Coneus, Katja; Schleife, Katrin

Working Paper

Online but still divided: Inequality in private internet use in Germany

ZEW Discussion Papers, No. 10-042

Provided in cooperation with:
Zentrum für Europäische Wirtschaftsforschung (ZEW)

Suggested citation: Coneus, Katja; Schleife, Katrin (2010) : Online but still divided: Inequality in private internet use in Germany, ZEW Discussion Papers, No. 10-042, http://hdl.handle.net/10419/36387

Discussion Paper No. 10-042

Online But Still Divided – Inequality in Private Internet Use in Germany

Katja Coneus and Katrin Schleife

ZEW
Zentrum für Europäische Wirtschaftsforschung GmbH
Centre for European Economic Research

Discussion Paper No. 10-042

Online But Still Divided – Inequality in Private Internet Use in Germany

Katja Coneus and Katrin Schleife

Non-technical summary

Although information and communication technologies (ICT) have diffused into all economic and private sectors during the last decade, large disparities still exist. The disparities between individuals with regard to their access and use of new technologies are facets of the so-called digital divide. Until today, research on the digital divide has mainly focused on the access decision of individuals, that is on gaps between those who do and those who do not have access to ICT (first-level digital divide), and less on differences in usage patterns (second-level digital divide).

In this paper, we investigate inequalities in internet use behaviour for the year 2004. Our analysis is based on the German ALLBUS data set which allows us to consider specific individual attitudes and preferences towards ICT. We focus on two different aspects: First, we estimate the impact of socio-economic factors on the internet access decision (first-level digital divide). Second, we analyse the determinants of the intensity of internet use and the frequency of various internet activities, such as downloading from the internet and creating web pages (second-level digital divide).

Comparing the determinants of the first- and the second-level digital divide, our empirical results show two important patterns: First, socio-economic characteristics, like education, age and migration explain inequalities in internet access, while for the second-level digital divide, socio-economic factors hardly seem to play a role. Second, preferences and attitudes regarding new technologies as well as peer effects turn out to be important for both.

Das Wichtigste in Kürze

Obwohl die private Nutzung von Informations- und Kommunikationstechnologien (IKT) in den letzten Jahren weiter zugenommen hat, gibt es bis heute Ungleichheiten sowohl hinsichtlich des Zugangs zu IKT (first-level digital divide) als auch hinsichtlich des IKT-Nutzungsverhaltens (second-level digital divide). Während sich die ökonomische Literatur bisher intensiv mit Fragen zum 'first-level digital divide' beschäftigt hat, sind Untersuchungen, die Unterschiede im IKT-Nutzungsverhalten von Individuen betrachten, bisher selten.

In dieser Arbeit untersuchen wir für das Jahr 2004 sowohl Determinanten des first- als auch des second-level digital divide. Unsere empirische Analyse basiert auf den Daten der Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS) für Deutschland, die zahlreiche individuelle Informationen zur IKT-Nutzung und zur Einstellung gegenüber neuen Technologien beinhaltet.

Die Analyse der Unterschiede und Gemeinsamkeiten zwischen den Determinanten des first- und des second-level digital divide in dieser Studie liefert zwei zentrale Ergebnisse: Erstens zeigt sich, dass sozio-ökonomische Faktoren wie Bildung, Alter und Migration Ungleichheiten im IKT-Zugang erklären – nicht jedoch in der Nutzungsart und -intensität. Zweitens erweisen sich Einstellungen und Präferenzen hinsichtlich IKT sowie der Einfluss des Freundeskreises sowohl für den first-, als auch für den second-level digital divide als von entscheidender Bedeutung. Sie tragen somit zur Erklärung von Ungleichheiten im Zugang wie auch in den genutzten Anwendungen bei.

Online But Still Divided – Inequality in Private Internet Use in Germany

Katja Coneus*, Katrin Schleife†

June 17, 2010

Abstract

In this analysis, we compare the determinants of the first-level and the second-level digital divide for private internet use in Germany. Our work offers three important innovations. First, we use the exact weekly duration of internet use to explain inequalities in internet intensity, explicitly controlling for non-users. Secondly, we use the frequencies of five different internet applications to further investigate the determinants of the second-level digital divide. Thirdly, we estimate selection models to control for unobserved characteristics of all individuals. Comparing the determinants of the first- and second-level digital divide shows that socio-economic characteristics (age, education, migration) explain inequalities of the first- but not of the second-level digital divide. By contrast, preferences and attitudes regarding new technologies as well as peer effects turn out to be important for both.

Keywords: digital divide, internet use, peer effects, inequality.
JEL-classification: I1, I12, J13

*Centre for European Economic Research, Mannheim. E-mail: coneus@zew.de

†Berlecon Research, Am Kupfergraben 6a, D-10117 Berlin, Germany. E-mail address: ks@berlecon.de. This article was prepared during the authors' work in the ICT Research Group at the Centre for European Economic Research (ZEW), Mannheim.

1 Introduction

Although information and communication technologies (ICT) have diffused into all economic and private sectors during the last decade, large disparities still exist. The disparities between individuals with regard to their access and use of new technologies are facets of the so-called digital divide.

The digital divide is a complex and continuously altering phenomenon which needs to be carefully examined. The OECD defines it as "the gap between individuals, households, business and geographic areas at different socio-economic levels with regard to both their opportunities to access information and communication technologies (ICTs) and to their use of the internet for a wide variety of activities" (OECD/DSTI, 2001, p. 5). This definition already points to the fact that the term digital divide does not describe a simple gap between two clearly defined groups but that there are many facets incorporated in it.

Until today, research on the digital divide has mainly focused on the access decision of individuals, that is on gaps between those who do and those who do not have access to ICT ("first-level digital divide"), when analyzing the impact of ICT on social and economic inequality. Significant differences in internet access between social groups are examined, for example, by Lenhart et al. (2003), Korupp and Szydlik (2005), and Schleife (2010). They show that access barriers might depend on economic and non-economic constraints, such as income, education, or age. On average, internet users turn out to be younger, more highly educated, employed, and to have a higher income.

However, during the last years, having internet access has become much more common even in economically and educationally disadvantaged groups. The recent (N)onliner Atlas, for example, observes a higher-than-average increase in the proportion of onliners for individuals with a lower educational degree in Germany, now reaching a value of more than 50 percent in 2009 (TNS Infratest, 2009). A large increase can also be observed for people aged 60 or older, a result originating from two important developments: First, former middle-aged users grow older and second, hard- and software becomes increasingly user-friendly. However, for several population groups, such as the

older and the less educated people, the proportion of internet users is still considerably below the German average of 69 percent. [1]

There have always been concerns whether having internet access alone can guarantee social and economic participation. Usage patterns, that is the duration and especially the used applications of ICT, have become at least as relevant for the individuals' opportunities to participate in the information and knowledge society. For example, today, accessing and selecting information, services, and tools of education on the internet, as well as dealing with advanced software applications, has become extremely important. But these tasks demand specific IT skills and usage experience. Differences in ICT usage patterns due to socio-economic characteristics and factors like motivation, IT skills, internet literacy, and peer effects have become known as the "second-level digital divide".

While the gap between individuals in accessing ICT is narrowing, large differences exist in how intensively the internet and the many new online contents are used when access is provided. A German study in 2004 shows that about half of the German population aged 14 or older uses the internet only sporadically and selectively (van Eimeren et al., 2004). Looking at Internet users there are in 2009 small differences between age groups in the proportion of people who are searching information online (around 95 percent) or those who are exchanging e-mails (around 90 percent)(Statistisches Bundesamt, 2009). Larger age-specific differences are observable for proportions of users providing own online content (text, pictures, videos) or participating in blogs and forums. This is more often done by younger age groups. In addition, those aged between 14 and 29 spend much more time per day on the internet than any other age group (van Eimeren and Frees, 2009).

For the US, the Falling Through the Net study finds large differences in internet usage by education, income, migration background, and other characteristics. The descriptive statistics show that with rising educational levels, internet activities like searching for information, doing job-related tasks, searching for jobs, and exchanging e-mail increase (van Dijk and Hacker, 2003). They additionally find that income seems to be a less important determinant of performing specific internet activities than the educational background.

---

[1] Moreover, it has to be kept in mind that it is not only the difference between, but also within groups that can be significant. Even among wealthy and well educated strata, there are still people who refrain from using the internet. And although today children are often seen to be part of the 'internet generation' and even called 'the digital natives' there are still many of them who, on the one hand, do not access the internet at all (or in school only) or who, on the other hand, use the internet much less or for other activities than their peers. The digital divide among children and young people is analyzed, for example, by Livingstone and Helsper (2007).

Policy makers and the public are highly interested in lowering access constraints to the internet because it becomes more and more relevant for social interactions or the exchange of knowledge. In addition, internet literacy becomes a crucial factor, as it is not only access and the time spent on the internet that is important, but especially what the internet is used for. Less educated individuals are more likely to fail keeping up with technological developments because individuals who have already gained experience in ICT use may have an advantage in efficiently applying and using newly emerging technologies. On average, individuals with a higher level of education begin using ICT earlier and therefore become more experienced users.

This paper addresses the issue of inequalities in internet use behaviour. We focus on two different aspects: First, we estimate the impact of socio-economic factors on the intensity of internet use. Secondly, we analyze the determinants of various internet activities, such as the frequency of downloads from the internet or the frequency of creating web pages.

Thus, we extend the previous research activities on the first-level digital divide in Germany by showing differences and similarities between factors influencing the first-level and the second-level divide. To our knowledge, this has scarcely been done by researchers so far. A literature overview about research on the different facets of the digital divide is provided by DiMaggio et al. (2004). However, this study finds that while the determinants of internet access are already studied by many authors, econometrically analyzing the factors influencing the differences in internet use intensity and activity is less common. Our paper aims at closing this gap.

The analyses are based on the German ALLBUS [2] data set of the year 2004. It is one of the few data sets available that provide representative data on aspects of the second-level digital divide.

We support prior findings that socio-economic characteristics like education, age, income, family composition, and migration are important determinants of the first-level digital divide and that attitudes and leisure preferences also have an impact. In contrast, inequalities in the second-level digital divide (internet intensity and activities) arise mostly due to differences in attitudes toward ICT and new technologies, while age, education, family composition, and migration background only play a minor role. The number of friends using the internet turns out to be an important determinant of both first- and second-level digital divide.

The rest of the paper is organized as follows: In the next section, we explain the data used for our analysis and present descriptive statistics. Section 3 presents the empirical methods. In section 4, we econometrically analyze

[2] Allgemeine Bevölkerungsumfrage der Sozialwissenschaften (German General Social Survey).

the determinants of the first- and second-level digital divide, and section 5 concludes.

2 Data

In order to compare the first- and second-level digital divide in Germany, detailed information on internet use, especially regarding the time spent on the internet and the frequency of various internet activities, is derived from the ALLBUS data set of 2004. Besides internet-related variables, the representative survey provides comprehensive information on demographic and socio-economic characteristics as well as on individual attitudes and behaviour of residents in Germany. [3] Our working sample comprises about 2,000 individuals aged 18 years and older.

Table 1 shows that in Germany in 2004, every second individual does not use the internet at home. [4] Among those individuals who use the internet, we are interested in two specific issues regarding the second-level digital divide: In a first step, we examine the intensity of internet use. It is defined as the time spent on the internet at home and is measured in hours per week. It ranges between 0 and 70 hours. As can be seen in Table 1, individuals in Germany use the internet nearly two hours per week on average. [5] If only internet users are taken into account, the average time spent on the internet is more than four hours per week (not presented in Table 1).

In a second step, we take a closer look at the individuals' frequency of several activities that can be performed on the internet and are covered by the survey. These activities comprise 'getting and sending e-mails', 'searching for information on the web', 'creating web pages', 'downloading content from the internet', and 'playing online games'. The frequencies are ranked on a scale of one to five, where 1 = 'never', 2 = 'infrequent', 3 = 'at least once in a month', 4 = 'at least once in a week', 5 = 'daily'. Table 1 shows that individuals who use the internet, most frequently use it for information search and e-mail exchange. The other activities are less frequent.

Table 2 gives an overview about the attitudes towards ICT, leisure preferences, the amount of friends using the internet, demographic and socio-economic characteristics for all individuals in our sample as well as differentiated between internet users and non-users.

---

[3] ALLBUS is conducted by the GESIS institution every two years. In every wave, data on various topics are collected. In 2004, one of these main topics was ICT usage behavior and technology attitudes. Unfortunately, individuals are not repeatedly observed in ALLBUS.

[4] It can be assumed that most of them have no internet access at all.

[5] See Figure 1 in section 3.2 for the distribution of the internet intensity.

Table 1
Descriptives – Internet intensity and activities

	mean	std.dev.	min	max
internet use (N=2,013)				
internet use (1=yes)	0.54	0.50	0	1
duration of internet use (hours/week)	1.85	4.53	0	70
activities on the internet (N=886)				
e-mail^{i}	3.67	1.31	1	5
information search^{i}	3.55	1.05	1	5
internet downloads^{i}	1.44	0.89	1	5
creating web pages^{i}	1.28	0.75	1	5
internet games^{i}	1.26	0.68	1	5

Source: ALLBUS 2004. Own calculations.
Notes: ^{i}): never, infrequent, at least 1x/month, at least 1x/week, daily.

Considering the attitudes towards ICT, which are measured on a scale between 1 (strongly agree) and 7 (strongly disagree), it can be shown that users significantly more often agree that they are interested in technology and in computers than non-users. The amount of friends who are online, a variable which might reflect peer effects, is measured on a scale between one and four with 1= nobody, 2= some, 3= many, 4= almost all of the friends use the internet. Internet users have significantly more friends using the internet (3.16) than non-users (2.11). It is interesting to see that a proportion of 90 percent of all individuals declare that at least 'some' of their friends use the internet (not shown in the table).

As the propensity to accept and adopt new technological trends varies between generations, we follow the age differentiation of Sackmann and Weymann (1994, p. 9) which specifies four types of generations, depending on phases of technological development: the 'pre-technical generation' (born before 1939), the 'generation of the household revolution' (born between 1939 and 1948), the 'generation of advanced household technology' (born between 1949 and 1964), and the 'computer generation' (born after 1964). In our sample of all individuals, 36 percent belong to the computer generation, 17 percent to the oldest, the pre-technical generation. The share of individuals belonging to the computer generation is more than twice as high among users as among non-users.

Similar to the total German population, half of the sample is female and six percent are foreigners. Most of the individuals in our sample left middle school or high school and got a vocational degree as highest educational degree (70 percent). 10 percent completed university, while 21 percent have no

Table 2
Descriptive statistics – Explaining variables

	all		internet users		internet non-users	
	mean	std.dev.	mean	std.dev.	mean	std.dev.
amount of friends using the internet[i]	2.68	0.97	3.16*	0.83	2.11*	0.81
no interest in computers[ii]	2.82	2.11	1.91*	1.51	3.90*	2.20
interested in technology[ii]	4.27	1.94	4.73*	1.80	3.71*	1.95
technological development is too fast[ii]	3.68	2.05	2.91*	1.72	4.59*	2.04
age <40 years	0.36	0.48	0.52*	0.50	0.17*	0.38
age 40-55 years	0.31	0.46	0.36*	0.48	0.25*	0.43
age 56-65 years	0.16	0.36	0.09*	0.29	0.24*	0.43
age >65 years	0.17	0.37	0.02*	0.15	0.34*	0.47
female	0.50	0.50	0.45*	0.50	0.56*	0.50
foreign	0.06	0.24	0.05*	0.21	0.08*	0.28
no vocational degree	0.21	0.40	0.15*	0.35	0.28*	0.45
vocational degree	0.60	0.49	0.57*	0.50	0.65*	0.48
A-level & vocational degree	0.09	0.29	0.13*	0.34	0.04*	0.20
university	0.10	0.30	0.16*	0.36	0.04*	0.19
employed	0.51	0.50	0.71*	0.46	0.28*	0.45
monthly net income of household (in 1,000 EUR)	2.28	1.42	2.69*	1.59	1.79*	0.98
single adult in household[iii]	0.19	0.39	0.17*	0.38	0.22*	0.41
number of children in household [iv]	0.50	0.87	0.62*	0.91	0.35*	0.81
East Germany	0.32	0.47	0.29*	0.45	0.36*	0.48
urban region[v]	0.92	0.27	0.94*	0.24	0.90*	0.30
reading books[vi]	3.11	1.38	2.91*	1.33	3.35*	1.41
sports activities[vi]	3.39	1.43	2.94*	1.33	3.92*	1.37
political activities[vi]	4.80	0.58	4.74*	0.64	4.87*	0.50
N	2,013		1,094		919	

Source: ALLBUS 2004. Own calculations.
Notes: [i]): nobody, some, many, almost all (1-4). [ii]): 1=strongly disagree ... 7=strongly agree. [iii]): Only one adult in household, independent from children in household or life partner living elsewhere. [iv] The number of children living in the household varies between zero and four. [v]): Regions with 5,000 or more inhabitants. [vi]): never, infrequent, at least 1x/month, at least 1x/week, daily (1-5). * indicates significant mean differences between users and non-users at 1 percent level.

vocational degree. The proportion of employed persons is 51 percent in 2004 and therefore much lower than the employment rate in Germany. That is not surprising, as the German employment rate is based on people aged between 18 and 64, but in our sample, there are many individuals aged 65 and above who usually are retired.

On average, the net income of households in our sample amounts to about 2,277 EUR per month, ranging from 1,500 to 15,000 EUR. The proportion of individuals living in East Germany is overrepresented in the sample: While it is about one third in the sample, the proportion is 20 percent in Germany as a whole.[6] Moreover, a proportion of 19 percent lives in single adult households. Single adult in household is independent from children in household or life partner living elsewhere. When comparing the sample of internet non-users with the one of internet users, Table 2 indicates that the latter are significantly younger, more highly educated, more often employed and more interested in new technologies.

In the next sections, we investigate which explanatory variables are important for using or not using the internet in a multivariate context (first-level digital divide). In a next step, we analyze the determinants of the second-level digital divide using the same explanatory variables. We then compare the results between the first- and second-level digital divide.

3 Methods

3.1 Modelling Internet Participation

In order to assess the first-level digital divide, we model internet participation by the following equation using a maximum likelihood approach:

$$P(U_i) = f(X_i'\alpha + I_i'\beta + u_i). \tag{1}$$

U is a binary outcome variable which takes the value one if individual i uses the internet and zero otherwise. The vector X_i comprises specific individual characteristics that are widely believed to be significant determinants of the internet intensity. These characteristics include dummy variables for education, age (generation), being female, living in a single household, having a migration background, living in East Germany, living in an urban area, being employed, and discrete variables like the household income, and the number of children living in the household. In addition, it contains leisure preferences, such as the frequencies of sports and political activities as well as reading books.

[6] A finding which is a result of the sample design of the ALLBUS data set.

Vector I_i summarizes attitudes towards ICT which might be related to the internet intensity: the statements 'no interest in computers', 'interested in technology', and 'technological development is too fast'. Moreover, I_i also comprises the number of friends who use the internet. The estimation results are presented in section 4.1.

3.2 Modelling Internet Intensity

When analyzing the second-level digital divide, we firstly examine the determinants of internet intensity. We estimate the following equation:

$$Y_i = X_i'\alpha + I_i'\beta + u_i, \tag{2}$$

where Y_i denotes the internet intensity for individual i measured in usage hours per week. A limitation of model (2) is that a 'cluster' of individuals chooses zero internet hours per week. Figure 1 shows the variation in the hours of internet use in 2004. While there is a remarkable share of internet users being online more than 10 hours per week and even a small proportion among them using the internet more than 50 hours per week, about 50 percent of individuals do not use the internet at all.[7] Ignoring the decision of not using the internet when estimating the internet intensity leads to inconsistent and biased coefficients. In order to correct for the 'selection into zero internet hours', we apply corner-solution (Tobit) models. First, we are interested in the conditional expectation of *internet users*, that is $E(y|y > 0)$. Secondly, we analyze the conditional expectations of *all individuals*, that is: $E(y|x)$.[8] Maximum likelihood estimation results of both expectations are presented in section 4.2.

$$Y_i^* = X_i'\alpha + I_i'\beta + u_i$$

$$Y_i = max\{0, Y_i^*\} \quad Y_i = \begin{cases} 0 & \text{if} \quad Y_i^* \leq 0 \\ Y_i^* & \text{if } Y_i^* > 0. \end{cases} \tag{3}$$

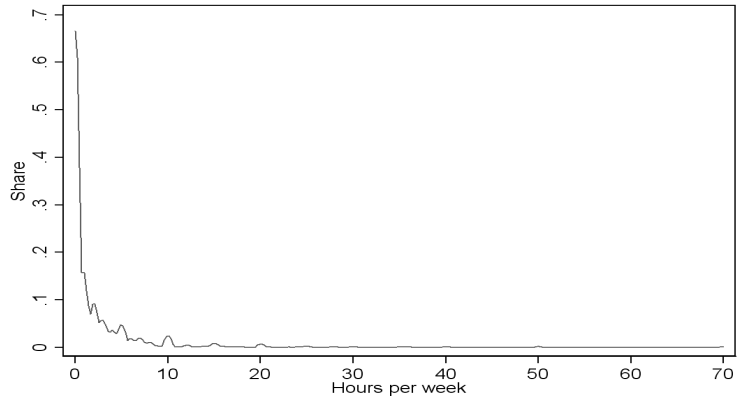
As every second individual is not using the internet, we would expect that the bias in model (2) is quite large. The correction for individuals not using the internet should lead to smaller coefficients in the Tobit model (3) compared to the coefficients in the OLS model (2).

---

[7] The large proportion of non-users in 2004 depends to some extend on the age structure of our sample population as we cannot observe individuals younger than 18 years.

[8] Note that in contrast to models ignoring the selection issue, both conditional expectations are no longer linear.

Fig. 1. Distribution of the internet intensity in 2004



Source: ALLBUS 2004. Own calculations and figure.

3.3 Modelling Frequencies of Internet Activities

In this section, we further examine the second-level divide by analyzing the determinants of different internet activities. We formulate model (4) in the following way:

$$A_i^j = X_i'\alpha + I_i'\beta + u_i \quad \forall j = 1, .., 5 \tag{4}$$

where A_i^j denotes the frequency of internet activity j, varying between never and daily, for each individual i. As stated above, we concentrate on the five activities 'exchanging e-mails', 'searching for information on the web', 'creating web pages', 'downloading content from the internet', and 'playing online games' (see Table 1). Figure 2 presents the distribution of their frequencies in 2004.

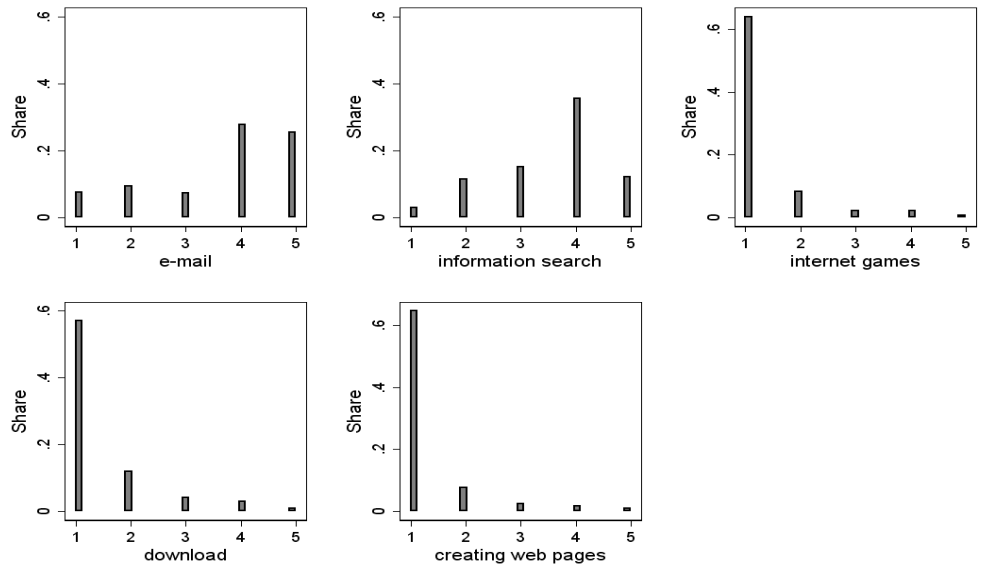
Vector X_i includes the same specific characteristics of individuals and I_i contains the same attitudes and preferences regarding new technologies as in model (2), because they might also be important determinants of the individuals' frequencies of different internet activities.

The limitation of model (4) is that internet activities are only observable for individuals with a positive internet intensity. If there are no differences in unobservable characteristics between both groups, users and non-users, there is no selection issue and model (4) yields consistent and unbiased results. However, if users differ from non-users with regard to unobservable characteristics, estimation results of model (4) are inconsistent. In order to account for a positive selection into 'internet use' (or even a selection into computer use), we estimate a Heckman selection model:

$$A_{1i}^j = X_{1i}'\beta_1 + u_{1i} \qquad \text{(activity equation)} \tag{5}$$

$$Y_{2i}^{j*} = X_i'\beta_2 + u_{2i} \qquad X_{1i} \leq X_i \quad \text{(selection equation)}. \tag{6}$$

Fig. 2. Distribution of the frequencies of internet applications in 2004



Source: ALLBUS 2004. Own calculations and figures. *Notes:* 1=never, 2=infrequent, 3=at least once per month, 4=at least once per week and 5=daily.

The model takes into account that both equations (5) and (6) are not independent, because the covariance of the error terms is not zero. The selection issue is the more relevant the higher the covariance of the error terms. For a positive covariance of the error terms, the OLS coefficients in model (4) are upward biased.

Equation (5) contains nearly the same individual characteristics as model (4). That is, the decision of using the internet depends on about the same factors which determine the frequency of internet activities. However, to correct for selection processes, there must be at least one factor which influences internet participation but not the internet activity itself (exclusion restriction). Besides this first condition, such an 'instrument' must also fulfil a second one: it should be exogenous.

All variables included in vector X_i, but not included in X_{1i}, serve as instruments. One source of valid instruments might stem from different kinds of usage constraints which are influencing the probability of accessing the internet. We use household income and differences between urban and rural areas as instruments. The household income can be considered as reflecting economic constraints. Living in rural or urban areas can be seen as a valid instrument because in 2004, the availability of internet access is much lower in rural regions compared to urban areas, therefore reflecting strong access constraints (Schleife, 2010).

4 Results

4.1 Results: Internet Use

Table 3 presents our estimation results for the first-level digital divide. Our findings indicate that both individual characteristics as well as attitudes towards ICT and leisure preferences combined explain nearly 50 percent of the first-level digital divide. Individuals who belong to the internet generation have a 19 percentage points higher probability of using the internet than the reference age group (40-55 years). Education is also crucial for the probability of using the internet. Individuals with no vocational degree and individuals with a vocational degree are significantly less likely to use the internet in 2004 than individuals with a university degree.

In addition, the household income matters: Increasing income by 1,000 EUR per month increases the probability of using the internet by seven percentage points. Obviously, individuals with migration background and individuals living in rural areas are less likely to use the internet. Besides socioeconomic characteristics, attitudes towards ICT and leisure preferences are significantly related to the probability of using the internet. For example, individuals with no interest in computers have an eight percentage points lower probability of using the internet. Finally, the number of friends using the internet explains the probability of using the internet to a large extent. We interpret this result as an online peer effect.

Table 3
Probit results: Probability of using the internet

	marg.eff.	std.err.
internet use among friends	0.48***	(0.05)
no interest in computers	-0.19***	(0.02)
interest in technology	0.05*	(0.02)
technical development too fast	-0.13***	(0.02)
no vocational degree	-0.69***	(0.18)
vocational degree	-0.60***	(0.15)
A-level & vocational degree	-0.25	(0.20)
employed	0.29***	(0.09)
net income of household (in 1000 Euro)	0.18***	(0.04)
age <40 years	0.49***	(0.09)
age $56-65$ years	-0.51***	(0.12)
age >65 years	-1.17***	(0.17)
number of children in household	-0.08*	(0.05)
single adult in household	0.04	(0.12)
female	-0.14	(0.09)
migration	-0.40**	(0.16)
East Germany	0.03	(0.09)
urban region	0.34**	(0.15)
reading books	-0.07**	(0.03)
sports activities	-0.08***	(0.03)
political activities	-0.09	(0.06)
N	2,013	

Source: ALLBUS 2004. Own calculations.
Notes: Reference categories: university; age $40-55$ years. Marginal effects. Pseudo $R^2 = 0.4891$. Standard errors are in parentheses: *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

4.2 Results: Internet Intensity

The estimation results regarding the determinants of internet intensity of individuals are shown in Table 4. Column 1 provides the OLS results according to equation (2) for comparison. Columns 2 and 3 show the Tobit results according to equation (3). While column 2 presents the estimation results con-

ditional on those individuals with a positive internet intensity ($E(y|y > 0)$), column 3 reports the unconditioned Tobit results ($E(y|x)$). The Tobit results differ only slightly.

The OLS results support the assumption that models that do not adequately account for the active decision of individuals to 'use' the internet zero hours per week can be misleading. As mentioned above, the OLS model leads to absolute larger coefficients for all continuous variables compared to the Tobit models. Although the OLS results are misleading, sign and significance of most coefficients are equivalent to the Tobit results.

Our Tobit results emphasize the importance of educational and financial resources: individuals with a university degree and individuals with a higher income use the internet significantly more hours per week than individuals with at most a vocational degree and individuals who are less wealthy. In addition, age and the number of children living in the household also influence the time spent online. Younger individuals have a higher internet intensity. As mentioned above, this result is underlined by many other studies. Possibly as a result of time restrictions, the duration of internet usage is lower the more children live in the individual's household. Additionally, after controlling for education, age, and children at home, women still spend significantly less time online than men.

Besides these socio-economic factors, preferences and attitudes regarding new technologies are utmost important when explaining inequality in internet intensity. Individuals with less interest in computers and technology or those who perceive technological development as being too fast use the internet significantly less often, on average. Moreover, the results reveal strong peer effects. The higher the number of friends who use the internet, the more time individuals spend online themselves. This can be explained by a higher individual propensity and frequency of communicating online the more friends can be reached this way.

A further interesting aspect is whether individuals with specific hobbies, like reading books, participating in political activities, or doing sports use the internet as a complement or a substitute. In our sample, results indicate only a weak relationship. While political participation or the frequency of reading books show insignificant effects, doing sports shows a significant complementary relationship to the time spent online.

Table 4
Estimation results: Intensity of internet use (hours per week)

	OLS	Tobit I	Tobit II
internet use among friends	0.78***	0.58***	0.84***
	(0.26)	(0.20)	(0.29)
no interest in computers	-0.47***	-0.44***	-0.64***
	(0.14)	(0.11)	(0.16)
interest in technology	0.23*	0.17*	0.25*
	(0.14)	(0.10)	(0.15)
technical development too fast	-0.32***	-0.29***	-0.41***
	(0.12)	(0.09)	(0.14)
no vocational degree	0.48	0.39	0.56
	(0.81)	(0.62)	(0.90)
vocational degree	-0.12	-0.08*	-0.11
	(0.60)	(0.46)	(0.66)
A-level & vocational degree	-0.26	-0.12	-0.17
	(0.73)	(0.56)	(0.81)
employed	-2.45***	-1.88***	-2.59***
	(0.51)	(0.39)	(0.56)
net income of household (in 1,000 Euro)	0.11	0.07	0.11
	(0.14)	(0.11)	(0.16)
age <40 years	-0.18	0.10	0.15
	(0.45)	(0.34)	(0.50)
age 56-65 years	-1.77**	-1.04*	-1.58*
	(0.79)	(0.61)	(0.87)
age >65 years	3.03**	-2.03*	-3.41**
	(1.39)	(1.08)	(1.56)
number of children in household	-0.68***	-0.49***	-0.70***
	(0.23)	(0.18)	(0.25)
single adult in household	1.38**	1.08**	1.49**
	(0.65)	(0.49)	(0.71)
female	-1.27**	-1.15***	-1.67***
	(0.49)	(0.38)	(0.55)
migration	0.76	0.58	0.81
	(1.02)	(0.78)	(1.13)
East Germany	-0.37	-0.36	-0.52
	(0.46)	(0.36)	(0.51)
urban region	-0.30	-0.09	-0.13
	(0.86)	(0.66)	(0.96)
reading books	-0.04	0.07	-0.10
	(0.17)	(0.13)	(0.19)
sports activities	0.35**	0.28**	0.40**
	(0.16)	(0.12)	(0.18)
political activities	0.02	-0.09	-0.13
	(0.31)	(0.24)	(0.34)
constant	4.51*	3.41*	4.92*
	(2.31)	(1.76)	(2.54)
N	875	875	875

Source: ALLBUS 2004. Own calculations.
Notes: Reference categories: university; age 40-55 years. Tobit: marginal effects. OLS: adjusted $R^2 = .190$. Tobit I: $E(y|y > 0)$, Tobit II: $E(y|x)$. Standard errors are in parentheses: *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

4.3 Results: Frequency of Internet Activities

Table 5 and Tables A.2 - A.5 (Appendix) present the estimation results regarding the determinants of our five different internet activities. In every table, the OLS estimation results (column 1) according to equation (4) are compared with the Heckit estimation results (column 3) according to equations (5) and (6).

For the frequency of information search on the internet, we exemplarily present the first-stage regression results in Table A.1. They show that all instruments are significantly related to the internet access decision. Individuals with a relatively high income and those who live in an urban area are significantly more likely to use the internet. [9]

Even though it could be expected that unobserved preferences might be related to a selection into internet use, e.g. individuals with high (unobservable) preferences towards communications are more likely to use computers and the internet, comparing the OLS and Heckman results indicates only small differences between the coefficients. In addition, only for downloading and creating web-pages, the null hypothesis of 'no selection' can be rejected (λ is significant), indicating only quite a small selection problem.

In contrast to the results regarding the first-level digital divide (section 4.1) and those regarding internet intensity (section 4.2), education and age turn out to be insignificant. Thus, these characteristics have an impact on the decision to use the internet and on the internet intensity, but they are without importance when explaining how often individuals use the internet for information search. However, attitudes and preferences regarding new technologies are of utmost importance. Being interested in computers and new technologies accompanies a high frequency of internet search. People complaining about the fast development of technologies have a lower frequency of online search activities. Again, we find a strong peer effect: the number of friends using the internet has a significantly positive impact.

Perhaps as a result of time restrictions, the frequency of reading books is significantly negatively related to the frequency of online search. However, at the same time, hobbies like political activities or sports are without significant impact. They do not seem to significantly compete for time with online activities. Interestingly, for individuals living in East Germany, the frequency of online information search is significantly lower than for those in West Germany.

[9] With a value of 981.02, the $LR - Chi^2$- test indicates a significant influence of the two exclusion restrictions on the probability of using the internet. Hence, we can conclude that there is no weak instrument problem (Stock et al., 2002).

Table 5
Estimation results: Frequency of information search on the internet

	OLS		Heckit	
	coef.	std.err.	coef.	std.err.
internet use among friends	0.17***	(0.04)	0.16**	(0.07)
no interest in computers	-0.13***	(0.02)	-0.13***	(0.03)
interest in technology	0.07***	(0.02)	0.05**	(0.02)
technical development too fast	-0.06***	(0.02)	-0.06***	(0.02)
no vocational degree	0.003	(0.11)	-0.05	(0.15)
vocational degree	-0.16**	(0.08)	-0.15	(0.11)
A-level & vocational degree	-0.06	(0.11)	-0.10	(0.12)
employed	-0.11*	(0.07)	-0.08	(0.11)
age <40 years	0.04	(0.11)	0.11	(0.07)
age $56-65$ years	-0.01	(0.11)	0.03	(0.13)
age >65 years	-0.08	(0.20)	-0.11	(0.23)
number of children in household	-0.08**	(0.03)	-0.06	(0.04)
single adult in household	0.09	(0.09)	0.07	(0.11)
female	-0.20***	(0.07)	-0.20**	(0.08)
migration	0.06	(0.15)	0.19	(0.17)
East Germany	-0.19***	(0.06)	-0.21***	(0.07)
reading books	-0.08***	(0.02)	-0.07**	(0.03)
sports activities	0.01	(0.02)	0.01	(0.03)
political activities	0.002	(0.04)	-0.02	(0.05)
constant	3.64***	(0.30)	3.79***	(0.40)
$\hat{\lambda}$			-0.06	(0.24)
Chi^2/F-test:	13.26***		629.64***	
adjusted R^2	0.1717			
(uncensored) N			886	
censored N			1,128	
N	1,125		2,014	

Source: ALLBUS 2004. Own calculations.

Notes: Reference categories: university; age 40-55 years. Instruments: income of household; urban. Standard errors are in parentheses: *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

Another very common internet activity is the exchange of e-mails, presented in Table A.2. Similar to online information search, preferences and attitudes as well as 'online-peers' are more important than socio-economic characteristics like education, age, and family composition.

Playing internet games is more common among individuals with a lower school degree and among single households, as the results provided in Table A.3 reveal. In contrast to the exchange of e-mails and online search activities, the results suggest that attitudes and preferences towards ICT are of no significant impact when explaining the frequency of playing internet games. As the overall explanatory power of the model is below five percent, there might

be other characteristics not observable in our data set, which are correlated with this internet activity.

A similar pattern of correlations is observable for the intensity of internet downloads (see Table A.4). However, in contrast to the results of playing online games, individuals with a rather high interest in technology have a significantly higher frequency of downloading files from the internet. Moreover, a significantly negative difference arises between men and women.

Finally, our results in Table A.5 suggest that again, especially attitudes and preferences towards ICT are important for explaining inequalities in the frequency of using the internet for creating web pages. The interest in computers and new technologies significantly promotes the intensity of this activity. As creating web pages needs advanced IT skills, this result is quite obvious: Individuals interested in new technologies rather tend to invest in these skills compared to others. Moreover, the peer effect is strong: The more friends use the internet, the more frequently individuals create web pages. Learning from others could be a reliable explanation for this finding.

To summarize, while socio-economic factors are of great importance when explaining the use as well as the intensity of internet use measured by the hours per week spent online, they have less impact on the frequency of specific online activities. An outstanding factor which is important for the first-level as well as for the second-level digital divide is the internet participation of friends. Our results therefore indicate strong online peer effects.

5 Concluding Remarks

When analyzing the digital divide, various studies look at the determinants of participation (first-level digital divide), but only a few focus on particular activities and the time spent on the internet. Differences between individuals and population groups regarding these aspects have become known as the second-level digital divide. In this study, we focus on both first- and second-level digital divide in private internet use in Germany for the year 2004 by analyzing individuals aged 18 years or older. We use the ALLBUS data set which allows us to consider specific attitudes and preferences towards ICT. Such variables are often missing in other data sets.

Our results imply that socio-economic characteristics, like education, age, income and migration background, as well as usage preferences and the affinity regarding new technologies explain half of the first-level digital divide. Accordingly, in 2004, in particular less trained individuals, such as the elderly and migrants, as well as individuals with a low income, are excluded from internet usage - a result that supports the findings of other studies. Moreover,

strong peer effects occur: The probability of using the internet is significantly higher the more friends use the internet.

However, in contrast to the first-level digital divide, socio-economic factors hardly seem to play a role for the duration of usage and the kind of internet applications the individuals perform. These factors strongly depend on the individuals' interests and IT affinity. Disparity with regard to the duration and the kind of usage can also be traced back to the individuals' online-peers and leisure time preferences.

As the usage of internet increasingly dominates everyday life with regard to all spheres, an important economic and social task is to minimize access difficulties. Encouraging remaining non-users to become part of the information society by providing them with access to online information and respective skills will remain an important challenge to policy in the next few years. In addition, for individuals to fully realize the benefits of the internet, they need to move from basic activities, such as e-mail and browsing, to more advanced activities, such as e-learning and transactional online activities like buying, banking and accessing governmental services. Therefore, IT interests need to be stimulated and IT literacy to be improved. Strong efforts should focus on encouraging essential skills and competencies for example by providing age-specific courses.

Technology is changing rapidly and more and more specific IT skills become a prerequisite for successfully participating in the labor market. They therefore need to be strongly encouraged. Thus, future research should concentrate much more on the variety of applications the internet is used for. Moreover, in order to obtain a deep insight in emerging and persistent, divides more recent representative data are crucial, not least focussing on new media such as mobile internet and the appropriate skills.

Acknowledgments

We gratefully acknowledge the helpful comments by Irene Bertschek, Marianne Saam and participants at the ZEW-ICT Seminar.

Appendix

Table A.1
First stage results – sample selection

	coefficient	std.error
net income of household (in 1,000 Euro)	0.16***	(0.03)
urban	0.29**	(0.14)
internet use among friends	0.48***	(0.04)
no interest in computers	-0.16***	(0.02)
interest in technology	0.01	(0.02)
technical development too fast	-0.11***	(0.02)
reading books	-0.03	(0.03)
sports activities	-0.08***	(0.03)
political activities	-0.01	(0.06)
no vocational degree	-0.52***	(0.15)
vocational degree	-0.42***	(0.13)
A-level & vocational degree	-0.14	(0.16)
employed	0.51***	(0.07)
number of children in household	0.18***	(0.04)
single adult in household	-0.19*	(0.10)
female	-0.04	(0.08)
migration	-0.36**	(0.15)
East Germany	0.02	(0.08)
N		2,014

Source: ALLBUS 2004. Own calculations.

Notes: Reference categories: university. Standard errors are in parentheses: ***significant at 1% level; **significant at 5% level; * significant at 10% level.

Table A.2
Estimation results: Frequency of exchanging e-mails

	OLS		Heckit	
	coef.	std.err.	coef.	std.err.
internet use among friends	0.21***	(0.05)	0.19**	(0.09)
no interest in computers	-0.08***	(0.03)	-0.07**	(0.04)
interest in technology	0.05**	(0.02)	0.02	(0.03)
technological development too fast	-0.10***	(0.02)	-0.09***	(0.03)
reading books	-0.13***	(0.03)	-0.12***	(0.03)
sports activities	-0.04	(0.03)	-0.04	(0.03)
political activities	-0.14***	(0.06)	-0.15**	(0.06)
no vocational degree	-0.26*	(0.14)	-0.03	(0.18)
vocational degree	-0.25**	(0.11)	-0.20	(0.13)
A-level & vocational degree	-0.18	(0.13)	-0.19	(0.15)
employed	-0.26***	(0.09)	-0.23*	(0.14)
age <40 years	-0.01	(0.14)	0.20**	(0.09)
age $56-65$ years	-0.17	(0.14)	0.25	(0.16)
age >65 years	-0.35	(0.25)	-0.39	(0.28)
number of children in household	-0.12***	(0.04)	-0.13**	(0.14)
single adult in household	0.20*	(0.11)	0.14	(0.14)
female	-0.26***	(0.09)	-0.29***	(0.10)
migration	-0.14	(0.19)	-0.11	(0.21)
East Germany	-0.26***	(0.08)	-0.26***	(0.09)
constant	5.04***	(0.37)	5.10***	(0.50)
$\hat{\lambda}$			-0.21	(0.30)
Chi^2/F-test:	12.51***		632.09***	
adjusted R^2	0.1629			
(uncensored) N			886	
censored N			1,128	
N	1,125		2,014	

Source: ALLBUS 2004. Own calculations.

Notes: Reference categories: university; age 40-55 years. Standard errors are in parentheses: *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

Table A.3
Estimation results: Frequency of playing internet games

	OLS		Heckit	
	coef.	std.err.	coef.	std.err.
internet use among friends	0.01	(0.03)	-0.03	(0.05)
no interest in computers	-0.02	(0.02)	-0.01	(0.02)
interest in technology	-0.01	(0.01)	-0.02	(0.02)
technological development too fast	-0.02	(0.01)	-0.01	(0.02)
reading books	0.02	(0.02)	0.02	(0.02)
sports activities	-0.002	(0.02)	-0.002	(0.02)
political activities	-0.02	(0.03)	-0.04	(0.04)
no vocational degree	0.49***	(0.08)	0.44***	(0.10)
vocational degree	0.19***	(0.06)	0.23***	(0.08)
A-level & vocational degree	0.16**	(0.08)	0.15*	(0.08)
employed	0.01	(0.05)	-0.03	(0.08)
age <40 years	0.03	(0.08)	0.08	(0.05)
age $56-65$ years	-0.03	(0.08)	0.04	(0.09)
age >65 years	0.001	(0.15)	-0.06	(0.16)
number of children in household	-0.02	(0.03)	-0.02	(0.03)
single adult in household	0.19***	(0.06)	0.20**	(0.08)
female	-0.06	(0.05)	-0.06	(0.06)
migration	-0.11	(0.11)	-0.03	(0.12)
East Germany	0.07	(0.05)	0.11*	(0.05)
constant	1.19***	(0.22)	1.50***	(0.28)
$\hat{\lambda}$			-0.17	(0.17)
Chi^2/F-test:	3.89***		541.50***	
adjusted R^2	0.0465			
(uncensored) N			886	
censored N			1,128	
N	1,125		2,014	

Source: ALLBUS 2004. Own calculations.

Notes: Reference categories: university; age 40-55 years. Standard errors are in parentheses: *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

Table A.4
Estimation results: Frequency of internet downloads

	OLS		Heckit	
	coef.	std.err.	coef.	std.err.
internet use among friends	0.03	(0.03)	0.09	(0.07)
no interest in computers	-0.02	(0.02)	-0.04	(0.03)
interest in technology	0.04**	(0.02)	0.04**	(0.02)
technological development too fast	-0.03*	(0.02)	-0.04**	(0.02)
reading books	0.00	(0.02)	-0.002	(0.02)
sports activities	0.03*	(0.02)	0.03	(0.02)
political activities	-0.02	(0.04)	-0.02	(0.04)
no vocational degree	0.51***	(0.10)	0.57***	(0.13)
vocational degree	0.09	(0.07)	0.07	(0.10)
A-level & vocational degree	0.15*	(0.09)	0.13	(0.11)
employed	-0.05	(0.06)	0.02	(0.10)
age <40 years	0.38***	(0.09)	0.15**	(0.06)
age $56-65$ years	0.23**	(0.10)	-0.19*	(0.11)
age >65 years	0.04	(0.17)	-0.20	(0.20)
number of children in household	-0.08***	(0.03)	-0.07*	(0.04)
single adult in household	0.17**	(0.07)	0.17*	(0.10)
female	-0.24***	(0.06)	-0.22***	(0.07)
migration	-0.06	(0.13)	-0.02	(0.15)
East Germany	0.02	(0.06)	0.05	(0.06)
constant	0.99***	(0.25)	1.01***	(0.35)
$\hat{\lambda}$			0.16	(0.21)
Chi^2/F-test:	8.86***		628.50***	
adjusted R^2	0.1175			
(uncensored) N			885	
censored N			1,129	
N	1,123		2,014	

Source: ALLBUS 2004. Own calculations.

Notes: Reference categories: university; age 40-55 years. Standard errors are in parentheses: *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

Table A.5
Estimation results: Frequency of creating web pages

	OLS		Heckit	
	coef.	std.err.	coef.	std.err.
internet use among friends	0.08***	(0.03)	0.18***	(0.06)
no interest in computers	-0.03*	(0.02)	-0.06**	(0.02)
interest in technology	0.04***	(0.01)	0.05***	(0.02)
technological development too fast	-0.04***	(0.01)	-0.05***	(0.02)
reading books	-0.02	(0.02)	-0.04	(0.02)
sports activities	-0.001	(0.02)	-0.03	(0.02)
political activities	-0.09***	(0.03)	-0.08**	(0.04)
no vocational degree	0.02	(0.08)	0.01	(0.12)
vocational degree	0.01	(0.06)	-0.03	(0.08)
A-level & vocational degree	0.09	(0.08)	0.08	(0.09)
employed	-0.03	(0.05)	0.07	(0.08)
age <40 years	0.14*	(0.08)	-0.02	(0.06)
age $56-65$ years	0.15*	(0.08)	-0.19*	(0.10)
age >65 years	0.05	(0.15)	-0.11	(0.17)
number of children in household	-0.02	(0.03)	-0.003	(0.03)
single adult in household	0.04	(0.06)	-0.05	(0.08)
female	-0.13**	(0.05)	-0.14**	(0.06)
migration	-0.08	(0.11)	-0.12	(0.13)
East Germany	0.08*	(0.05)	0.06	(0.06)
constant	1.37***	(0.22)	1.14***	(0.31)
$\hat{\lambda}$			0.35*	(0.19)
Chi^2/F-test:	5.53***		557.60***	
adjusted R^2	0.0712			
(uncensored) N			885	
censored N			1,129	
N	1,123		2,014	

Source: ALLBUS 2004. Own calculations.

Notes: Reference categories: university; age 40-55 years. Standard errors are in parentheses: *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

References

DiMaggio, P., Hargittai, E., Celeste, C., Shafer, S., 2004. From Unequal Access to Differentiated Use: A Literature Review and Agenda for Research on Digital Inequality. Report prepared for the Russell Sage Foundation.

Korupp, S.E., Szydlik, M., 2005. Causes and Trends of the Digital Divide. European Sociological Review 21(4), 409-422.

Lenhart, A., Horrigan, J., Rainie, L., Allen, K., Boyce, A., Madden, M., OGrady, E., 2003. The Ever-Shifting Internet Population − A New Look at Internet Access and the Digital Divide. Pew Internet & American Life Project Report, http://www.pewinternet.org/reports.asp. accessed: February 15, 2008.

Livingstone, S., and Helsper, E., 2007. Gradations in Digital Inclusion: Children, Young People and the Digital Divide. New Media & Society 9(4), 671-696.

OECD/DSTI, 2001. Understanding the Digital Divide. http://www.oecd.org/dataoecd/38/57/1888451.pdf. accessed: March 03, 2008.

Sackmann, R., Weymann, A., 1994. Die Technisierung des Alltags. Generationen und technische Innovationen. Frankfurt/New York: Campus.

Schleife, K., 2010. What Really Matters: Regional versus Individual Determinants of the Digital Divide in Germany, Research Policy, 39, 173–185.

Statistisches Bundesamt, 2009. Private Haushalte in der Informationsgesellschaft − Nutzung von Informations- und Kommunikationstechnologien. Fachserie 15 Reihe 4.

Stock, James H. and Wright, Jonathan H. and Yogo, Motohiro, 2002. A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments, Journal of Business and Economic Statistics, 20 (4), 518–529.

TNS Infratest, 2009. (N)Onliner-Atlas 2009 − Eine Topographie des digitalen Grabens durch Deutschland.

van Dijk, J., Hacker, K., 2003. The Digital Divide as a Complex and Dynamic Phenomenon. The Information Society, 19, 315-326.

van Eimeren, B., Frees, B., 2009. Der Internetnutzer 2009 multimedial und total vernetzt? − Ergebnisse der ARD/ZDF-Onlinestudie 2009. Media Perspektiven 7/2009.

van Eimeren, B., Gerhard, H., Frees, B., 2004. Internetverbreitung in Deutschland: Potenzial vorerst ausgeschöpft? − ARD/ZDF-Onlinestudie 2004. Media Perspektiven 8/2004.