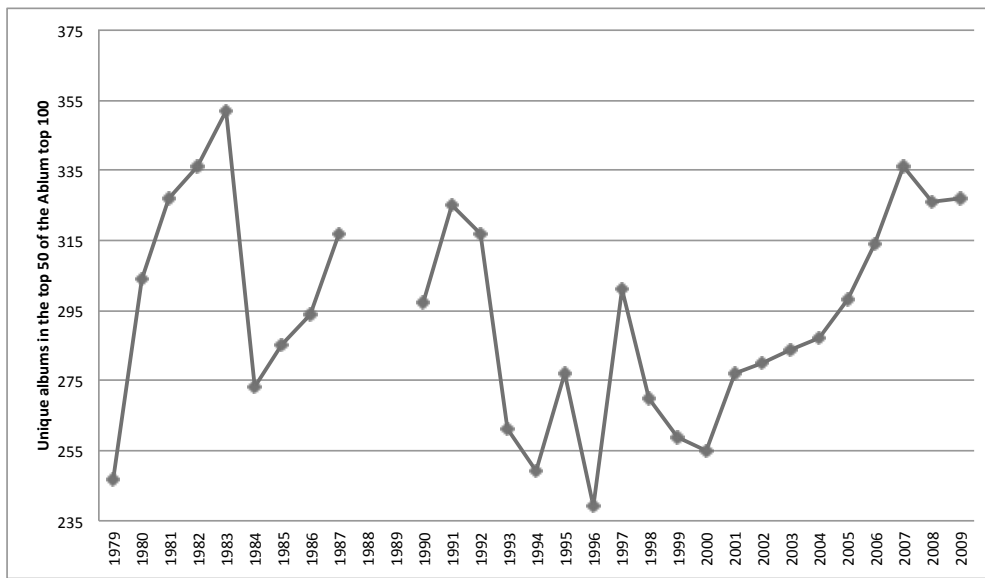
A sizable drop in the amount of different artists is shown in 1992. I did not find any apparent reason in the literature. Nonetheless, when I looked at all the 1992 editions of the Album top 100 individually it appeared to me that in this year, for whatever reason, there were many compilation albums in the charts. The data does not include a value for artists when there is a compilation album quoted.

¹³ http://nl.wikipedia.org/wiki/Album_Top_100, accessed on August 4, 2011. This page is checked by Michel Admiraal.

It was just in 1993 that compilation albums were relegated from the Album top 100. So, compilation albums have elbowed out albums from bands or solo artists in the top 50 of the charts. This observation checks out when you look at figure 4.2: in 1992 there was no drop in the amount of different album titles.

From the second high in 1997 diversity drops three consecutive years. In 2000 it reaches its modus: 209. Without an exception the number of artists increases every year the following seven years. This is the only sequence of years that could be labeled as a trend, judging visually from this graph. In the upcoming analysis I will test for trend in the data. For now we could keep this period in mind. Another observation: The all time high in 2007 is the first year in the 300s: 300 different artists were quoted in the Album top 100.

Figure 4.2: Development unique albums in the top 50 of the Album top 100



The development of the amount of different albums each years in figure 4.2 shows roughly the same picture as that of artists. The graph looks a bit more stretched though. Per year the changes in the numbers are larger than their artist's counterpart. Note that the Y axis does not have another interval of increment, so it is not the way the data is displayed that causes this image to arise.

The entry of the new firms in 1982 does have a much greater effect on the course of development than it had on the development of the amount of different artists. Yet, it is not the year 1982 that is the all time high, but 1983 instead. This picture is a reversed one of that in figure 4.1. Apparently, in 1983 the new firms released more albums of the same artists.

There are two years with relatively exceptional decline: 1984 and 1993. Both can be explained. The first is simply the effects of the dried up extra supply from the newly entered companies in the market. The second drop in the amount of different albums is due to the exclusion of compilation albums in the charts effected in the early months of 1993. The lowest point in this 31 year period is 1996. In this year 239 different titles were present in the charts.

4.1. Checking for trends

Judging from figure 4.1, ostensibly no trend is visible in the pre period from 1979 until 1994. We do however not know for sure. There are several ways to check trend shifts. I checked the trend in the pre period by running a simple linear regression against UA and time (YearREG, 1979 = 1).

4.1.1. Pre period: 1979 - 1993

R^2 is .026 and this means that time accounts for 2,6 per cent of the variation in the amount of unique artists in the top 50 of the Album top 100.

The F-ratio is .291, $P = .600$ and is not statistically significant at $P < .05$. The chance that our F-ratio would occur if the null hypothesis were true is greater than 5% and therefore we can conclude that this model does not fit the data well.

If there were a trend in the pre digitisation period, the slope of this simple regression has to be significant: or in other words: the chance it is 0 has to be as small as possible. I on the other hand want the slope to be insignificant, since this would imply no trend.

The slope (-.035) is not statistically significant at $P < .05$. Its observed significance is 0.600 which is rather high. This means that time does not make a significant contribution to predicting UA in the pre period. ¹⁴

Next objective to establish whether there is a trend in the unique number of different albums in the Album top 100. The procedure is the same as with the unique artists.

The test showed the same results as before: there is no significant correlation with time and diversity in albums. R^2 is .001, indicating that 1% of the variation in albums is shared by time. The F-ratio 0.08 is not significant at $P < .05$. The slope is -.004 and its significance .930. Also not significant at $P < 0.05$. ¹⁵

Considering these first two observations I infer that there is no trend in the data before starting point for digitisation. Beginning in the year 1979 until 1993 (with the exclusion of the year 1988 and 1989) diversity did not develop under the influence of time. There could always be other variables which influence this development, however. An ANOVA analysis does not tell us about the individual contribution of variables in the model (Field, 2009: 207).

4.1.2. The post period: 1994 - 2009

Next task to carry out is the check whether there is a correlation of time with diversity in the period after digitisation set in. I will again test this with a simple linear regression to see whether there is a significant slope in the results. Figures 4.1 and 4.2 show a positive trend in the charts data beginning in 2000. We will now find out whether this trend is still there if we incorporate the early years of the post period. These years (1994-1999) show no apparent logic: numbers go up and down with no clear pattern.

¹⁴ See SPSS output 1.

¹⁵ See SPSS output 2.

I analyzed the year 1994 until 2009 and from this analysis follows that, in contrast to the pre period, the gradient of the regression line is significant. This slope .132 is highly significant at $P < .001$. R^2 is .721 and this tells us that time shares 72,1 percent of the variation with UA. The F-ratio 39.714 is significant at $P < .001$. The model is a good fit to the data.¹⁶

There is a trend in diversity of unique artists in the post period which implies that something has changed. Before I will the relation of digitisation with the diversity, the number of different albums should also be tested for a trend in the post period.

As with the number of different artists in the Album top 100, the unique albums also show a significant slope at $P < .001$ with .135. Time shares 67,8 per cent of the variation in diversity in the album charts, $R^2 = .678$. The F-ratio is 29.494 and statistically significant at $P < .001$. The model predicts the number of albums significantly well.¹⁷

I identified trends in the pre and post period. There is no trend in the pre period and there is a positive trend in the post period. The aim of this thesis is to uncover a relationship between digitisation and diversity, if there is any. Since we now know that were no developments which causes a trend in the years before digitisation, we can maybe assign the positive trend in diversity - either in the number of artists and albums - found in the post period to digitisation.

4.3. Correlation analysis

The method I use for this determination is a correlation analysis. As I want to discover the relation between two variables, this will be a bivariate correlation analysis. SPSS provides the calculation of three different coefficients to determine the strength of relationships between variables. These three include Pearson's product-moment r , Kendall's τ and Spearman's ρ . Before I can carry out the correlation analysis some assumptions have to be tested. As for the first parametric correlation efficient, Pearson's r , the data is required to be normally distributed.

Pearson's r is able to check for a linear relationship between two variables, but if the data is not normally distributed and measured at an interval level we cannot draw any conclusions whether the correlation is significant. (Field, 2009: 177) And this is exactly what we do want to know. The two other correlation coefficients do not require the data to be normally distributed since these are two non-parametric tests. The second assumption for Pearson is met, since the data is at an interval measurement. Now, we will look at the first assumption of normality.

16 See SPSS output 3.

17 See SPSS output 4.

4.3.1. Checking normality

With the Shapiro-Wilk test I can check whether the distribution as a whole deviates from the normal distribution. It is an exact test, so it can be used with small samples¹⁸. ($N < 20$) Results smaller than 0.05 are significant deviations from the normal distribution. I ran the test on the years of the post period, so from 1994 until 2009. The results will show 16 df (degrees of freedom). First I will test the IU variable which functions as the indicator of digitisation.

The Shapiro-Wilk test for IU is not significant at $P < 0.05$ since its .086 value for the test is larger. We can thus assume this variable is normally distributed. Next I will also run this test for the variables which indicate the diversity in the charts: UA and UAL.

Can we assume that the data is normally distributed? Yes, we can. Both scores, .088 for UA and .753 for UAL - show for the strict Shapiro-Wilk test no significant deviation.

Therefore we can use all the three correlation coefficients in the analysis of the impact of digitisation on diversity in the charts.¹⁹ That is not to say every test will be useful. Pearson's r gives more informative results as it is parametric and therefore I will report on this statistic here. In the Appendix I included test results for the other two correlation tests for control purposes.

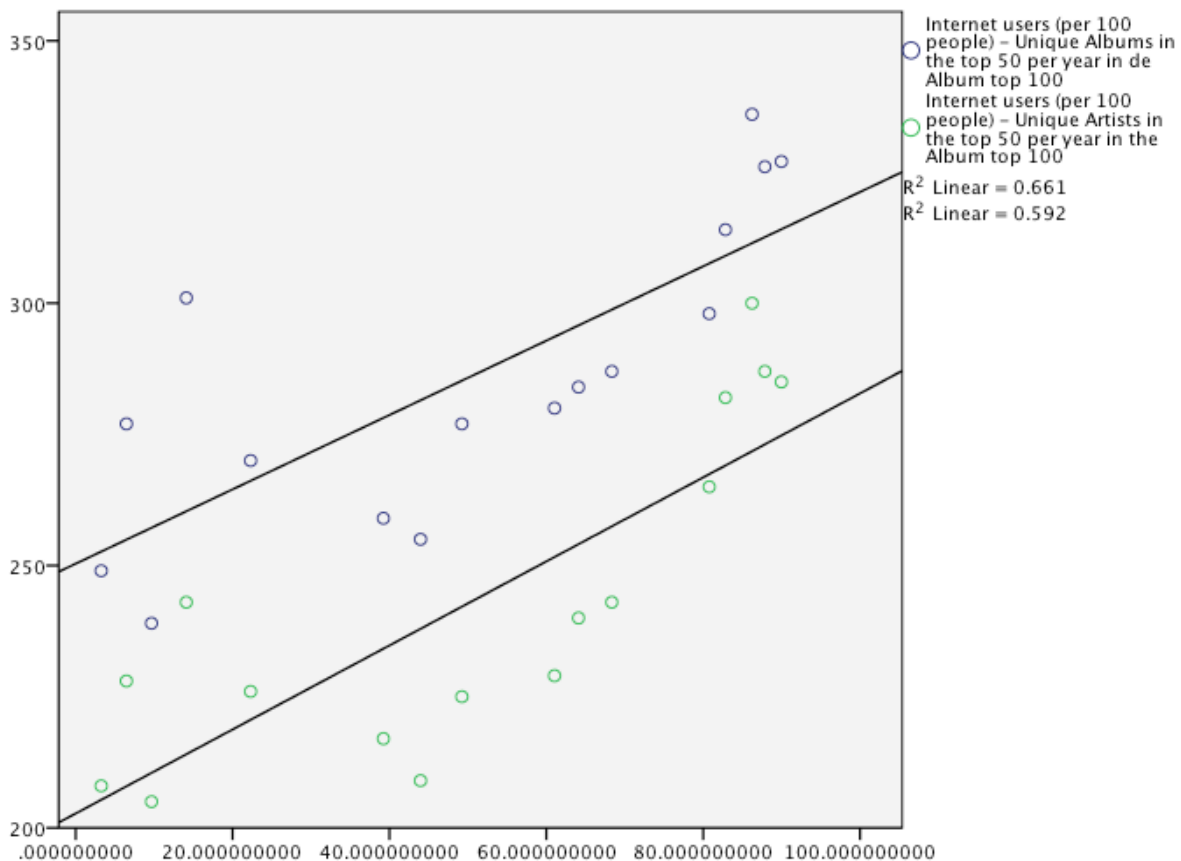
4.3.2. A first exploration of the data

As figure 4.3 shows there seems to be some general trend in the data. When IU goes up, the corresponding values for both diversity indicators go up as well. There are no obvious outliers which cannot be explained with external factors. See for example the 1982 case I made earlier in this chapter. I plotted a linear regression line into this graph. The points on the plots seem to fall nearby the line, indicating that the regression line could be a good fit to the data. Furthermore, there are no clear points with high diversity and low IU. Figure 4.3. confirms the presupposition that digitisation and diversity are related.

1818 Field, A. <http://www.uk.sagepub.com/field3e/Downloads/addmaterial/DSUS3%20Chapter%205%20Web%20Material.pdf>, page 18.

19 See SPSS output 5.

Figure 4.3: Scatterplot of the relation between Internet users per 100 people in the Netherlands and diversity in the Album top 100 from 1994 to 2009.



4.3.3. Pearson's r .

Since I test a directional hypothesis my correlation tests are one-tailed. UA and UAL are both approached positively correlated with the number of Internet users in the Netherlands. It is assumed that IU influences UA and UAL and not the other way around.

The number of Internet users is significantly correlated with the amount of different artists in the top 50 of the Album top 100, $r = .813$, P (one-tailed) $< .001$. There is a genuine relationship between these two variables.

From the test we can also conclude that the unique number of albums in the top 50 of the Album top 100 shares a large amount variation with the number of Internet users. Its Pearson correlation coefficient is $r = .770$ and is significant at P (one-tailed) $< .001$.

The probability of getting correlation coefficients this size when there is in fact no relation is close to zero and so we may conclude that both dependent variables have strong relationships with their independent variable, Internet users per 100 people in the Netherlands.

4.3.4. Using R² for interpretation²⁰

We now know that digitisation and diversity are correlated with each other. To what extent? What can we say about how much variation in diversity is shared by the development of digitisation? If we square the *r*-values we will get the coefficient of determination (Field, 2009: 179; Bryman, 2009: 329) R². This measure tells us the amount of variation in one variable shared by the other. In our case R² for unique artists is $(.813)^2 = .661$. We saw this value earlier in figure 4.3. Although we cannot say anything about the direction of causality, we can say that 66,1 per cent of the variation in diversity is shared with the variation in digitisation. That leaves 33,9 per cent unexplained. There must be other factors which influence diversity in the charts. Still, we can conclude that digitisation has a large effect on how many different artists show up in the charts on a yearly basis.

When we look at the relationship between digitisation and the amount of unique albums, we notice that the coefficient is slightly less. Still, $R^2 = (.770)^2 = .592$. Digitisation and the albums share 59,2 per cent of the variation. Here 40,8 per cent remains unexplained.

4.3.5. Correlation Digitisation and Diversity in the period 2000 - 2009

Figures 4.1, 4.2 and 4.3 hinted to a somewhat perfect linear trend in the relation between IU and UA and UAL in the years from 2000 until 2009. I do like to test this 2000 - 2009 period as well to see how large the shared variance is between the variables. Furthermore, I will also test the period 1994-1999 separately.

Internet users per 100 people (2000-2009) is significantly correlated with UA, $r = .964$, P (one-tailed) $< .001$. $R^2 = .929$. This result implies that these variables share 92,9 of their variance. This is a very high score, and maybe a bit unrealistic. This result leaves no room for other factors that could influence diversity.

Also, the amount of different album titles is significantly related in 2000-2009 with IU, $r = .945$, P (one-tailed) $< .001$. $R^2 = .893$.²¹

So the last 10 years of the researched period the variables show a nearly perfect linear relationship with each other. The high scores for R² suggest that there are not many other factors responsible for growth of diversity in the charts or that they have a very small impact. We must however stay cautious to not over-interpret these results.

Where the 2000-2009 period showed significant correlations, the period 1994-1999 attests to the opposite. IU is not significantly correlated with UA, $r = .104$, sig (one-tailed) = .422. IU is not significantly correlated as well with UAL, $r = .054$, sig (one-tailed) = .460.²²

It appears that the post digitisation period is characterized with two conflicting dynamics. The first period (1994-1999) shows no significant influence from digitisation upon diversity. The second period 2000-2009 is to the contrary very significantly correlated. These two periods together make that the overall influence of IU on UA and UAL is still strong as reported in my 1994-2009 analysis.

²⁰ This subtitle is borrowed from Field (2009: 179)

²¹ See SPSS output 8

²² See SPSS output 9.

The effect of digitisation is in the beginning absent, but as time passes its significance grows steadily. It seems that digitisation had to reach a critical value before it became a major influence on diversity. This value is around forty Internet users per 100 people in the Netherlands. In 1999 its value is under 40, in 2000 it is 43,95. 1999 is notorious for the advent of illegal file-sharing networks. (Handke, 2010: 58) Digital copying and distribution became that year accessible to the public at large. Whether or not these networks harm record sales is not the question in the paper. However, it looks like digitisation, apart from its alleged adverse effects to turnover in the market, has also a positive side. There are reasons to believe that this development made a more diversified spectrum of music available to the public.

4.4. Summary

My hypothesis - Digitisation leads to more diversity in the charts - cannot be proved by using statistics alone. However, statistics could provide us with useful tools to gain more insights. The highest attainable achievement would be here that I find evidence to support my hypothesis. I can do this by rejecting the null hypothesis connected to my alternative hypothesis: Digitisation will not lead to any change in diversity.

I can reject this null hypothesis, and so I will present support for my alternative hypothesis. (Field, 2009: 27) Based on my findings, the null hypothesis can be rejected. I found strong relationships between digitisation and the variables which indicate diversity. The amount of Internet users in the Netherlands account for respectively 66,1 per cent and 59,2 percent of the variation in the number of different artists and the number of different albums in the top 50 of the Album top 100.

Any statements about the direction of causality I cannot infer from the statistical results. Also, there is always the problem of the third variable: there are maybe other factors which cause more diversity in the album charts in the Netherlands.

5. Conclusion

In this paper I studied the relation between digitisation and diversity in the supply of record in the Netherlands. I tried to find support for the following hypothesis: "Digitisation leads to more diversity in the supply of records in the Netherlands". My test results show that digitisation has a strong relationship with diversity. Yet, any statements about whether digitisation has caused more diversity in the supply of records I cannot make. It seems plausible that digitisation is causing the found increase in diversity however. The opposite conclusion, more diversity will lead to more digitisation, seems rather unrealistic. As my indicator for digitisation was Internet users per 100 people it is hard to see why more choice in music would cause more Internet users.

My study was deductive and began from theory. I used the theory of product searches of Nelson (1970), the 'frictionless market' hypothesis (Brynjolfsson & Smith, 2000) and the Long-tail theory of Anderson (2004). All these theories served my purpose. They were all connected to the Internet. Since I used Internet users as a proxy for digitisation my theoretical framework and empirical part could be nicely fused together.

To generate my results I used basic statistical tests, such as correlation analysis and simple linear regression. The assumptions underlying the employed tests are met by the data, so results from the tests are to be robust and should give a reasonable representation of the data.

There are several shortcomings in my research. I did not have a control group at my disposal. Although I have tried to control for any pre digitisation upward trends in diversity we will never know whether diversity would have increased regardless of digitisation. Furthermore, using the charts have clear benefits, but also disadvantages. In this study real diversity will likely to be underestimated since the charts only capture the best selling fracture of the market. There must have been undoubtedly more artists and albums published the previous 31 years. Still, this study was able to identify a clear positive trend in the diversity. This trend exists, even if we do not know the absolute numbers attached to it.

The nature of the data for diversity did not allow me to measure diversity in a very sophisticated way. I could only touch upon the first layer of the multi-dimensional definition of cultural diversity: variety. This paper investigated only the sheer multiplicity of types. Another shortcoming is that I was not able to investigate the Album top 100 entirely, but only the top 50 of these chart. This was due the different compositions throughout the years.

5.1 Suggestions for further research.

My study is one of the first studies that relate diversity to digitisation. The first results are in, but we are not there yet. Future studies should try to find a way to measure more dimensions of cultural diversity. In addition to sheer variety, balance and diversity are still underexposed. Also, at the other dimensions such as producer diversity progress can be made.

Furthermore, finding different sources of data could benefit the diversity analysis. Charts are a good indicator of the market, but much diversity remains hidden. Researchers should employ these other data sources to increase our understanding of the topic and the current state of cultural diversity. Other researchers may also look into other factors which influence diversity other the digitisation.

To data we can only guess what influences there are on diversity. Also, about the extent of their impact we have only faint ideas. More substantial future research projects should concentrate on multi variate analysis of this subject matter.

Ultimate goal is to enrich our knowledge of diversity. With better knowledge we can make better decisions. After all, “Everyone has the right freely to participate in the cultural life of the community, to enjoy the arts, and to share in scientific advancement and its benefits”.

References

- Acheson, K.K. 2004. Convention on cultural diversity. *Journal of Cultural Economics* 28 (4), 243-256.
- Achterberg, P. 2011. A cultural globalization of popular music? american, dutch, french, and german popular music charts (1965 to 2006). *The American Behavioral Scientist* (Beverly Hills) 55 (5), 589-608.
- Akerlof, G.A. 1970. The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics* , 488.
- Allain, M. & Waelbroeck, P. 2006. 'Music Variety and Retail Concentration', Working Paper ESS-06-06, ENST
- Baker, A.A.J. 1991. A model of competition and monopoly in the record industry. *Journal of Cultural Economics* 15 (1), 29-53.
- Bala, V. 2005. International trade and cultural diversity with preference selection. *European Journal of Political Economy* 21 (1), 143.
- Baumol, W.J. 1982. Contestable markets: An uprising in the theory of industry structure. *The American Economic Review* 72 (1), 1.
- Benhamou, F. and S. Peltier. 2007. 'How should cultural diversity be measured? An application using the French publishing industry', *Journal of Cultural Economics* 31(2); pp. 85-107.
- Benhamou, F., Flôres, R.G., Peltier, S. 2009. Diversity analysis in cultural economics: theoretical and empirical considerations. Document de travail du GRANEM n° 2009-05-017
- Burnett, R.R. 1992. The implications of ownership changes on concentration and diversity in the phonogram industry. *Communication Research* 19 (6), 749-769.
- Bryman, A.B. 2009. *Social Research Method*. Oxford University Press: New York.
- Chung, K.H. 1994. A stochastic model of superstardom: An application of the yule distribution. *The Review of Economics and Statistics* , 771.
- Curtis, J.J.S. 2006. Culture and the digital copyright chimera: Assessing the international regulatory system of the music industry in relation to cultural diversity. *International Journal of Cultural Property* 13 (01), 59.
- Daniel, E. 1999. The impact of electronic commerce on market structure: An evaluation of the electronic market hypothesis. *European Management Journal* 17 (3), 318.
- Darby, M.R. 1973. Free competition and the optimal amount of fraud. *The Journal of Law & Economics* 16 (1), 67.
- Dixit, A.K. 1977. Monopolistic competition and optimum product diversity. *The American Economic Review* 67 (3), 297.
- Dyen, I., Kruskal J.B., & Black P. (1992). An Indo-European Classification: a Lexicostatistical Experiment, *Transactions of the American Philosophical Society*, 82(5), Philadelphia: American Philosophical Society
- Field, A. 2009. *Discovering Statistics Using SPSS*. Third Edition. Sage: London.
- Flôres, R.G. 2006. The Diversity of Diversity: further methodological considerations on the use of the concept in cultural economics.

- Giles, D. 2007. Survival of the hippest: Life at the top of the hot 100. *Applied Economics* 39 (15), 1877-1887.
- Gopal, R.R.D. 2006. Do artists benefit from online music sharing?*. *The Journal of Business* (Chicago, Ill.) 79 (3), 1503-1533.
- Hamlen Jr, W.A. 1994. Variety and superstardom in popular music. *Economic Inquiry* 32 (3), 395.
- Handke, C.W. 2010. The Creative Destruction of Copyright Innovation in the Record Industry and Digital Copying. Doctoral dissertation, EUR.
- Hoch, S.J. & Ha, Y. 1986. Consumer learning: Advertising and the ambiguity of product experience. *The Journal of Consumer Research* 13 (2), 221.
- Hotelling, H. 1929. Stability in competition. *The Economic Journal* (London) 39 (153), 41.
- Isar, Y. 2006. Cultural diversity. *Theory, Culture & Society* 23 (2-3), 372-375.
- Kawashima, N.N. 2011. Are the global media and entertainment conglomerates having an impact on cultural diversity? A critical assessment of the argument in the case of the film industry. *International Journal of Cultural Policy* : CP 99999 (1), 1-15.
- Klein, L.R. 1998. Evaluating the potential of interactive media through a new lens: Search versus experience goods. *Journal of Business Research* 41 (3), 195.
- Klein, L.R. 1998. Evaluating the potential of interactive media through a new lens: Search versus experience goods. *Journal of Business Research* 41 (3), 195.
- Kretschmer, M.M. 1999. Increasing returns and social contagion in cultural industries. *British Journal of Management* 10 (s1), 61-72.
- Lancaster, K. 1990. The economics of product variety: A survey. *Marketing Science* (Providence, R.I.) 9 (3), 189.
- Lertwachara, K. 2007. Stochastic dynamics of music album lifecycle: An analysis of the new market landscape. *International Journal of Human-Computer Studies* 65 (1), 85-93.
- Lopes, P.D. 1992. Innovation and diversity in the popular music industry, 1969 to 1990. *American Sociological Review* 57 (1), 56.
- Malone, T.T.W. 1987. Electronic markets and electronic hierarchies. *Communications of the ACM* 30 (6), 484-497.
- Mas-Colell, A.A. 1999. Should cultural goods be treated differently? *Journal of Cultural Economics* 23 (1/2), 87-93.
- Moreau, F.F. 2004. Cultural diversity in the movie industry: A cross-national study. *Journal of Media Economics* 17 (2), 123-143.
- Morton Fiona Scott, P.P. 2009. Searching for experience on the web: An empirical examination of consumer behavior for search and experience goods. *Journal of Marketing* 49 (2), 55-69.
- Nelson, P. 1974. Advertising as information. *The Journal of Political Economy* 82 (4), 729.
- Peterson, R.A. 1975. Cycles in symbol production: The case of popular music. *American Sociological Review* 40 (2), 158.
- Peterson, R.A. 1975. Cycles in symbol production: The case of popular music. *American Sociological Review* 40 (2), 158.
- Peterson, R.A. 1996. Measuring industry concentration, diversity, and innovation in popular music. *American Sociological Review* 61 (1), 175.
- Ranaivoson, H. 2007. Measuring cultural diversity: a review of existing definitions.

- Rothenbuhler, E.E.W. 1982. Popular music: Concentration and diversity in the industry, 1974? 1980. *Journal of Communication* 32 (1), 143-149.
- Schatz, B.B.R. 1994. NCSA mosaic and the world wide web: Global hypermedia protocols for the internet. *Science (New York, N.Y.)* 265 (5174), 895-901.
- Simpson, E.H. 1949. Measurement of diversity. *Nature (London)* 163 (4148), 688.
- Stigler, G.J. 1961. The economics of information. *The Journal of Political Economy* 69 (3), 213.
- Stigler, G.J. 1962. Information in the labor market. *The Journal of Political Economy* 70 (5), 94.
- Stirling, A. 1999. On the Economics and Analysis of Diversity. *Electronic Working Papers Series Paper No. 28*.
- Weitzman, M.M.L. 1992. On diversity. *The Quarterly Journal of Economics* 107 (2), 363.
- Weitzman, M.M.L. 1993. What to preserve? an application of diversity theory to crane conservation. *The Quarterly Journal of Economics* 108 (1), 157.
- Wright, A.A. 1995. Communication effects of advertising versus direct experience when both search and experience attributes are present. *The Journal of Consumer Research* 21 (4), 708.
- Wurff, R.J.W. 2001. Impact of moderate and ruinous competition on diversity: The dutch television market. *Journal of Media Economics* 14 (4), 213.

Appendix 1: SPSS output

1.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics			
					R Square Change	F Change	df1	df2
1	.161 ^a	.026	-.063	4.610	.026	.291	1	11

a. Predictors: (Constant), Unique Artists in the top 50 per year in the Album top 100

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6.184	1	6.184	.291	.600 ^a
	Residual	233.816	11	21.256		
	Total	240.000	12			

a. Predictors: (Constant), Unique Artists in the top 50 per year in the Album top 100

b. Dependent Variable: YearREG

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	15.428	13.830		1.116	.288
	Unique Artists in the top 50 per year in the Album top 100	-.035	.065	-.161	-.539	.600

a. Dependent Variable: YearREG

2.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics			
					R Square Change	F Change	df1	df2
1	.027 ^a	.001	-.090	4.669	.001	.008	1	11

a. Predictors: (Constant), Unique Albums in the top 50 per year in de Album top 100

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.175	1	.175	.008	.930 ^a
	Residual	239.825	11	21.802		
	Total	240.000	12			

a. Predictors: (Constant), Unique Albums in the top 50 per year in de Album top 100

b. Dependent Variable: YearREG

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	9.199	13.428		.685	.507
	Unique Albums in the top 50 per year in de Album top 100	-.004	.044	-.027	-.090	.930

a. Dependent Variable: YearREG

3.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics			
					R Square Change	F Change	df1	df2
1	.860 ^a	.739	.721	2.516	.739	39.714	1	14

a. Predictors: (Constant), Unique Artists in the top 50 per year in the Album top 100

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	251.382	1	251.382	39.714	.000 ^a
	Residual	88.618	14	6.330		
	Total	340.000	15			

a. Predictors: (Constant), Unique Artists in the top 50 per year in the Album top 100

b. Dependent Variable: YearREG

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-8.532	5.122		-1.666	.118
	Unique Artists in the top 50 per year in the Album top 100	.132	.021	.860	6.302	.000

a. Dependent Variable: YearREG

4.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics			
					R Square Change	F Change	df1	df2
1	.823 ^a	.678	.655	2.796	.678	29.494	1	14

a. Predictors: (Constant), Unique Albums in the top 50 per year in de Album top 100

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	230.559	1	230.559	29.494	.000 ^a
	Residual	109.441	14	7.817		
	Total	340.000	15			

a. Predictors: (Constant), Unique Albums in the top 50 per year in de Album top 100

b. Dependent Variable: YearREG

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-15.120	7.146		-2.116	.053
	Unique Albums in the top 50 per year in de Album top 100	.135	.025	.823	5.431	.000

a. Dependent Variable: YearREG

5.

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Internet users (per 100 people)	.143	16	.200*	.902	16	.086
Unique Artists in the top 50 per year in the Album top 100	.191	16	.123	.902	16	.088
Unique Albums in the top 50 per year in de Album top 100	.114	16	.200*	.965	16	.753

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

6.

Correlations

		Internet users (per 100 people)	Unique Artists in the top 50 per year in the Album top 100	Unique Albums in the top 50 per year in de Album top 100
Internet users (per 100 people)	Pearson Correlation	1	.813**	.770**
	Sig. (1-tailed)		.000	.000
	N	16	16	16
Unique Artists in the top 50 per year in the Album top 100	Pearson Correlation	.813**	1	.973**
	Sig. (1-tailed)	.000		.000
	N	16	16	16
Unique Albums in the top 50 per year in de Album top 100	Pearson Correlation	.770**	.973**	1
	Sig. (1-tailed)	.000	.000	
	N	16	16	16

** . Correlation is significant at the 0.01 level (1-tailed).

7.

Correlations

			Internet users (per 100 people)	Unique Artists in the top 50 per year in the Album top 100	Unique Albums in the top 50 per year in de Album top 100
Kendall's tau_b	Internet users (per 100 people)	Correlation Coefficient	1.000	.678**	.695**
		Sig. (1-tailed)	.	.000	.000
		N	16	16	16
	Unique Artists in the top 50 per year in the Album top 100	Correlation Coefficient	.678**	1.000	.941**
		Sig. (1-tailed)	.000	.	.000
		N	16	16	16
	Unique Albums in the top 50 per year in de Album top 100	Correlation Coefficient	.695**	.941**	1.000
		Sig. (1-tailed)	.000	.000	.
		N	16	16	16
Spearman's rho	Internet users (per 100 people)	Correlation Coefficient	1.000	.834**	.821**
		Sig. (1-tailed)	.	.000	.000
		N	16	16	16
	Unique Artists in the top 50 per year in the Album top 100	Correlation Coefficient	.834**	1.000	.987**
		Sig. (1-tailed)	.000	.	.000
		N	16	16	16
	Unique Albums in the top 50 per year in de Album top 100	Correlation Coefficient	.821**	.987**	1.000
		Sig. (1-tailed)	.000	.000	.
		N	16	16	16

** . Correlation is significant at the 0.01 level (1-tailed).

It will be instructive to check whether the other two correlation coefficients, Kendall's τ and Spearman's r give the same image as the analysis with Pearson's s . The choice of correlation for a certain correlation coefficient can make a difference in the size of the effect. (Field, 2009: 193). I ran the same bivariate correlation test for these two indicators of relationships between variables. It is however not required to do this extra analysis, but I think that this gives a better picture of the relationships and it also verifies the results from the Pearson's correlation coefficient. Maybe, the test will bring up interesting results. The two tests are non-parametric and therefore often "back-up" tests if data is not measured at an interval level or not normally distributed. I did not need the specific advantages of Spearman's r - useable with ranked ordinal and non-normal data - and those of Kendall's τ - useable with a small dataset with a large number of tied ranks (Field, 2009: 181). Whereas the correlation coefficient for unique artists in the charts with Pearson's r $R^2 = .661$ was higher than the one for unique albums $R^2 = .592$, Kendall's tau test results in the opposite. $\tau = .678$, $P < (\text{one-tailed}) .001$, for the relation between digitisation and unique artists, while its higher $\tau = .695$, $P < .001$ for digitisation and different albums in the top 100. Kendall's τ does not however tell us anything about the amount of shared variance. Spearman's r_s values show the same image as Pearson's r : a larger part of variation in diversity is explained by digitisation for artists than albums. The values are $r_s = .834$, $P < (\text{one-tailed}) .001$ for artists and $r_s = .821$, $P < (\text{one-tailed}) .001$ for the unique number of albums.

As with Pearson's r , we could square r_s to see how much variation the variables share. Spearman's r reports about the same relationship between digitisation and diversity, albeit a bit stronger. The value for the correlation between digitisation and the artists in the charts is $r_s^2 = (.834)^2 = .695$ and for digitisation and the amount of different albums in the charts $r_s^2 = (.821)^2 = .674$. Over sixty per cent of the variation in diversity is explained by digitisation, according to this statistic. But as Field (2009: 181) notes: Kendall's τ is actually a better estimate.

8.

Correlations

		Internet users (per 100 people)	Unique Artists in the top 50 per year in the Album top 100	Unique Albums in the top 50 per year in de Album top 100
Internet users (per 100 people)	Pearson Correlation	1	.964**	.945**
	Sig. (1-tailed)		.000	.000
	N	10	10	10
Unique Artists in the top 50 per year in the Album top 100	Pearson Correlation	.964**	1	.986**
	Sig. (1-tailed)	.000		.000
	N	10	10	10
Unique Albums in the top 50 per year in de Album top 100	Pearson Correlation	.945**	.986**	1
	Sig. (1-tailed)	.000	.000	
	N	10	10	10

** . Correlation is significant at the 0.01 level (1-tailed).

9.

Correlations

		Internet users (per 100 people)	Unique Artists in the top 50 per year in the Album top 100	Unique Albums in the top 50 per year in de Album top 100
Internet users (per 100 people)	Pearson Correlation	1	.104	.054
	Sig. (1-tailed)		.422	.460
	N	6	6	6
Unique Artists in the top 50 per year in the Album top 100	Pearson Correlation	.104	1	.994**
	Sig. (1-tailed)	.422		.000
	N	6	6	6
Unique Albums in the top 50 per year in de Album top 100	Pearson Correlation	.054	.994**	1
	Sig. (1-tailed)	.460	.000	
	N	6	6	6

** . Correlation is significant at the 0.01 level (1-tailed).

Appendix 2: data set

Year	IU	UA	UAL
1979		170	247
1980		205	304
1981		223	327
1982		245	336
1983		239	352
1984		207	273
1985		214	285
1986		204	294
1987		225	317
1988			
1989			
1990	0,334403	208	297
1991	0,530856	208	325
1992	1,317697	182	317
1993	1,963467	205	261
1994	3,250736	208	249
1995	6,468305	228	277
1996	9,666817	205	239
1997	14,096239	243	301
1998	22,295834	226	270
1999	39,228092	217	259
2000	43,954854	209	255
2001	49,233174	225	277
2002	61,044439	229	280
2003	64,101873	240	284
2004	68,351755	243	287
2005	80,772562	265	298
2006	82,828125	282	314
2007	86,229228	300	336
2008	87,856072	287	326
2009	89,964161	285	327