12-1-2017

Internet and Income Inequality: A Research Note

Yongzheng Liu
Renmin Univeristy of China, yonghzeng.liu@ruc.edu.cn

International Center for Public Policy

International Center for Public Policy
Working Paper 17-23
December 2017

Internet and Income Inequality: A Research Note

Yongzheng Liu

Georgia State University
Andrew Young School of Policy Studies

Internet and Income Inequality: A Research Note[1]

Yongzheng Liu
School of Finance
China Financial Policy Research Center
Renmin University of China
E-mail: yonghzeng.liu@ruc.edu.cn

Abstract

By using a sample of 51 developed and developing countries, this research note empirically examines the impact of internet diffusion on income inequality. To address the potential endogeneity issue of internet diffusion, I employ lightning density as an instrument for internet diffusion and use an instrumental variable method for the estimations. I find that internet diffusion significantly reduces income inequality. The results are robust across alternative specifications.

Keywords: Internet diffusion; Income inequality; Lightning density; Endogeneity
JEL classifications: D63; L86

[1] This work was supported by the National Natural Science Foundation of China (Nos. 71403276; 71773128; 71533006) and the Fok Ying Tung Education Foundation (No.151085). I would like to thank Jia Li and Yan Li for able research assistance.

1. Introduction

As a worldwide phenomenon, income inequality has received increasing attention from academic, policy, and media circles in the past decades. Our understanding of this issue, especially the determinants of income inequality, has been considerably improved by the growing discussion and empirical work on this subject. The current literature has identified a wide range of factors that may affect income inequality in a region, including selected macroeconomic variables such as economic growth, unemployment, and inflation (e.g., Blejer and Guererro 1990; Deininger and Squire 1996; Mocan 1999), political and institutional factors (e.g., Bourguignon and Morrisson 1998; Li *et al.* 2000), and specific public policies (e.g., Feenberg and Poterba 1993; Auten and Carroll 1999). In this research note, I contribute to the literature by offering a novel determinant of income inequality, namely internet diffusion, which has spread globally and influenced the world economy.

In theory, internet diffusion may affect income inequality in a region either positively or negatively, depending on who benefits in which way from the use of the internet. On the one hand, several reasons support the internet's potential positive effect in reducing income inequality. First, internet diffusion may help lower living costs through global supply chains. By building a more convenient global medium for information exchange, the internet helps allocate materials and manufactures on a global scale more efficiently, which lowers production costs and selling expenses and eventually leads to a more reasonable price for customers. Meanwhile, electronic commerce and internet marketplaces serve as alternatives or supplements to traditional retail markets, which benefits consumers by improving their convenience and expanding their choices (McQuivey *et al.* 1998; Litan and Rivlin 2001). Additionally, internet diffusion tends to reduce consumers' search costs, resulting in a significant growth in their market power and thus larger welfare gains (Brown and Goolsbee 2002). Insofar as the reduction in living costs benefits the poor more, the internet is likely to have a positive effect on reducing income inequality. Second, the extensive use of the internet and its related technologies creates a large number of new job opportunities that mostly increase the incomes of the low-income group, lifting millions of workers out of poverty and leading to an adjustment of the income distribution (Bauer 2015). Third, internet diffusion improves the democratic environment by allowing citizens to express their views and offer feedback on public policies in a more effective way and with a higher degree of freedom (Lee and Heshmati 2017). These measures largely reduce bureaucratic discretion and increase governments' responsiveness to the needs and rights of ordinary people.

On the other hand, internet diffusion may increase income inequality when online access is unequally distributed among populations and favored toward people with higher social status (Hargittai 1999). This is so because the internet divide may act as an additional avenue to exclude people that lack the internet from participation in the national online economy; at the same time, it may also exclude those people from the social and human capital that flows online (Lentz and Oden 2001). As a result, informational illiteracy and a lack of internet access may reinforce other economic and cultural disparities (Howard *et al.* 2010).

Given these opposite predictions, the net impact of internet diffusion on income inequality remains an open question for empirical examination. To date, the literature has only provided some suggestive and indirect evidence in this regard. For instance, by using national survey data, Willis and Tranter (2006) find that internet use in Australia is structured by complex inequalities in terms of users' income, age, gender, education, and occupational class, which has significant implications for the internet's impacts on income inequality. Similarly, by using survey data from the Netherlands, van Deursen and Helsper (2015) suggest that highly educated individuals benefit more from the internet than those with less education, potentially implying that existing offline inequalities could be amplified by internet diffusion. Nevertheless, Howard *et al.* (2010) provide evidence that in Canada, the concentration of internet access among wealthy educated populations has been significantly reduced, in part because of the active role of governments in supporting the provision of culturally relevant digital content.

In this research note, I add to the literature by directly testing the causal influence of internet diffusion on income inequality by using cross-country data; equally importantly, I address the endogeneity issue in the estimations to obtain an unbiased estimate of internet diffusion.

2. Empirical Strategy and Data

The basic specification I employ here takes the following form:[2]

$$\Delta Gini_i = \alpha + \beta \Delta Internet_i + \gamma Gini_{0i} + \delta \Delta \boldsymbol{X}_i + \varepsilon_i \quad (1)$$

where $\Delta Gini_i$ is the change in income inequality (i.e., the Gini coefficient in my study) between an initial ($Gini_{0i}$) and a final ($Gini_{1i}$) year, namely, $Gini_{1i} - Gini_{0i}$; $\Delta Internet_i$ is the corresponding change in internet diffusion between the two years; and ε_i is an idiosyncratic error term. Support of my previous hypothesis would predict a negative coefficient for the change in internet diffusion (i.e., β), implying that the internet is a useful technology for reducing income inequality.

I include the initial level of income inequality $Gini_{0i}$ as the main control variable in the specification. The inclusion of the initial level of income inequality captures a series of time-invariant institutional and structural characteristics that may persistently affect income inequality; thus, the inclusion of this variable helps reduce omitted variable bias to a large extent. Additionally, I include in the specification other general factors of significance (i.e., $\Delta \boldsymbol{X}_i$) to determine income inequality based on the extant empirical literature. These include real GDP per capita, total population, and the level of education. Similarly, I measure these variables by taking their changes between the initial and the final year.

An important concern for estimating specification (1) is the potential endogeneity of internet diffusion. This issue may arise because of both reverse causality and the potential omitted variables. For instance, internet diffusion may increase the visibility of income inequality in society by accelerating information exchange, which threatens incumbent politicians. Consequently, countries with severe income inequality may have stronger incentives to block the adoption of new technology. In addition, some common factors may affect both internet diffusion and income inequality in a country simultaneously. To circumvent the endogeneity issue, I use an instrumental variable approach. I follow Andersen *et al.*'s (2011) approach to use lightning density in the country as an instrument for internet diffusion. The rationale here is that lightning activity is a *natural source* of power disruption,[3] which increases the user cost of IT capital by damaging IT equipment and thus lowers the

[2] The specification is similar to that in Andersen *et al*. (2011).
[3] As pointed out by Andersen *et al*. (2011), one-third of all power disruptions in the United States are related to lightning activity.

speed of internet diffusion. However, while the relevance of lightning density and internet diffusion appears to be easy to justify, this instrument may still be invalid if conditional on internet diffusion, lightning activities affect income inequality through other channels. I argue that by controlling for general determinants of income inequality such as economic development, total population, and openness in the specification, I reduce this concern to a large extent. Nevertheless, in the next section, I present the results of formal tests to confirm the validity of the instrument.

The data I use include 51 developed and developing countries for 1991 and 2005. My measure of income inequality, the Gini coefficient, is obtained from the UNU-WDIER World Income Inequality database. Internet diffusion for the cross-country sample is calculated as the number of internet users per 100 people and is provided by the World Development Indicators (WDI) database. Since the internet, in the sense of the appearance of the first World Wide Web, was launched in 1991, I use year 1991 as the starting period in my analysis. For the same reason, the initial value for internet diffusion is zero. Real GDP per capita and total population are obtained from the WDI database. Openness is calculated as the ratio of total trade (imports plus exports) to GDP and this is also provided by the WDI database. The education level of a country is captured by the average years of schooling for populations 15 years old and above and is obtained from Barro and Lee (2001).

The instrument, lightning density, is captured by satellite data on lightning intensity, for which the raw data (strikes per km^2 per year) are provided by the National Aeronautics and Space Administration (NASA). Specifically, given data availability, these data on lightning density are the average flash density for each country over a five-year period (i.e., April 12, 1995 to December 31, 1999).[4] They are obtained directly from Andersen et al. (2011).[5] Table 1 presents the summary statistics of the key variables of interest.

Table 1. Summary Statistics

Variable	**Obs.**	**Mean**	**Std. Dev.**	**Min**	**Max**
Δlog(Gini)	51	0.051	0.160	-0.273	0.426
Δlog(Internet)	51	3.004	1.196	-1.344	4.335

[4] Recall that the initial value of internet diffusion (in 1991) is zero. Thus, the “change” in internet diffusion between 1991 and 2005 is actually the same as the internet diffusion level in 2005. For this reason, I use only one period of lightning data to predict the “change” in internet diffusion.

[5] For more details about the construction of the variable, see Andersen *et al.* (2011).

log(lightning)	51	1.51	1.27	-1.76	3.32
log($Gini_{1991}$)	51	3.512	0.302	2.890	4.075
Δlog(GDP per capita)	49	0.294	0.171	-0.284	0.612
Δlog(population)	51	0.110	0.138	-0.148	0.366
Δlog(openness)	50	0.324	0.321	-0.588	1.169
Δlog(education)	49	0.188	0.130	0.000	0.524

Note: Δ represents the differences between 2005 and 1991. For example, Δlog(Gini) is the difference of logarithm of Gini in 2005 and 1991.

3. Main Results and Robustness Checks

3.1 Main Results

Table 2 reports both the OLS and the 2SLS estimation results, with and without controlling for the initial level of income inequality and other determinants of income inequality.

First, I examine the effect of internet diffusion on income inequality without adding any other explanatory variables into the model. As shown in Column (1) of Table 2, the coefficient of internet diffusion is positive but statistically insignificant. When I include the initial level of income inequality in the specification in Column (2), the estimate becomes negative, although it is still statistically insignificant. I then continuously add to the model with a set of other explanatory variables in Column (3). I find a negative and statistically significant coefficient of internet diffusion (i.e., Δlog(Internet)), suggesting that internet diffusion helps reduce income inequality.

Table 2. OLS and 2SLS Regression Results

	OLS			2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)
Δlog(Internet)	0.003	-0.034	-0.070**	0.010	-0.085**	-0.108***
	(0.024)	(0.030)	(0.034)	(0.029)	(0.037)	(0.035)
log($Gini_{1991}$)		-0.321***	-0.160*		-0.414***	-0.187**
		(0.065)	(0.081)		(0.078)	(0.081)
Δlog(GDP per capita)			0.045			0.087
			(0.122)			(0.109)
Δlog(population)			-0.440**			-0.494***
			(0.190)			(0.176)
Δlog(openness)			0.069*			0.064*
			(0.039)			(0.038)
Δlog(education)			-0.309**			-0.397**
			(0.143)			(0.191)
Constant	0.041	1.279***	0.906**	0.023	1.761***	1.128***

	(0.078)	(0.295)	(0.342)	(0.092)	(0.358)	(0.344)
Cragg-Donald F Statistic	-	-	-	29.13	14.63	12.30
Observations	51	51	47	51	51	47
R-squared	0.001	0.293	0.462	-0.001	0.175	0.410

Note: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Next, I turn to the 2SLS estimation results. Before I proceed, I provide some evidence for the validity of the selected instrument. I first estimate the following first-stage specification to show the relevance between the instrument and internet diffusion:

$$\Delta Internet_i = \alpha + \beta Z_i + \gamma Gini_{0i} + \delta \Delta \boldsymbol{X}_i + \varepsilon_i \quad (2)$$

where Z_i represents my measure of lightning density in country i. All the other variables remain the same as in specification (1). Table 3 reports the corresponding first-stage estimation results. As shown, lightning density is negatively and significantly correlated with internet diffusion, confirming the prediction that lightning activities may damage IT equipment and hence lower the speed of internet diffusion. Meanwhile, for all three specifications, the F-statistic is always over 10, suggesting that my instrumental variable estimates are not prone to the weak instrument concern.

Table 3. First-stage Estimation Results for the IV Estimations

	(1)	(2)	(3)
log(lightning)	-0.576***	-0.487***	-0.477***
	(0.122)	(0.130)	(0.117)
log($Gini_{1991}$)		-0.677	0.225
		(0.443)	(0.628)
Δlog(GDP per capita)			1.061
			(0.770)
Δlog(population)			-1.252
			(1.551)
Δlog(openness)			-0.114
			(0.340)
Δlog(education)			-1.539
			(1.419)
Constant	3.872***	6.116***	3.156
	(0.161)	(1.467)	(1.877)
Observations	51	51	47
Cragg-Donald Wald F statistic	29.13	14.63	12.31

Note: The dependent variable in the first-stage regressions is the Δlog(Internet). Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Next, I formally check that the instrumental variable satisfies the exclusion condition; in other words, conditional on internet diffusion, lightning density is not correlated with the residual error (i.e., ε_i). I conduct a test following Andersen *et al.* (2011). The premise for the test is that if lightning density affects income inequality only through internet diffusion, then it should have no impact on income inequality before the inception of the World Wide Web in 1991. To validate this, in Table 4 I examine two pre-internet periods, 1975–1990 and 1980–1990. I include lightning density as an explanatory variable in specification (1) and exclude internet diffusion for obvious reasons.[6] The results reveal that lightning density is not correlated with the changes in income inequality in these two pre-internet periods, shedding some light on the satisfaction of the exclusion restriction of the selected instrument. As a comparison, I also report the result for 1991–2005 in Column (3) of Table 4, where I find a positive estimate of lightning density, confirming the validity of the instrument and the potential impact of lightning density on income inequality through internet diffusion.

Table 4. Tests for Exclusion Restriction

	(1)	(2)	(3)
	1975-1990	1980-1990	1991-2005
log(lightning)	0.042	-0.004	0.051**
	(0.034)	(0.016)	(0.021)
log($Gini_0$)	-0.482***	-0.341***	-0.212***
	(0.135)	(0.078)	(0.076)
Δlog(GDP per capita)	0.067	-0.222	-0.027
	(0.106)	(0.132)	(0.127)
Δlog(population)	0.943***	0.731***	-0.359
	(0.242)	(0.201)	(0.231)
Δlog(openness)	-0.004*	0.000	0.076
	(0.002)	(0.002)	(0.049)
Δlog(education)	-0.361*	-0.153	-0.231
	(0.211)	(0.192)	(0.156)
Constant	1.561***	1.206***	0.788***
	(0.450)	(0.264)	(0.254)
Observations	35	43	47
R-squared	0.534	0.398	0.384

Note: The dependent variable is the Δlog(Gini). Δ represents the difference between the initial and the final year as indicated on the top of each column. log($Gini_0$) represents logarithm of Gini index in 1975, 1980, and 1991 for Columns (1)-(3), respectively. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Having shown the validity of the instrument, I estimate the second-stage specification of

[6] 1991 was the founding year of the World Wide Web, and hence internet diffusion did not exist before then.

the 2SLS estimation as follows:

$$\Delta Gini_i = \alpha + \beta \Delta \widehat{Internet}_i + \gamma Gini_{0i} + \delta \Delta \boldsymbol{X}_i + \varepsilon_i \quad (3)$$

where $\Delta \widehat{Internet}_i$ is the predicted value of the dependent variable in the estimation of the first-stage specification (2). As shown in Columns (4)–(6) of Table 2, I find the coefficient of internet diffusion to be positive but statistically insignificant in Column (4). Nevertheless, after I control for the initial level of income inequality in Column (5), the coefficient of internet diffusion becomes negative and statistically significant at the 5% level, which is consistent with the OLS results. This result persists when I add to the model with other explanatory variables in Column (6). Quantitatively, this finding indicates that, on average, a one percentage point increase in internet users in a country is associated with a 0.108 point reduction in the Gini coefficient of the country.

Finally, the initial level of income inequality has a negative coefficient, statistically significant at the 1% level, which may be interpreted as the effects of a series of heterogeneous institutional factors in explaining the trend of income inequality across countries. All the other control variables tend to have statistically significant coefficients and the results are mostly consistent with the existing literature.

3.2 Robustness Checks

To check the robustness of the main results, I conduct a sensitivity analysis along two dimensions. First, I examine an alternative time period (1991–1999) to see if the aforementioned results might be driven by the particular time period selected. Second, given that economic development and income inequality may be correlated at a certain geographical level, I respond to this concern by alternatively considering standard errors at the regional level.[7] Panels A and B of Table 5 report the corresponding robustness results along with the above two dimensions, respectively. As shown, the results are virtually unchanged with these alternative samples and standard errors,[8] confirming the main finding that internet diffusion acts as an effective tool for reducing income inequality.

7 I borrow the World Bank's classification of economic regions: Sub-Saharan Arica, East Asia and the Pacific, Europe and Central Asia, Latin America and the Caribbean, Middle East and North Africa, South Asia, and Advanced Economies.

8 As shown in Panel B of Table 5, although the OLS estimates of internet diffusion only become statistically significant at the margin, the 2SLS estimates remain statistically significant.

Table 5. Robustness Checks

	OLS			2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Alternative time period (1991-1999)					
Δlog(Internet)	-0.004	-0.021**	-0.025***	-0.002	-0.058***	-0.060***
	(0.009)	(0.010)	(0.009)	(0.014)	(0.020)	(0.022)
log($Gini_{1991}$)		-0.286***	-0.148**		-0.391***	-0.178**
		(0.067)	(0.068)		(0.081)	(0.078)
Other controls	No	No	Yes	No	No	Yes
Observations	77	77	68	77	77	68
R-squared	0.002	0.261	0.315	0.002	0.056	0.202
Cragg-Donald F Statistic	-	-	-	26.42	13.02	12.77
	Panel B: Alternative standard errors					
Δlog(Internet)	0.003	-0.034	-0.070	0.010	-0.085*	-0.108***
	(0.023)	(0.040)	(0.039)	(0.026)	(0.044)	(0.015)
log($Gini_{1991}$)		-0.321**	-0.160**		-0.414***	-0.187***
		(0.114)	(0.053)		(0.106)	(0.054)
Other controls	No	No	Yes	No	No	Yes
Observations	51	51	47	51	51	47
R-squared	0.001	0.293	0.462	-0.001	0.175	0.410
Cragg-Donald F Statistic	-	-	-	29.13	14.63	12.30

Note: Panel A reports the estimation results for an alternative time period (1991-1999); Panel B reports the estimation results with standard errors clustering at regional level. Other control variables include Δlog(GDP per capita), Δlog(population), Δlog(openness), and Δlog(education). Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

4. Conclusion

By drawing on cross-country data for 1991 and 2005, I examine the causal impact of internet diffusion on income inequality in a country by using an instrumental variable estimation approach. My estimation result suggests that internet diffusion does serve as a useful technology for reducing income inequality, which is consistent with some existing theoretical arguments. Thus, I contribute to the literature by adding additional evidence on the potential social impacts of internet diffusion. From a policy perspective, promoting the development of the internet worldwide will have a positive consequence on improving social fairness. For further research, it would be interesting to explore the exact channels through which the internet has driven down income inequality.

References

Andersen, T. B, Bentzen, J, Dalgaard, C J, and P. Selaya (2011) "Does The Internet Reduce Corruption? Evidence from US States and Across Countries" *The World Bank Economic Review* 25, 387-417.

Auten, G and R. Carroll (1999) "The Effect of Income Taxes on Household Income" *Review of Economics and Statistics* 81, 681-693.

Barro, R J and J. W. Lee (2001) "International Data on Educational Attainment: Updates and Implications" *Oxford Economic Papers* 53, 541-563.

Bauer, J. (2015) "Internet Connectivity and Income Inequality" *Paper presented at the 27th Annual Meeting of Society for the Advancement of Socio-Economics*, London, England.

Blejer, M I and I. Guerrero (1990) "The Impact of Macroeconomic Policies on Income Distribution: An Empirical Study of the Philippines" *The Review of Economics and Statistics* 72, 414-423.

Bourguignon, F and C. Morrisson (1998) "Inequality and Development: The Role of Dualism" *Journal of Development Economics* 57, 233-257.

Brown, J R and A. Goolsbee (2002) "Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry" *Journal of Political Economy* 110, 481-507.

Deininger, K and L. Squire (1996) "A New Data Set Measuring Income Inequality" *The World Bank Economic Review* 10, 565-591.

Feenberg, D R and J. M. Poterba (1993) "Income Inequality and the Incomes of Very High-income Taxpayers: Evidence from Tax Returns" *Tax Policy and the Economy* 7, 145-177.

Hargittai, E. (1999) "Weaving the Western Web: Explaining Differences in Internet Connectivity Among OECD Countries" *Telecommunications policy* 23, 701-718.

Howard, P N, Busch, L, and P. Sheets (2010) "Comparing Digitai Divides: Internet Access and Social Inequality in Canada and the United States" *Canadian Journal of Communication* 35, 109.

Lee, Min-Kyu and A. Heshmati (2017) "Analysis of the Multinational Diffusion of the Internet", unpublished manuscript.

Lentz, R G and M. D. Oden (2001) "Digital Divide or Digital Opportunity in the Mississippi Delta Region of the US" *Telecommunications policy* 25, 291-313.

Li, H, Xu, L C, and H. F. Zou (2000) "Corruption, Income Distribution, and Growth" *Economics & Politics* 12, 155-182.

Litan, R E and A. M. Rivlin (2001) "Projecting the economic impact of the internet" *American Economic Review* 91, 313-317.

McQuivey, J, Delhagen, K, Levin, K, and M. L. Kadison (1998) "Retail's Growth Spiral" *The Forrester Report* 1, 1-34.

Mocan, H. N. (1999) "Structural Unemployment, Cyclical Unemployment, and Income Inequality" *Review of Economics and Statistics* 81, 122-134.

van Deursen, A J A M and E. J. Helsper (2015) "The Third Level Digital Divide: Who Benefits Most from Being Online?" in *Communication and Information Technologies Annual* by L. Robinson, S. R. Cotten, and J. Schulz, Eds., Emerald Publishing, 29-52.

Willis, S and B. Tranter (2006) "Beyond the 'Digital Divide' Internet Diffusion and Inequality in Australia" *Journal of Sociology* 42, 43-59.