Regional Versus Individual Aspects of the Digital Divide in Germany

Katrin Schleife[*]

February 2007

Abstract

This paper analyzes the regional dimension of the German digital divide. It considers the impact of regional characteristics on differences in the share of Internet use between German counties. In addition, it studies the influence of regional factors as well as individual characteristics on the individual probability of becoming a new Internet user. Based on two large data sets, SOEP and INKAR, the analyses show that it is not the rusticity of a region itself that explains regional differences in Internet use. The results rather indicate that it is the different composition of the population between rural and urban areas that accounts for the regional digital divide.

Keywords: digital divide, Internet use, regional differences
JEL-Classification: O33, R20, O18

[*]**Address:** ZEW Mannheim, Centre for European Economic Research, Research Group Information and Communication Technologies, P.O.Box 103443, D-68034 Mannheim, Germany. Email: schleife@zew.de

1   Introduction

In Germany, Internet access has spread over the whole country during the last years. In 2006, a total of nearly 60 percent of the German population older than 14 years uses the Internet at home or on the job, compared to 37 percent in 2001 (Infratest [9]). But, measuring the shares of Internet users among various population groups, the Nonliner-Atlas 2006 shows a difference of 11 percentage points between small communities having 51 percent users and large cities having 62 percent (Infratest [9]), a difference that was even growing in recent years.[1] That development could be the result of the uneven regional diffusion of broadband connection.[2] However, rural-urban Internet use differences were already prevalent in 2001 when broadband was used by scarcely anybody in German private households.[3] Thus, infrastructure differences are less likely to represent the main source of the rural-urban digital divide[4] with respect to Internet use observed in Germany. As closing the digital divide is one of the current economic challenges within and between countries the dimensions and sources of the digital divide need to be studied in order to find appropriate policies targeted at closing the gap.

This paper analyzes regional differences in Internet penetration rates in Germany and therefore contributes to the digital divide research. Research studies focussing on the regional dimension of the digital divide mostly analyze differences between countries. Regional aspects within countries that go *beyond* population density and *beyond* mere technical access considerations are much less prevalent.

---

[1] Small communities are those with a population of less than 5,000 inhabitants, large cities have a population of more than 500,000. The difference in the share of Internet users between small communities and large cities decreased between 2001 and 2004 from 8.5 to 6.4 percentage points, but grew strongly afterwards.

[2] As Internet providers are attracted by market size they focus their activities on densely populated areas. Thus, the availability of broadband access that allows the use of a wide range of Internet applications at a high rate of data transfer is limited and more expensive in rural areas.

[3] At the beginning of 2001, only 4 percent of German user households had broadband access (Infratest [8]).

[4] The OECD defines it as follows: "The term 'digital divide' refers to the gap between individuals, households, business and geographic areas at different socio-economic levels with regard both to their opportunities to access information and communication technologies (ICTs) and to their use of the Internet for a wide variety of activities."

Thus, when studying the determinants of Internet use diffusion in private households in Germany this paper takes various regional characteristics into account. It shows how regional aspects are related to differences in Internet use between regions as well as between individuals. Moreover, it takes account of network effects.

The empirical work is based on two large data sets: the German Socio-Economic Panel (SOEP) and the INKAR data.[5] This provides the opportunity of merging detailed individual and regional information. The analyses are carried out in two steps. In the first step, it is studied which regional characteristics determine the degree of home Internet use in German counties. Corresponding to other studies, the results confirm that in less densely populated German counties Internet diffusion is less advanced. However, part of this population density effect can be explained by other regional characteristics: The county-wide share of highly qualified employees and the share of one-person households are positively related to the penetration of Internet use at home. The regional unemployment rate as well as the share of foreigners show a significantly negative correlation.

In a second step, further analyses are carried out on the individual level where the decision of becoming a new Internet user is studied. Besides factors that characterize the county the individuals live in, the impact of individual characteristics is analyzed. The results show that the effect of only two of the regional factors are of significant size when individual factors are additionally considered: Besides the share of one-person households, the local share of experienced Internet users (usershare) reveals a highly significant positive impact on the decision of becoming a new user, signalling positive network effects. Thus, individual characteristics are much more important than regional determinants when analyzing the individual probability of becoming a new Internet user.

This paper proceeds as follows: Section 2 gives a short overview on the results of previous studies. The data are described in Section 3. Section 4 exhibits hypotheses and the empirical framework and Section 5 presents the results on the regional as well as on the individual level. Section 6 concludes.

---

[5]INKAR denotes 'Indikatoren und Karten zur Raumentwicklung' (indicators and maps on land development).

2 Related Literature

Within wealthy nations, such as the United States, Germany, the United Kingdom, and Scandinavian countries, Internet connectivity is widely spread. But even within these countries sizable discrepancies exist between population groups regarding their Internet use. Research in various countries ascertains that individual characteristics, such as age, income or educational background can generate major usage barriers. An OECD analysis in 2001 based on national statistics of OECD countries shows that Internet access and use is higher for individuals and households with a higher income and a higher level of education (OECD [15]). The authors conclude that policies should aim at improving Internet literacy by expanding the related skill base in educational institutions and vocational training programs. Analyzing the online population in the U.S. Lenhart [13] reaches similar conclusions with survey data of 2002.[6]

Largely similar results are obtained for Germany. Using German SOEP data of the years 1997, 2001 and 2003, the empirical analyses of Korupp and Szydlik [12] indicate a positive impact of educational background and income on home Internet use. They additionally show that being a young adult as well as being male significantly increase the Internet use probability. The recent study of Korupp et al. [11] based on a special SOEP ICT survey supports the negative impact of age and the positive impact of education on the probability of using ICT.

The results of Jäckel et al. [10] confirm the important role of education analyzing German data for 2003. Besides individual characteristics they consider the role of community size and analyze differences between German cities and rural areas in accessing and using the Internet. However, in their multivariate discriminant analysis the rural-urban impact on Internet use diminishes once education is taken into account.

The role of regional factors on the probability of using ICT is researched more intensely in non-German studies. They show, for example, that the attainable

---

[6]Lenhart [13] additionally states that there may be voluntary and involuntary patterns of non-use. Some people choose not to use the Internet, others are faced with circumstances that make Internet access difficult or unreachable. Thus, policies should be targeted at sub-groups in order to motivate people to use the Internet or to remove other group-specific access barriers.

benefits of using the Internet are determining regional Internet penetration rates. As Sinai and Waldfogel [16] state, the Internet may act as both, a substitute and a complement for cities. The substitute function originates from individuals using the Internet to overcome local isolation regarding communication and product availability. If that function prevails, Internet use should be higher in rural areas. On the other hand, the Internet complements cities as it offers local websites containing local news and information. The amount of these sites grows with size and population density of the region. Besides, local sellers may provide special services or may offer additional products via the Internet. This is also part of the complement function as the density of sellers is much higher in city areas. Using data for the U.S., Sinai and Waldfogel [16] conclude that both, the substitute as well as the complement function are observable. The strength of these functions varies between population groups.

Gaspar and Glaeser [4] provide some evidence that new information technologies and cities are rather complements than substitutes by stating that face-to-face contacts and electronic contacts are positively correlated. The authors suggest that new information technologies increase the number of contacts, as even long distance collaborations involve significant face-to-face activities. Thus, cities attract ICT as a higher population density enables more and easier face-to-face contacts. However, the authors do not take a possible selection bias into account that is due to people with specific characteristics that self-select into cities.

Mills and Whitacre [14] consider such household characteristics when explaining the large differences in Internet use between metropolitan and non-metropolitan areas in the U.S. They test the relative importance of household attributes versus region-based differences in analyzing the rural-urban digital divide. Their results suggest that nearly two third of the divide can be explained by differences in household characteristics like education (of household head) and income. One third stems from place-based characteristics, especially from network externalities. Mills and Whitacre [14] show that positive externalities exist as a higher regional density of Internet use significantly increases a household's probability of using the Internet. Given the importance of network externalities and differences in household characteristics the authors state that policies that are targeted solely at creating infrastructure and high-speed Internet access are not sufficient in order to close the digital gap.

The importance of network externalities in the diffusion of ICT is studied in more detail by Goolsbee and Klenow [6] focussing on the diffusion of home computers in the U.S. They find that households living in regions with a higher share of people that already own a computer are more likely to buy their first computer even if various individual characteristics are considered. By carrying out a multitude of tests Goolsbee and Klenow [6] show that the strong network effect is robust as it cannot be explained by common traits or common economic environments.

This paper contributes to the empirical research on the various dimensions of the digital divide. It studies the determinants of home Internet use in Germany on the level of counties as well as on the level of individuals by merging two large data sets. The regional determinants included in the estimation approaches are not limited to population density figures as they are in most of the other studies but comprise many more characteristics, such as the regional shares of youths, foreigners, and highly qualified employees. Below, the data are described in detail.

3   Data

The empirical analyses of this paper are performed on the basis of two data sets: the Socio-Economic Panel (SOEP) that provides detailed information on individuals, and INKAR providing a wide range of official regional figures for Germany. With the combined data set various individual socio-economic as well as region-based characteristics can be considered in order to analyze the inequalities in home Internet use between regions and individuals.

The SOEP is a representative longitudinal survey of private households collected by the German Institute for Economic Research (DIW).[7] Annually, since 1984, the same individuals have been asked for the development of their living and working conditions. Since the German reunification in 1990, East German households have been added to the survey. The 2001 SOEP wave covers more than 22,000 individuals aged 16 years or older. For the analyses of this paper the 2001 SOEP wave is considered, as besides other socio-economic variables it provides informa-

---

[7]See Haisken-DeNew and Frick [7] for a detailed description of the SOEP.

tion on the individual's computer and Internet use.[8] However, the SOEP data contain hardly any regional information. Thus, a second data set is needed to fill this gap.

The INKAR data are provided by the German Federal Office for Building and Regional Planning[9] and contain a wide range of regional figures, for instance regarding the structure of population, employment and industry, or levels of education, production and wages.[10] Thus, INKAR does not only allow the regions to be classified as rural or non-rural, but provides a much more detailed description of regions. The INKAR data are given on several regional levels. The county level is the smallest and it is used in this paper's analyses. As most variables are given with a time lag of one to three years, INKAR data of 2002, 2003 and 2004 are used. That provides the information for the year 2001 used on the regional level or for the year 2000 used on the individual level.[11] By merging the two data sets regional information can be assigned to individuals. On the individual level of the analyses, the regional information can therefore be treated as person-specific determinants.

Germany consists of 440 counties. Seven of these counties cannot be considered in the analyses as they do not match the SOEP data. Moreover, only counties that contain 20 or more observed SOEP individuals are taken into consideration. As the sample is additionally restricted to individuals with an age between 16 and 64, and to individuals who provide information regarding their Internet use, 312 counties and 16,662 individuals remain in the data set.

---

[8]The related questions were: 'Do you use a computer and the Internet for activities not related to work? If so, since when?' and 'Do you use a computer and the Internet at work or in your education? If so, since when?'. Questions regarding computer and Internet use are not included in every SOEP wave.

[9]Bundesamt fuer Bauwesen und Raumordnung - BBR.

[10]INKAR (indicators and maps on land development) does not include information on ICT use and access.

[11]Age structure figures and GDP per capita are only provided for the year 2000. As the share of foreigners is given for 1999 and 2001 only, the figures for the year 2000 are taken from an additional data set: 'Statistik regional 2002' provided by the Federal Statistical Office Germany.

4 Hypotheses and Estimation Strategies

4.1 The regional level

The analyses of Internet use determinants are carried out on a regional as well as on an individual level. At first, on the regional level, differences in the shares of private Internet use between German counties are analyzed. The main research questions are: What are the regional characteristics that determine the degree of home Internet use in German counties? Can population density explain differences in Internet use between regions? Can cities therefore be seen as a complement or a substitute for Internet use?

The individual Internet use information given by the SOEP data has been aggregated by county, providing the respective share of Internet users, which is the outcome variable of the econometric model. One of the main regional explanatory variables is a rural-urban indicator: the share of population living in communities with less than 150 inhabitants per squared kilometer, the so called rusticity. It serves to find the impact of population density on Internet use shares. In recent papers it is often argued that technological differences, such as the availability and the quality of Internet connectivity, are one of the main reasons for differences in Internet use rates between rural areas and cities. In Germany, like in many other countries, especially the availability of broadband is much less prevalent or much more expensive in rural regions. However, at the beginning of the analyzed year 2001, there was only the very small fraction of 4 percent of private user households in Germany that used a broadband Internet connection (Infratest [8]). Thus, when explaining the underlying causes of the digital divide, Internet infrastructure differences do not seem to be among the most important factors a few years ago.

Other local characteristics, such as the attainable benefits of using the Internet can determine regional Internet penetration rates. As Sinai and Waldfogel [16] state, the Internet may act as both a substitute and a complement for cities. Depending on which function outweighs the other, the relationship between population density and Internet use rates is positive or negative.

Part of the rural-urban differences in Internet use caused by the Internet's complement or substitute properties could be explainable by miscellaneous region-specific characteristics, such as the regional size of minority groups like foreigners. Thus, the hypothesis is that not population density itself is the crucial factor but the regional size of specific population groups that have specific preferences regarding communication and consumption. Therefore, the significance of rusticity is expected to decrease if additional regional factors are considered in the estimation approach. Further variables that may have an impact on Internet use shares and that are therefore added in a further specification are: the share of population with an age between 18 and 29, the population shares of foreigners and women, the share of one-person households on all households, the share of highly qualified employees on all employees in the county, the unemployment rate, and GDP per capita.[12]

The local share of young people is expected to show a significantly positive impact on regional Internet use rates, as today adolescents get in touch with ICT very early by using it in school, in apprenticeship, or for communication as well as playing activities among friends. They encourage using a computer and the Internet within their families, therefore increasing the regional home Internet use share.

Many studies find complementarities between skills and new technology use and show that the educational level increases the probability of using a computer or the Internet at work.[13] An important reason is that as computers and the Internet were introduced first in academic and research institutions highly educated people got used to these new technologies earlier than others.[14] Thus, by accumulating ICT skills at work, highly qualified workers become also more likely to

---

[12]While all determinants are taken for 2001 the regional share of young people with an age between 18 and 29 years and per capita GDP are taken for the year 2000 due to data limitations. Unfortunately, no data are available regarding the activities the Internet is used for or regarding the amount of region-specific websites.

[13]For example, this result is obtained by Korupp et al. [11] for Germany, by Borghans and ter Weel [1] using data of Germany, the UK, and the USA, and by Entorf et al. [3] analyzing French data.

[14]The important role of universities in the process of Internet diffusion is analyzed by Goldfarb [5]. He provides evidence that in the mid-1990s universities in the U.S. intensely taught students how to use the Internet. "These students then brought the technology into their homes..." (p. 203).

use computers and the Internet at home. Assuming that they work and live in the same county, a higher regional share of highly qualified employees is therefore expected to lead to a higher regional share of home Internet use.

As the availability of time and money - two important requirements for using the Internet - is often high in one-person households, a positive relationship with the county-wide share of Internet use should be observable.

Unemployed people could highly benefit from using the Internet for their job search activities. But many studies ascertain that they are less likely to own a computer at home and to have access to the Internet, presumably because of higher financial restrictions. Thus, the share of home Internet users is expected to be reduced by a higher local share of unemployed inhabitants.

Following the arguments of Sinai and Waldfogel [16], foreigners tend to use the Internet to overcome local isolation. In areas where many foreign individuals live, face-to-face communication as well as the exchange of information is much easier for them to perform. In addition, as the group of foreigners may have its own preferences, its size determines the amount of preferred products available on the local offline market. The higher the local foreign minority group the more foreign newspapers and shops exist, and the smaller are the benefits of using the Internet to get some information and to buy preferred products. Thus, the higher the foreigners' share the less the Internet is needed for their online communication or shopping activities. As a result, this may intentsify the anticipated negative impact of the population share of foreigners probably caused by linguistic problems or shortcomings in education.

Resuming all these considerations, the hypotheses to be analyzed are: i) A higher rusticity leads to a smaller regional share of home Internet users. ii) The impact of rusticity declines when other regional characteristics are additionally considered. iii) The share of young people, of one-person households and of highly qualified employees have a positive impact on the regional Internet use shares. iv) A high unemployment rate and a higher share of foreigners in a county lead to a smaller share of Internet users.

On the regional level, the hypotheses are tested by using an OLS approach. The analysis is carried out on county level and can be modelled as

$$S_r = \alpha + \beta \cdot rusticity_r + \gamma \cdot C_r + u_r \qquad \text{with} \quad r = 1 \dots k, \tag{1}$$

where the dependent variable S_r is the population share of home Internet use in county r in 2001. The coefficient β measures the impact of a region's *rusticity* on the share of Internet users, γ denotes a vector of coefficients measuring the sensitivity of Internet use to various other county characteristics C_r, α is a constant, and u_r is the normally distributed error term.

4.2 The individual level

On the individual level, differences in the probability of becoming a new Internet user are analyzed. Besides several individual demographic and employment-related characteristics, regional determinants that characterize the individual's home county are taken into consideration. The main questions are: Do the regional factors that are found to have an impact on the county-wide home Internet use also show a relation to the individual's probability of starting to use the Internet at home? And if that is the case, is the impact still of significant size when additionally considering individual characteristics? What is the role of network effects?

Contrary to the regional level, where differences in Internet use shares are explored, the focus here is on the individual decision of *starting* to use the Internet. By exploiting the information on the duration of Internet use provided by the SOEP data, 'beginners' are defined as those individuals who declared to use the Internet at home since 2000 or 2001.[15] The data show that 2,346 individuals are new users, compared to 11,280 individuals who did not start using the Internet so far. The remaining 3,036 individuals are those with more than one year usage experience.

---

[15]The question concerning the duration of usage is addressed to users only. In order to increase the number of new users in the data set, not only one but two years are taken into account.

The determinants initially entering the estimation analysis on the individual level are again population density figures. As the rusticity variable is not provided for the year 2000, a county type variable is used: For the analyses the three types of German counties provided by INKAR are defined as 'urban', 'suburban', and 'rural'.[16] In a next step, the above-mentioned regional characteristics are included. After this, individual characteristics, such as age, gender, educational background, and occupational status, are additionally considered.

The positive impact of network effects as well as of possibilities of learning from others on the adoption process of ICT is described by several authors. For example, Goolsbee and Klenow [6] analyze the importance of such local spillovers in the diffusion process of computers in private households. They find that these spillovers appear to come from experienced and intensive users. Also Mills and Whitacre [14] find a positive association between the regional share of households with Internet access and the individual household's propensities to use the Internet. These positive network effects arise when the individual's benefits of participating in a network increase with the size of the network. Regarding the Internet these effects are obvious as a larger network increases the individual's communication possibilities (especially if family members or friends join the network) as well as the content that is available online. Learning from others is additionally important as experienced users can teach a hitherto non-user how to use the Internet and for what it is useful. Therefore, in a further specification of the analyses on the individual level, the effect of the regional share of experienced users is studied.

The hypotheses analyzed on the individual level are: i) Individuals living in rural areas are less likely to become a home Internet user than individuals in urban areas. ii) The impact of regional characteristics resembles those that can be found on the regional level. iii) The impact of regional determinants declines when individual characteristics are additionally considered. iv) Young, male and highly qualified individuals have a higher probability of starting home Internet use. v) There is a positive network effect: In counties with high shares of experienced Internet users non-using individuals have a higher probability of accessing the Internet for the first time.

---

[16]For details, see BBR [2].

The impact of several individual and regional factors on the individual's decision of becoming a new Internet user is examined by including these variables in a probit model of the form:

$$
\begin{aligned}
prob(Y_i = 1) \;=\; & F(\alpha + \beta \cdot countytype_r + \gamma \cdot C_r + \delta \cdot X_i \\
& + \lambda \cdot usershare_r + \varepsilon_{ri})
\end{aligned}
\tag{2}
$$

$$
\text{with} \quad r = 1 \ldots k \quad \text{and} \quad i = 1 \ldots n,
$$

where Y_i is the dependent variable indicating whether an individual i connects to the Internet at home for the first time in the years 2000 or 2001 ($Y_i = 1$) or not ($Y_i = 0$). The coefficient β describes the impact of the *countytype* of the region r the individual lives in on the connection decision. The coefficient vector γ shows the effects of various county characteristics C_r, δ represents a vector of coefficients regarding individual observables X_i.[17] The variable *usershare* indicates the regional share of experienced Internet users, that is, individuals who declared to use the Internet since 1999 or earlier. The usershare effect on the probability of becoming a new user is measured by λ. The error term ε_{ri} covers individual and regional unobservable characteristics.

5 Empirical Results

5.1 The regional level

Table 3 in the Appendix shows the summary statistics of the characteristics of the 312 analyzed German counties. On average, the county-wide share of home Internet use is 32 percent. It is significantly higher in urban counties (34 percent) than in rural regions (26 percent). Moreover, urban counties have significantly higher shares of highly qualified employees, one-person households, and foreigners, as well as a higher GDP per capita.

---

[17]As in the SOEP the Internet use question was asked in 2001, data on individual characteristics are all taken from this SOEP wave. Regional characteristics included in the estimations are of the year 2000.

The results of the OLS estimations of the share of home Internet use according to equation (1) are depicted in Table 1. The first specification includes rusticity as the only explanatory variable. It shows a negative bivariate correlation with Internet use shares indicating that rural regions experienced a lower Internet penetration compared to city regions. Thus, without considering further determinants, the Internet's complementary function with respect to cities seems to outweigh the substitute function.

As expected, the impact of rusticity declines when including further regional characteristics in the estimation approach, although it remains statistically significant on the 10 percent level (see specification (2)). The results show a significantly positive effect of the share of highly qualified employees on the regional share of home Internet use. An increase of the share of highly qualified by one percentage point increases the rate of home Internet use by about 0.8 percentage points. That finding supports the hypothesis that human capital is an important factor for the technological diffusion process. The regional rates of Internet use are also positively affected by the share of one-person households. This is not surprising as the availability of time and money - two important requirements for using the Internet - is often relatively high in one-person households. An increase in the share of one-person households by one percentage point leads to an increase of the Internet use rate by 0.38 percentage points.

The youths share does not show any significant effect. Thus, the data for German counties do not support the hypothesis that young people induce their families to use new technologies, although adolescents are very likely to use ICT (see Section 5.2).

The regional unemployment rate as well as the share of foreign population show a significantly negative correlation with Internet use penetration, therefore supporting the hypotheses. Unemployed people are often faced with financial rectrictions which make them less likely to own a computer at home and to have access to the Internet. The negative impact of the population share of foreigners could be the result of linguistic problems or shortcomings in education and income. Moreover, it can be intensified by a strong Internet substitute function regarding communication, information, and shopping activities of local minorities, as described above. The more foreigners live in a region the less the Internet is needed to compensate for shortcomings in communication and information possibilities in

the county. A higher local share of foreigners may therefore result in lower local rates of Internet usage.

As mentioned above, the rusticity effect remains statistically significant, even after considering various county-specific characteristics. Internet infrastructure differences are assumed not to be a crucial factor when explaining the divide in private Internet use rates between rural and urban regions in 2001 (see Section 1). However, the regional divide may be driven, for example, by differences in the regional retail market structure, which unfortunately cannot be observed. As using the Internet requires owning a computer and technical accessories the availability of computer retailers can affect the regional rate of Internet diffusion and vice versa: More Internet users increase the demand for IT equipment, leading to an increasing number of computer retail stores, which itself can attract new Internet adopters.

Table 1: Determinants of home Internet diffusion at county level in 2001

dependent variable: population share of home Internet use		
regional characteristics	**(1)**	**(2)**
rusticity^i	-.139 (.028)***	-.082 (.043)*
share of highly qualified employees		.798 (.297)***
share of one-person households		.384 (.154)**
share of population aged betw. 18 and 29^{ii}		-.641 (.689)
unemployment rate		-.809 (.171)***
share of foreign population		-.512 (.253)**
share of female population		-1.358 (1.427)
per capita GDP^{ii}		-.000 (.001)
constant	.349 (.010)***	1.054 (.733)
adjusted R^2	.070	.139
number of observations	312	312

Notes: OLS estimation. ***, **, * depict significance at the 1%, 5% and 10% level. Standard errors in parentheses.
^{i}) Share of population in communities with a population density of less than 150 inhabitants per square kilometer.
^{ii}) Figures for the year 2000. All others are for 2001.
Source: Author's calculations based on SOEP 2001, INKAR 2002, 2003 and 2004.

5.2 The individual level

Table 4 in the Appendix depicts average individual characteristics of new Internet users, non users, and experienced users. It shows that compared to non-users, new users are significantly younger, better educated, and richer. In addition, new users are significantly more likely to work in a full-time job as well as to be male, single, and German. Similarly, compared to new users, those individuals who already use the Internet for more than one year are better educated, richer, and more likely to work in a full-time job as well as to be male, and single. As can be seen in Table 5 in the Appendix, in rural regions there is a higher share of individuals not using the Internet compared to urban areas. The share of experienced users is higher in urban areas. Taking only the hitherto non-using population into consideration it turns out that the share of new users is significantly higher in urban areas (see Table 6). On average, a share of 15 percent of hitherto non-users in rural counties starts using the Internet in the years 2000 or 2001, in urban counties the share is 18 percent.

A probit model (see equation 2) is used to find out whether regional and individual characteristics have a statistically significant impact on the individual's probability of becoming a new Internet user. The results are depicted in Table 2.[18] The first specification examines how the county type affects the tendency to connect to the Internet without considering further control variables. As described by Sinai and Waldfogel [16], such population density figures can be interpreted as a measure of local offline product variety and information availability.[19] Thus, they are useful for analyzing the substitute and complement functions of the Internet.

The results reveal that compared to city regions, individuals living in rural areas have a significantly lower probability of becoming a new user. Similar to the results on the regional level, this denotes a predominant complementarity of the Internet to cities, caused for example by the number of websites offering local information. In order to find out, whether this is an effect of population density

---

[18]All of the standard errors of regional determinants are corrected for the fact that they do not vary between individuals living in the same county.

[19]The higher the local population density the more products and information are available offline, therefore reducing the incentives to use the Internet. On the other hand, a higher population density may increase the amount of local websites offering regional information, therefore increasing incentives of using the Internet.

itself or whether other determinants are able to explain that relationship, several regional and individual factors are included in further specifications. As can be seen by the results of specification (2) in Table 2, the county type effect does not remain significantly different from zero when including further regional determinants. Among these additional variables, only two show significant results: The share of one-person households positively affects the usage probability of hitherto non-users, the unemployment rate has a negative effect. These results were also found on the regional level, where Internet use shares by county were analyzed. The share of foreigners as well as the share of highly qualified employees, which had a significant impact on Internet use rates on county level, do not have coefficients significantly different from zero on the individual level. Also the share of young people, which was anticipated to have a positive impact on the probability of becoming an Internet user, does not show a significant effect. Thus, comparable to the results on the regional level, the data again do not support the hypothesis that young people who are assumed to get in touch with computers very early do measurably act as an Internet use promoter.

It is interesting to see that the regional share of one-person households as well as the unemployment rate remain of importance when individual characteristics are additionally considered (see specification (3) in Table 2). The effects of the individual characteristics largely show the expected signs: The probability to connect to the Internet for the first time decreases with age but increases with the level of education as well as with the level of household income. Males are more likely to connect to the Internet than females, German individuals are more likely to become a new user than foreigners. In addition, part-time employees have a higher probability of starting to use the Internet at home than individuals working full-time. This may be an effect of the additional time that is available at home when individuals are working part-time. Individuals that are retired have a significantly smaller probability of accessing the Internet at home than those working full-time, a result that is found in many other studies, too. Moreover, individuals living alone have a significantly higher probability of connecting to the Internet than individuals in more-person households. Being single shows the opposite effect. However, as the bivariate effect between being single and starting Internet use is significantly positive, this could be the result of the large number of other explaining variables. The results indicate that individual characteristics are more important than regional factors when explaining the individual decision

of becoming a new Internet user.

In order to examine the regional network effect, specification (4) additionally includes the county share of *experienced* Internet users (usershare), that is individuals who are using the Internet for at least one year. The results support the hypothesis that the probability of becoming an Internet user increases with the size of the regional network even after considering various individual and county characteristics. On average, an increase in the usershare by 10 percentage points increases the probability of accessing the Internet by 1.5 percentage points. Thus, the size of the network effect is much smaller than the effect found by Goolsbee and Klenow [6] when analyzing the diffusion of home computers in the U.S. They do not consider the range of regional characteristics included here, but even excluding those determinants from the model does not change the results very much. The increase in the probability of using the Internet then amounts to 2.1 percentage points (not depicted in the table).

As the usershare effect remains significant even after considering the regional youths share and the share of highly qualified employees, the network effect seems to go beyond learning spillovers provoked by these groups. However, family members and friends who live in the same county and who already use the Internet are reasonable causes of the usershare effect. Moreover, it would be interesting to consider differences in ICT prices or in the structure of the ICT retail market between counties as this can additionally cause the differences in the usage probability detected by the usershare effect. Unfortunately, this information is not provided in the data set.

Besides the usershare effect, the share of one-person households is the only regional variable that remains with a significant impact on the probability of accessing the Internet for the first time. The underlying causes of this effect have to be further explored.

Table 2: Determinants of starting home Internet use in 2000 or 2001

dependent variable: probability of starting home Internet use				
variable (reference group)	(1)	(2)	(3)	(4)
regional characteristics				
county type (ref.: urban)				
rural	-.024 (.012)*	-.023 (.015)	-.023 (.014)	-.018 (.014)
suburban	-.003 (.011)	-.008 (.012)	-.013 (.011)	-.009 (.011)
share of one-person households		.260 (.121)**	.327 (.103)***	.280 (.106)***
unemployment rate		-.386 (.131)***	-.207 (.119)*	-.138 (.115)
share of foreign population		-.322 (.196)	-.231 (.176)	-.207 (.170)
share of highly qualified employees		.237 (.200)	.260 (.202)	.164 (.197)
share of population aged				
between 18 and 29		.629 (.503)	-.136 (.502)	-.054 (.517)
share of female population		-.544 (1.013)	-.992 (.973)	-.890 (.961)
per capita GDP		-.001 (.001)	-.001 (.001)	-.001 (.001)
usershare[i]				.146 (.058)**
individual characteristics				
age in years (ref.: age less than 25)				
25-34			.058 (.016)***	.058 (.016)***
35-44			-.042 (.010)***	-.042 (.010)***
45-54			-.091 (.010)***	-.092 (.010)***
55-64			-.144 (.009)***	-.145 (.009)***
male			.043 (.007)***	.044 (.007)***
single			-.021 (.010)**	-.021 (.010)**
one-person household			.044 (.015)***	.040 (.015)***
German nationality (ref: foreigner)			.095 (.010)***	.094 (.010)***
education (ref.: university degree)				
lower secondary education or less			-.123 (.011)***	-.123 (.011)***
other vocational education			-.091 (.017)***	-.092 (.017)***
apprenticeship			-.119 (.013)***	-.119 (.013)***
special. vocational school			-.083 (.011)***	-.082 (.011)***
technical school			-.069 (.015)***	-.068 (.015)***
civil servant school			-.058 (.019)***	-.060 (.019)***
polytechnical or college abroad[ii]			-.050 (.012)***	-.050 (.012)***
occup. status (ref.: employed full-time)				
employed part-time			.034 (.011)***	.034 (.011)***
apprentice			-.009 (.017)	-.008 (.017)
not employed			.002 (.009)	.002 (.009)
retired			-.049 (.013)***	-.048 (.013)***
log net income of household			.111 (.011)***	.109 (.011)***
pseudo-R^2	.001	.003	.095	.096
number of observations	13.626	13.626	12.480	12.480

Notes: Probit estimation, marginal effects, robust standard errors (clustered at regional level). Regional figures are for the year 2000, individual characteristics for 2001. ***, **, * depict significance at the 1%, 5% and 10% level. Standard errors in parentheses.

[i]) Share of Internet users with more than one year usage experience by county.

[ii]) College abroad: In the data it is not clear what kind of degree is meant.

Source: Author's calculations based on SOEP 2001, INKAR 2002 and 2003, Statistik regional 2002.

6 Concluding Remarks

Although in Germany the Internet has spread rapidly in recent years, there are still large differences in accessing the Internet between population groups. This so called 'digital divide' has many dimensions. Besides differences caused by individual characteristics such as age, educational background, and income, there is also a regional gap in ICT use: Rural regions possess smaller Internet use rates than cities. In order to reduce the existing divides it is crucial to know what the driving forces are behind that development. The attempt of this paper is to analyze descriptively as well as econometrically the relationship between regional characteristics and home Internet use.

On the regional level, the results of the multivariate analyses support the hypothesis that a higher rusticity is related to a lower Internet use rate. Besides, they identify several county characteristics that are correlated with home Internet use rates. As expected, the share of highly qualified employees as well as the share of one-person households have a positive impact on regional Internet use penetration. Regions with a higher unemployment rate show a smaller Internet use share. Also the share of foreign population shows a negative impact on the diffusion of Internet use at home. Although shrinking a bit the rusticity effect remains statistically significant after including several regional characteristics. The effect therefore captures differences between rural and urban areas that are not explainable by the other included characteristics.

On the individual level, regional characteristics show much less influence than on the regional level. After including various individual characteristics only the share of one-person households remains significant. In addition, there is a positive network effect observable. Thus, it is not the fact that individuals live in a rural region that reduces their probability of becoming an Internet user. The results rather indicate that it is the different composition of the population between rural and urban areas that causes the regional digital divide. Therefore, it can be concluded that policies targeted at decreasing the digital divide should primarily provide programs which encourage the Internet literacy of less qualified, unemployed, and older individuals. In addition, experienced users should be involved in order to motivate non-users by teaching them how to use the Internet and by showing them its usefulness.

Acknowledgments

References

[1] L. Borghans, B. ter Weel, Do Older Workers Have More Trouble Using a Computer Than Younger Workers?, Research in Labor Economics 21 (2002), pp. 139–173.

[2] BBR - Bundesamt für Bauwesen und Raumordnung, Aktuelle Daten zur Entwicklung der Städte, Kreise und Gemeinden, Berichte, Band 14, 2002.

[3] H. Entorf, M. Gollac, F. Kramarz, New Technologies, Wages, and Worker Selection, Journal of Labor Economics 17(3), 1999, pp. 464–491.

[4] J. Gaspar, E. Glaeser, Information Technology and the Future of Cities, Journal of Urban Economics 43(1), 1998, pp. 136–156.

[5] A. Goldfarb, The (Teaching) Role of Universities in the Diffusion of the Internet, International Journal of Industrial Organization, 24(2), 2006, pp. 203–225.

[6] A. Goolsbee, P. Klenow, Evidence on Learning and Network Externalities in the Diffusion of Home Computers, The Journal of Law and Economics 45 (2002), pp. 317–344.

[7] J. Haisken-DeNew, J. Frick, Desktop Companion to the German Socio-Economic Panel Study (DTC), http://www.diw.de/deutsch/sop/service/dtc/dtc.pdf, 2005.

[8] TNS Infratest, (N)Onliner-Atlas 2002 - Eine Topographie des digitalen Grabens durch Deutschland, 2002.

[9] TNS Infratest, (N)Onliner-Atlas 2006 - Eine Topographie des digitalen Grabens durch Deutschland, 2006.

20

[10] M. Jäckel, T. Lenz, N. Zillien, Die regionale digitale Spaltung Eine empirische Studie zu Unterschieden in der Internetnutzung in Stadt und Land Schriftenreihe zum E-Business (Band 4), 2005.

[11] S. Korupp, H. Künemund, J. Schupp, Digitale Spaltung in Deutschland: Geringere Bildung - seltener am PC. DIW Wochenbericht 73(19), 2006, pp. 289–294.

[12] S. Korupp, M. Szydlik, Causes and Trends of the Digital Divide, European Sociological Review 21(4), 2005, pp. 409–422.

[13] A. Lenhart, The Ever-Shifting Internet Population - A New Look at Internet Access and the Digital Divide, Pew Internet & American Life Project, 2003.

[14] B. Mills, B. Whitacre, Understanding the Non-Metropolitan-Metropolitan Digital Divide, Growth and Change 34(2), 2003, pp. 219–243.

[15] OECD/DSTI, Understanding the Digital Divide, 2001.

[16] T. Sinai, J. Waldfogel Geography and the Internet: Is the Internet a Substitute or a Complement for Cities? Journal of Urban Economics 56(1), 2004, pp. 1–24.

Appendix

Table 3: Regional characteristics of German counties, 2001

regional variable	total	rural	suburban	urban
share of Internet users	0.32	0.26***	0.31	0.34
rusticity[i]	0.24	0.55***	0.25***	0.12
share of highly qualified employees	0.07	0.05***	0.07***	0.09
share of one-person households	0.34	0.30***	0.33	0.35
share of population aged between 18 and 29	0.14	0.14	0.14***	0.14
unemployment rate	0.11	0.11	0.11	0.10
share of foreign population	0.07	0.04***	0.06***	0.10
share of female population	0.51	0.51**	0.51	0.51
per capita GDP (in 1,000 Euro)	22.20	17.93***	21.31***	24.83
number of counties	312	49	137	126

Notes: Mean values of regional figures for the year 2001.

***, ** indicate that means are significantly different from urban means at the 1% and 5% level.

[i]) Share of population in communities with a population density of less than 150 inhabitants per square kilometer.

Source: Author's calculations based on SOEP 2001, INKAR 2002 and 2003, Statistik regional 2002.

Table 4: Characteristics of new Internet users, non-users, and experienced users[i], 2001

individual variable	non-users	new users	experienced users
number of individuals	11,280	2,346	3,036
age in years	42.6	36.8***	36.7
male	0.45	0.51***	0.62***
single	0.20	0.27***	0.34***
one-person-household[ii]	0.10	0.08***	0.11***
German nationality	0.87	0.94***	0.95
education			
number of individuals[iii]	11,028	2,293	2,973
lower secondary education or less	0.26	0.22***	0.20
other vocational education	0.02	0.01	0.01
apprenticeship	0.42	0.35***	0.29***
specialized vocational school	0.11	0.11	0.08***
technical school	0.05	0.06**	0.07
civil servant school	0.02	0.03***	0.03
polytechnical or college abroad[iv]	0.08	0.11***	0.12**
university	0.05	0.13***	0.20***
occupational status			
number of individuals[ii]	11,159	2,299	2,984
employed full-time	0.46	0.54***	0.61***
employed part-time	0.15	0.17***	0.14***
apprentice	0.04	0.05***	0.04***
not employed	0.25	0.21***	0.20
retired	0.11	0.03***	0.02***
income of household (in 1,000 Euro)	4.44	5.44***	5.87***

Notes: Mean values of individual characteristics for 2001.

***, ** indicate that means are significantly different at the 1% and 5% level: means of new users are compared with those of non-users, means of experienced users are compared with those of new users.

[i]) Experienced Internet users are those with more than one year usage experience.

[ii]) In the data, two non-using individuals have missing values for that characteristic.

[iii]) Differences in the number of observations originate from missing values in the data set.

[iv]) College abroad: In the data it is not clear what kind of degree is meant.

Example: In 2001, the share of singles among the new users is 27 percent, among the non-users the single share is 20 percent.

Source: Author's calculations based on SOEP 2001.

Table 5: Characteristics of rural and urban regions, 2000

regional variable	total	rural	urban
new users[i]	0.14	0.13***	0.14
non-users	0.68	0.74***	0.65
experienced users[ii]	0.18	0.13***	0.21
share of highly qualified employees	0.08	0.05***	0.09
share of one-person households	0.35	0.29***	0.37
share of population aged between 18 and 29	0.14	0.14**	0.14
unemployment rate	0.11	0.13***	0.11
share of foreign population	0.09	0.04***	0.11
share of female population	0.51	0.51***	0.51
per capita GDP (in 1,000 Euro)	23.87	17.70***	26.72
number of individuals	16,662	1,839	9,036
number of counties	312	49	170

Notes: Mean values of regional figures for the year 2000, calculated by individuals.

***, ** indicate that rural means are significantly different from urban means at the 1% and 5% level.

[i]) Population share of Internet users who start usage in 2000 or 2001, by county.

[ii]) Population share of Internet users with more than one year usage experience by county.

Source: Author's calculations based on SOEP 2001, INKAR 2002 and 2003, Statistik regional 2002.

Table 6: Share of new users and non-users on hitherto non-users, 2000/2001

regional variable	total	rural	urban
new users[i]	0.17	0.15***	0.18
non-users	0.83	0.85***	0.82
number of individuals	13,626	1,603	7,173
number of counties	312	49	170

Notes: Mean values of regional figures for the period 2000/2001, calculated by individuals.

*** indicate that rural means are significantly different from urban means at the 1% level.

[i]) Share of those who become new users in 2000 or 2001 on hitherto non-users.

Source: Author's calculations based on SOEP 2001, INKAR 2002 and 2003.