The More You Know: Estimating the Information Effects of Internet Usage on Inequality in a Simultaneous Equations Framework

Blair Long
Queen's University

August 1, 2011

Abstract

I analyze effect on inequality of increased access to information via internet usage using a simple two-period model with no saving in which there are complementarities between education, income and internet access. This relationship is estimated as a single equation and as a system using instrumental variables. 2SLS, SUR and 3SLS estimation is employed to account for interdependence amongst key macroeconomic variables with internet usage substituted for technological progress. My findings indicate that the facilitated access to information from internet access may have a unique effect on inequality which is independent of complementarities with education and income, though it is clear that educational and income differences may polarize incomes in spite of internet access. This is interpreted as the extent to which information benefits the wealthy versus the poor and indicates the extent to which complementarities matter.

Contents

List of Figures

List of Tables

"If information and knowledge are central to democracy, they are conditions for development."

-UN Secretary General Kofi Anaan, Press Release, June 23, 1997

1 Introduction

Technological progress has been seen as a central determinant of economic growth and development since the era of classical political economy. The incorporation of technology in the growth engine has persisted through the analyses of Schumpeter, is common in neoclassical growth models (most notably Solow 1956), and is a common attribute of most contemporary growth models[1]. The impact of technological progress on economic growth is not as "cut and dry" as one may initially think. Certainly, innovation generates enormous efficiencies for firms and workers. This progress is often dependant upon complementarities with other drivers of growth such as education. Numerous authors have evaluated the human capital effects of technological spillovers in foreign direct investment (FDI) on economic development. Examples include Anwar and Nguyen (2010), Krammer (2010) and Kristjansdottir (2010). The instantaneous internet access provided by any modern "smartphone" has allowed information to influence daily consumer decisions in a way unforeseen even a few years ago. Prices can be compared amongst competitors, product reviews and personal finance information are available instantaneously. Still, the long-run and indirect impacts of technological diffusion on society extend further then its impact on productivity and growth. For example, the invention and subsequent mass adoption of the automobile is principally responsible for the shift towards "sprawl" in city planning seen in the twentieth century. Such a change in the geographic and spatial distribution of firms and other agents has clearly changed the nature of commerce.

The complicated process that governs the state of technology is analogous to the

[1]The most obvious example being Romer (1990).

contemporary paradigm of development economics. While the diffusion of certain technologies can impact numerous elements of society in the long-run, the modern view of development requires the coordination of multiple sectors of the economy, institutions and national endowments (natural resources, geography, soil quality, etc) to bring about growth alongside the alleviation of inequality and poverty. The pervasive impact of technology and its ability to reshape the way society is organized, compensate for weak links in the productivity chain[2] and provide access to information that is otherwise unavailable creates a requisite need to identify the determinants technological diffusion, the unique processes that govern diffusion in heterogeneous economic circumstances and the economic outcomes of adoption.

Amongst the most influential technologies of the past 20 years is the internet. The spread of internet usage worldwide has diffused at an unprecedented pace. Figure 1 depicts the trend of internet usage worldwide from 1990 to 2008. There are several active campaigns to increase internet usage amongst low-income individuals and in the developing world. The mission of the Connect-to-Learn Initiative is to provide education to children in impoverished areas (especially girls) via scholarships and broadband infrastructure. The economic impacts of such internet access are potentially enormous; increased connectivity and access to a wealth of information are essential components of rational decision making and internet access has consequently trumped numerous market failures. The focus of this paper is the unique information properties of internet access. It is assumed that information is received and used homogeneously across agents, bearing in mind that there are complementarities with income, education and institutions which may skew the extent to which different individuals can benefit from access.

The objective of this paper is to quantify the unique effect of increased access to information, outside of complementarities with income and education, on inequality. The central idea of which being that with reliable access to accurate information consumers can better make daily economic decisions. In a developing world context, the

[2] The representation of economic development as a multiplicative process which generates positive assortative matching in skills/productivity was popularized by Kremer (1993).

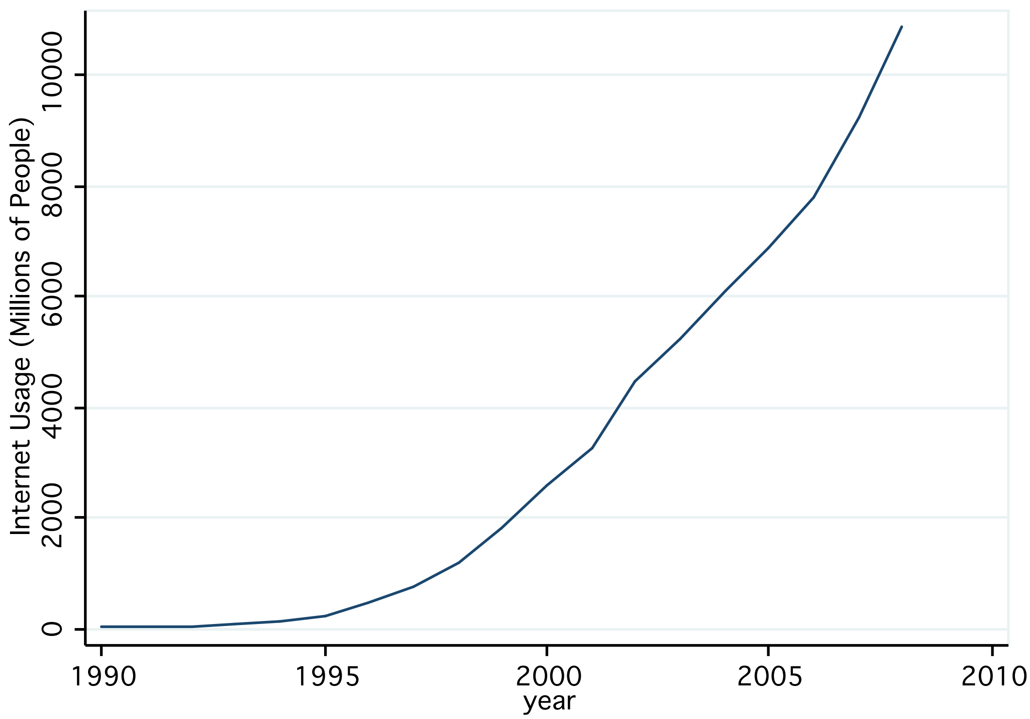

Figure 1: Internet Usage World Wide 1990-2008

integration of small businesses with world market data should increase productivity. As such, I propose that internet usage will impact inequality negatively.

This paper is organized as follows. Section 2 provides a review of the relevant literature concerning the digital divide, social mobility and inequality. Relevant growth studies are also noted. Section 3 details some simple two-period models which highlight the importance of complementarities with education and income Section 4 establishes the econometric framework in which the relationship of interest is assessed. Section 5 provides the results and testing. Section 6 contains discussion and Section 7 concludes. Further research is suggested in Section 8.

2 Background and Literature Review

This paper draws explicitly on two literatures. The first is related to growth, inequality and development; the second is the growing empirical literature on ICT adoption.

2.1 The Digital Divide

Inequality is an essential topic in modern development economics. A myriad of theoretical and empirical work has been conducted concerning the relationship between inequality and growth, living standards, human capital and social cohesion. Examples include Barro (2001), Odozi et al. (2010) and Haile et al. (2006), respectively. Although inequality is typically expressed in terms of income, the inequality that exists in ICT adoption (termed "digital divide") has several unique properties that are related to the characteristics of both technologies and receiving users. First, the unique properties of technologies influence the extent to which these technologies are adopted. Different consumers/firms have different needs in terms of the utility of a given technology. As such, the simple characterization that low-income countries will be less-digitalized is flawed. ICT adoption (for example) is a phenomenon characterized by several heterogeneous goods that serve many different purposes and as such, some technologies are feasible and fill a niche in countries where others are not and do not. Figure 2 presents two measures of technological distribution related to digital goods[3]. These measures are Gini coefficients[4] and coefficients of variation. It is clear from the figure that goods requiring significant levels of infrastructure such as secure internet servers (SIS) and international internet bandwidth (IIB) are distributed more inequitably than world GDP. On the other hand, accessible goods such as mobile phones are distributed far more equitably. A simple analysis of digitalization that does not account for the heterogeneity of technologies cannot adequately explain this. Figure 3 demonstrates that SIS are distributed similarly to GDP in a typical log-normal fashion, yet the sampling distribution of mobile phones is considerably

[3]The abbreviations used here are explained in section 3 of this paper

[4]The Gini coefficient used here is calculated on an international scale for specific digital goods

more equitable worldwide. Additionally, it is clear that internet is distributed somewhat equitably, with the exception of extremely low usage rates, which are common in many countries.

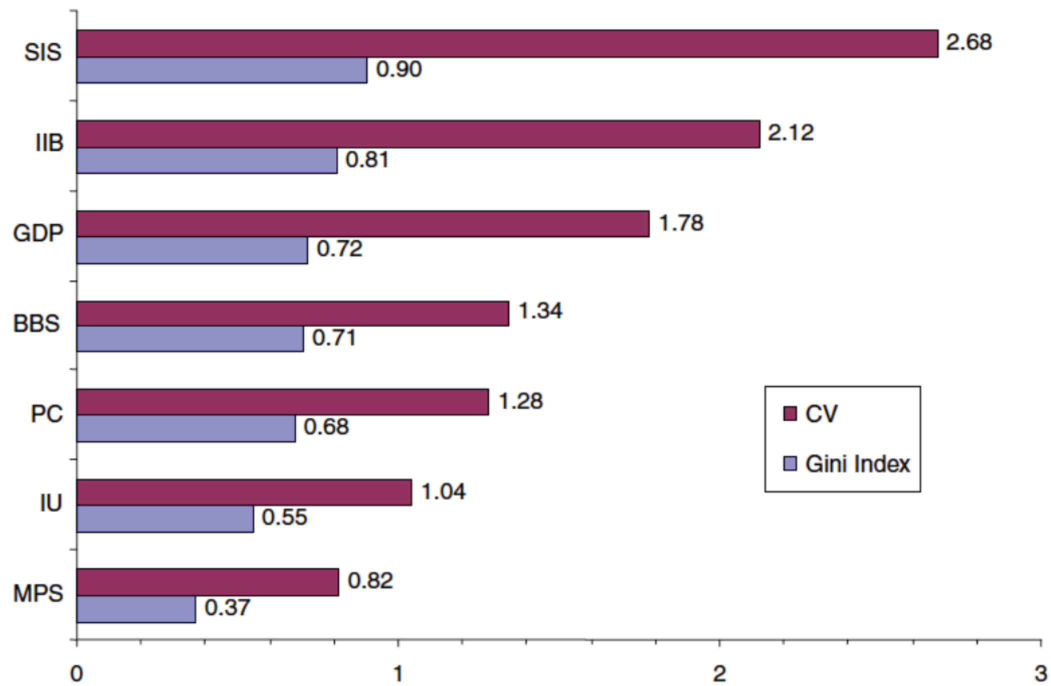

Figure 2: The Distribution of Digital Goods v. World Income Inequality (Source: Billon et al. 2010)

An additional characteristic of the digital divide that makes it unique compared to income inequality is the human capital requirements needed for the use of many technologies. The positive relationship between the rate of return to education and technological progress is first considered by Nelson and Phelps (1966). Though income inequality is clearly affected by the degree of education in a country, a majority of consumer goods do not require specific knowledge for their consumption. These skills are generally independent of conventional education and tend to develop alongside technological change. The internet is the ideal example of this. This complementarity creates a disincentive for diffusion at both the household and small firm level. As such, without coordination between physical goods and knowledge, it is difficult for households to adopt certain technologies. Without this preliminary, complementary knowledge, individuals are further constrained in their acquisition of human capital. This is especially detrimental to growth as it will only drive the incomes of rich and poor countries further apart as world commerce becomes more integrated and dependant on ICTs and e-commerce.

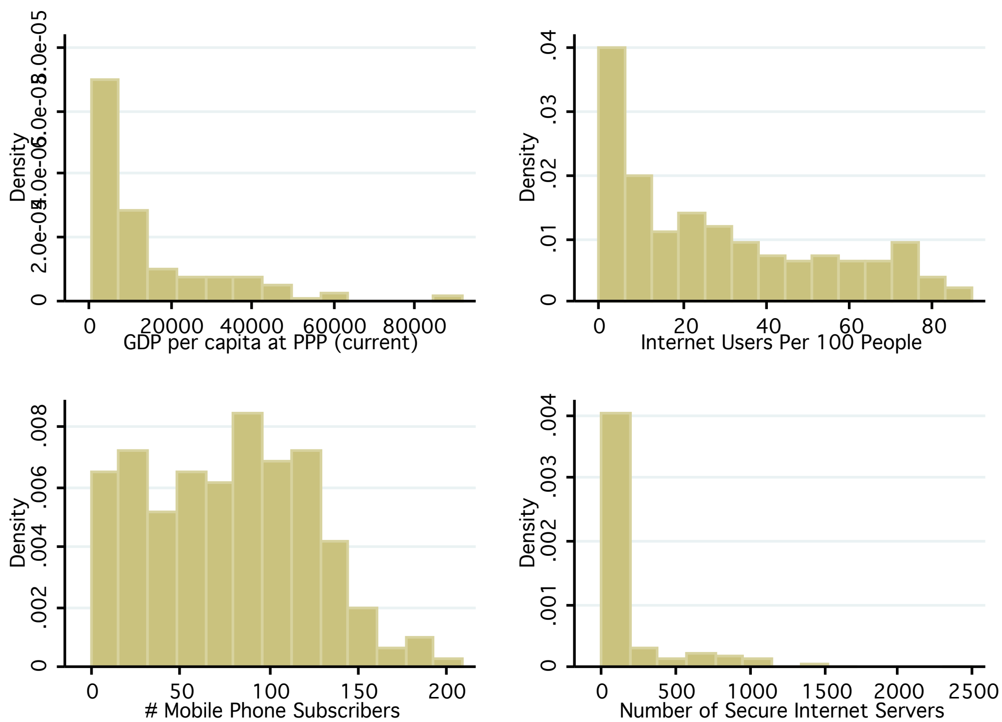

Figure 3: Sampling Distributions for Selected Technologies v. GDP

Given the persistent inequality in ICT adoption worldwide and the potential for this to worsen the already severe income inequality in the international community, the need arises to explain how ICT adoption is determined and how states of the world can be pushed from "bad" equilibria to "good" equilibria. Internet use having a positive effect on inequality would be an example of such a coordination failure. In the presence of significant income (and presumably education) inequality, increased access to information could further polarize incomes, due to complementarities. The "good" equilibrium in this circumstance would involve internet usage having a negative impact on inequality. As such, an efficient internet access policy would strive to reduce the complementarities between information and income/education.

2.2 Determinants of Digital Divide

2.2.1 Theory

As noted above, there is no dominant model of technological diffusion used by economists. While technological progress is generally captured in macroeconomics by models of endogeneous technological change, the micro-level household/firm adoption of diffusion is not consistently modelled using any single framework. This section summarizes ICT adoption models using diffusion theory, epidemic models, rank models and knowledge spillovers. A more comprehensive summary of theoretical models of technological diffusion is conducted by Baptista(1999). An important detail of this literature is that innovation is generally analyzed in the context of firms conducting R&D or firms adopting technologies into their production process. As such, these models are generally applied to the supply side of the economy. In the context of economic development, particularly in the case when production is somewhat informal, there is less of a distinction to be made between household adoption of technology and firm adoption. This section also provides a glance into evolutionary approaches to diffusion.

The workhorse of technological adoption is diffusion theory. Diffusion theory relies on the assumptions that a potential user is willing to adopt a technology once they are aware of its existence and that the technology is spread through direct contact with a current user. Diffusion theory posits that adoption over time will follow an s-shaped curve, akin to the logistic distribution. This approach allows for the identification of leaders in adoption and laggers of adoption, with the majority of eventual users adopting somewhere in between. The epidemic approach to diffusion theory was first popularized by Griliches (1988), who demonstrated that the diffusion of hybrid corn seeds in agriculture followed the logistic curve. Mansfield (1968) notes that the proportion of non-users adopting in a given period should increase, as the risk of adoption declines as an innovation is more widely received. As such, the rate of adoption is increasing in the number of users, as this increases the probability of contact between users and non-users.

Rank Models have been adopted to explain diffusion in the case of heterogeneous firms, proposing that this causes firms to adopt/innovate at different rates. Firm heterogeneity is taken as exogenous. In this framework, the demand function for a good is based on a vector of firm characteristics (Fusaro, 2009). Firm size, R&D expenditure and market structure data are typically included. Rank models are important in international comparisons of ICT adoption as the unique characteristics of countries are analogous to firm characteristics. Billon et al. (2009, 2010) use different characteristics such as population, education and institutional factors. Fusaro (2009) presents a technology adoption model for bounce protection programs[5], finding evidence supporting rank effects. The analytical framework used is based on specifying the incremental profit in time τ to firm i if the firm adopts in market j in period t, $g_{ij}^{t\tau} = g(R_i, F_j^t, F_j^\tau)$. R_i is a vector of characteristics specific to firm i which influence adoption (rank effect). F_j^t and F_j^τ refer to the number of current users of a technology at times t and τ respectively. Summation yields,

$$G_t^{ij} = \int_{\tau=t}^{\infty} e^{-(\tau-t)} g(R_i, F_j^t, F_j^\tau) \partial\tau$$

If the cost of adoption is P^t, then the arbitrage condition can be expressed as:

$$\max_t \quad G_t^{ij} - P^t$$

Knowledge spillovers have also been adopted by the diffusion literature. Billon et al. (2010) attempt to account for the heterogeneity of technologies by developing a digitalization index and using explanatory variables that are specific to components of the index. The extent to which a technology is adopted depends on the externalities

[5]Bounce protection is process/product whereby a bank pays overdrawn checks rather than returningbouncingthem

associated with it. For example, mobile phone use is easy to spread as it merely requires contact between users and non-users, given a minimal level of infrastructure. E-commerce is not as easily adopted by firms, as it requires SIS, which are specialized and costly to acquire.

There is no single paradigm with which to approach empirical work in digitalization, and many authors draw upon multiple literatures to construct econometric specifications. Theories span both adoption by firms and adoption by households, making diffusion generally adaptable to a macro-economic framework.

2.2.2 Empirics

Numerous empirical studies have attempted to explain the digital divide. There are generally two approaches in the literature. Some authors such as Lacovone and Crespi (2010) attempt to measure and quantify the gap that exists between countries in terms of ICT adoption. Others such as Billon et al. (2009, 2010), Chinn and Fairlie (2004) and Brown et al. (1979) seek to assess its determinants. The latter approach is prioritized in this paper.

Numerous indexes have been developed to quantify the overall level of digitalization in a country. Studies have been conducted using both panel data (Chinn and Fairlie, 2010; Chinn and Fairlie, 2007; Hultberg et al. 1998) as well as cross-sections (Billon et al. 2009, Billion et al. 2010). Samples range from groups of developing countries, groups of developed countries and combinations of both. Additionally, Hultberg et al. (1998) seek to assess what the world technological gap means for growth and convergence. The following section summarizes a comprehensive group of studies used to determine ICT diffusion across countries.

Billon et al. (2010) aim to identify and explain the differences in ICT adoption for a sample of 142 developed and developing countries. The main analysis of this paper concerns identifying determinants of overall digitalization, detecting unique patterns of digitalization across country groups and explaining these patterns. The crux of this analysis is statistical; theoretical models are briefly summarized but not explicitly drawn upon to develop an econometric model. The paradigm employed by these authors is to explain ICT diffusion by accounting for the nature of the technology, the characteristics of the receiving users and the means of transmission. This approach is consistent with a number of theoretical models, primarily those from economics. The authors' findings are consistent with the literature, primarily Chinn and Fairlie (2007, 2010). The authors also observe that the different coefficients across digitalization groups indicate that as country's economic circumstances vary, the technological preferences of consumers change as well. Though implemented using cross-sectional data, the results seem to imply that as a country undergoes economic development, consumer preferences change along with their economic circumstances.

Billion et al. (2009) study the determinants of ICT diffusion using multivariate analysis on a cross-section of countries. They use canonical correlation analysis to detect differences between groups of countries in terms of ICT patterns and their determining factors. Using data grouped by development levels the authors find that the major determinants of digitalization for each group correspond to development levels. Digitalization in countries with high levels of development is correlated with variables such as GDP, service sector, education and governmental effectiveness. For developing countries, age, internet costs and urban population are significant. The results could be used to assess ICT development policies.

Detecting different patterns of digitalization across countries is consistent with the market acting as a selection mechanism to allocate technologies across countries according to the level of development and needs of consumers. For example, government effectiveness having a positive impact on digitalization in developed countries is consistent with the types of specialized firms and institutions found in developed countries. Alternatively, urban population affecting digitalization positively is consistent with the minimal demand for technology in rural life in many developing countries. This latter scenario, however, could be the type of market failure mentioned above. Authors such as Aker and Mbiti (2010) note the effectiveness of mobile phones in lowering search and communication costs in rural areas. It is likely that costs, a lack of infrastructure or both may impede an efficient allocation of goods like mobile phones in rural areas.

An important study of the determinants of the digital divide is conducted by Chinn and Fairlie (2007). These authors focus on the determinants of PC and internet penetration by examining a panel of 161 countries from 1999-2001. As is the common approach, their explanatory variables use economic variables such as income per capita, years of schooling, illiteracy and trade openness, demographic variables related to dependency and urbanization, ICT infrastructure measures, ICT pricing variables and regulatory quality. For computer use, all variables are significant except telecom pricing and trade openness. A similar pattern holds true for internet usage, yet telephone density and age dependency are less significant. Both PC and internet use are mainly explained by income per capita, with regulatory quality and telephone density being the other primary factors. The authors additionally run individual regressions by region for East Asia and the Pacific, Europe and Central Asia, Latin America and the Caribbean, Middle East and North Africa, South Asia and Sub-Saharan Africa. The causes of region-specific disparities in PC and internet use are generally very similar. They advocate that public investment in human capital, telecommunications infrastructure and regulatory infrastructure could could mitigate the digital divide. The authors use a supply and demand framework to specify the model, estimating only the demand equation in a reduced form.

This work has been extended recently (Chinn and Fairlie 2010) to further evaluate differences in ICT adoption between developed and developing countries using a variant of an error-correction model using equilibrium values implied by current variable levels instead of lags. This model is applied to a similar panel as the one used previously. Findings reveal that the major factors determining ICT penetration in developing countries such as Brazil, China, Indonesia, India, Mexico, and Nigeria are income disparities, telephone density, legal quality, and human capital.

Dewan and Riggins (2005) examine both first and second order effects of the digital divide at three levels of analysis: individuals, organizations and the globe. The authors survey existing research noting the theoretical perspective taken in the work, the research methodology employed, and the key results that were obtained. The key theme found is that factors such as age income and education are successful predictors of ICT usage.

Hultberg et al. (1998) aim to address the impact of the worldwide technology gap on growth rates across countries. They develop a dynamic model which is an extension of the neoclassical growth model to include technological diffusion and institutional rigidities. They estimate panel data models using least squares with dummy variables, two-stage least squares and generalized method of moments for Europe, Latin America and East Asia. The authors instrument the current technology gap with its lagged value. The results indicate that the gap between a country and the leader nation is a significant determinant of growth, but that absorption capacity differs across countries. The authors use country-specific effects and absorption capacities to estimate the degree of efficiency across countries. Unsurprisingly, these estimates are significantly explained by measures of the corresponding country's institutional quality. This study is one of few that account for endogeneity problems by instrumenting independent variables.

Schleife (2010) attempts to explain the different determinants of ICT adoption in rural versus urban regions in Germany. The authors estimate a logit model using the probability of adopting home internet usage. Their findings indicate that there is a unique rural effect (negative) on internet access, independent of characteristics such as age, education and income. This effect diminishes as such variables are added. A key finding of this paper is the significance of the regional internet use rate as a regressor, indicating a network effect in which usage is more likely if an individual is surrounded by users.

The conclusions of these authors are incredibly important for assessing the digital divide. The importance of the absorptive capacity of countries in ICT adoption is noteworthy, as factors that influence absorptive capacity are included in many models that assess the digital divide. Factors such as education and other human capital components, as well as social and cultural characteristics, age and other demographic information constrain the demand for technology in a given country and how effective the technology can be in influencing growth.

Sandberg (2007) finds that the adoption patterns for email, internet connections, informative websites and interactive websites for Swedish public and private organizations all follow the logistic curve, with varying rates of adoption.

2.3 Inequality, Growth and Social Mobility

The other major literature this study falls under is that of growth and inequality. Numerous authors document the relationship between ICT, growth and inequality, and this is often applied to ICT.

Barro (2000) conducts an influential study on the impact of inequality on growth. The major finding of the paper is that inequality drives growth in rich countries but hinders it in poor countries. The author's results confirm the presence of Kuznet's inverted U-shaped relationship in cross-country panel data. However, factors other than GDP account for more of the variation in inequality.

Noh and Yoo (2008) note that the major economic outcomes associated with ICT diffusion are productivity increases and transaction cost decreases. Additionally, they insist that the relationship between inequality and growth is not perfectly understood, as it is unclear whether inequality is growth enhancing or growth-inhibiting. The authors analyze the impact on growth of internet usage when controlling for inequality using a pure-exchange overlapping generations model. Panel data estimation suggests that internet adoption affects growth adversely in the presence of high income inequality. This is attributed to the digital divide hindering growth associated with the internet.

Acemoglu (2002) attributes the recent increase in income inequality to the diffusion of information technology raising the wage premium for workers whose productivity is related to ICT. The author finds that for the majority of the twentieth century, technological change in the United States has been skill-biased, and this skill bias has been increasing. Traditionally, technological advances have been skills-complementing and hence raise productivity.

Lloyd-Ellis (1999) contends that the diffusion of ICT should increase worker productivity and will result in lower income inequality. The author develops a model of endogenous technological change and wage inequality, which suggests that wage inequality will rise when the rate of absorption is lower than the rate of introduction. This is attributed to competition for technologically-competent labour. This premium however raises the cost of innovation and thus slows adoption. This implies that in a circumstance where a population is largely devoid of technology-specific skills the rate of adoption will be much lower and income inequality will be high.

Martin and Robinson (2004) show internet diffusion has become polarized with income in the United States. Individuals with the highest family incomes had a significantly higher probability of accessing ICT than individuals with lower incomes. This result creates the need for further work investigating the relationship between ICT diffusion and income inequality, and serves as a benchmark study on the polarizing effects of incomes in ICT adoption.

Baddeley (2006) analyzes the effect of globalization on growth and inequality in less developed countries. A major finding of this paper is that the increased informational flows resulting from the spread of internet access has complemented economic outcomes, but less-developed countries have generally experienced increases in income inequality.

Andrews and Leigh (2009) analyze the relationship between inequality and social mobility using micro-data. Findings indicate that individuals from countries with higher inequality are less likely to experience social mobility. This supports the framework of this study which relies on an inverse relationship between inequality and social mobility.

3 A Simple Analytical Model

To illustrate the potential relationships between education, technology adoption and inequality, in this section I develop a simple formal example. Consider a two-period model with no saving. There are two types of individuals, those with higher initial incomes and higher levels of education and those with lower incomes and lower levels of education. Agents have initial income y and preferences $U(c_1, c_2) = c_1 + \beta c_2$ with $\beta \in [0, 1]$. The second period is analogous to working life and agents receive income based on their education $e \sim \{e_L, e_H\}$ and technology. Agents choose to invest in internet access in period one which complements education by θ in period two, yielding a higher income. It is assumed that $\theta > 1$. The cost of technology is x. An individual thus invests in internet access if:

$$U_I = y + \beta\theta e - x > U_{NI} = y + \beta e \tag{1}$$

which implies,

$$x < \beta e(\theta - 1) \tag{2}$$

Condition (2) implies that an individual adopts internet if the cost of access is less than the realized income gain of doing so. It is clear that if an individual has a higher level of education or the complementarity between internet access and education is high, then there is a greater chance an individual will invest in internet access. Condition (2) can be re-arranged to yield the minimum education level to justify investing in internet access.

$$e > \frac{x}{\beta(\theta - 1)} \tag{3}$$

Next, consider the case in which individuals with high education meet this threshold and invest while individuals with low education do not, that is:

$$e_L < \frac{x}{\beta(\theta - 1)} < e_H \tag{4}$$

This is the circumstance which characterizes a digital divide.

Proposition 1: The impact of a digital divide is the polarization of incomes, regardless of initial inequality.

Proof:

Because e_H-types adopt and e_L-types do not, their respective incomes over two period are $y + e_H\theta - x$ and $y + e_L$, respectively. It is clear that:

$$\text{(i) } y + e_L < y + \tfrac{x}{\beta(\theta-1)} = y + \tfrac{x}{\beta(\theta-1)}$$
$$\text{(ii) } y + e_H\theta - x > y + \tfrac{x\theta}{\beta(\theta-1)} - x = y + \tfrac{x\theta}{\beta(\theta-1)} - x$$

It follows that $y + \frac{x\theta}{\beta(\theta-1)} - x = y + \frac{x}{\beta(\theta-1)} = \lambda$. Combining (i) and (ii) yields:

$$y + e_L < \lambda < y + e_H\theta - x \tag{5}$$

Hence, incomes diverge with digital divide.

■

This result raises several policy questions. The first concerns the existence of an effective policy to alleviate this further polarization of incomes. The second relates to the nature of the policy, addressing whether or not intervention should be targeted and education-specific or whether the issue can be addressed via a public good such as infrastructure.

Proposition 2: Subsidization in public internet infrastructure that reduces consumer cost by s will decrease income inequality.

Proof:

Only $e = e_H$ types invest. That is:

$$\text{(iii) } y + \beta e_H\theta - x > y + \beta e_H$$
$$\text{(iv) } y + \beta e_L\theta - x < y + \beta e_L$$

This implies that s must satisfy

$$y + \beta e_L \theta - x + s = y + \beta e_L \quad (6)$$

As such, an effective s must compensate individuals with $e = e_L$ for the difference in income between adopting and not adopting. The difference in initial incomes is $e_H\theta - e_L - x$. After low-education individuals adopt, the difference in incomes becomes $\theta(e_H - e_L)$. The difference in inequality is thus[6]:

$$e_L(1 - \theta) + x < 0 \quad (7)$$

■

This framework highlights the potential for exogenous factors such as education and infrastructure to drive technology adoption.

Figure 4 demonstrates the impact of an increase in schooling for type e_L individuals. If a proportion ρ of the population are type e_H, then the average income can be expressed as $Y = \rho e_H \theta + (1 - \rho)e_L$. An exogenous increase in e_L will increase the average income, and reduce inequality accordingly. As such, the Lorenz curve will shift as indicated. As such, education variables are taken as exogenous in the empirical section of this paper.

An alternative framework involves comparable education levels amongst individuals, but different levels of initial wealth $w \sim \{w_L, w_H\}$. In the case of a borrowing constraint, the case when $w_L < x < w_H$ implies that incomes will diverge. If the cost of technology is reduced via an exogenous change in infrastructure such that $w_L > x$, then individuals with w_L will adopt. This will raise their incomes and reduce inequality. This highlights the complementarity between income and internet access in reducing inequality. The exogenous change in infrastructure that drives this increase in internet usage is the basis of the instrumental variable approach employed in the empirical section of this paper.

[6] It is clear that expression (7) holds as $\beta e_L(\theta - 1) < 0$.Taking the difference $\beta e_L(\theta - 1) - x - [e_L(1 - \theta) - x]$ yields $(\beta + 1)e_L(\theta - 1 > 0$. This implies that $0 > \beta e_L(\theta - 1) - x > e_L(1 - \theta) - x$.

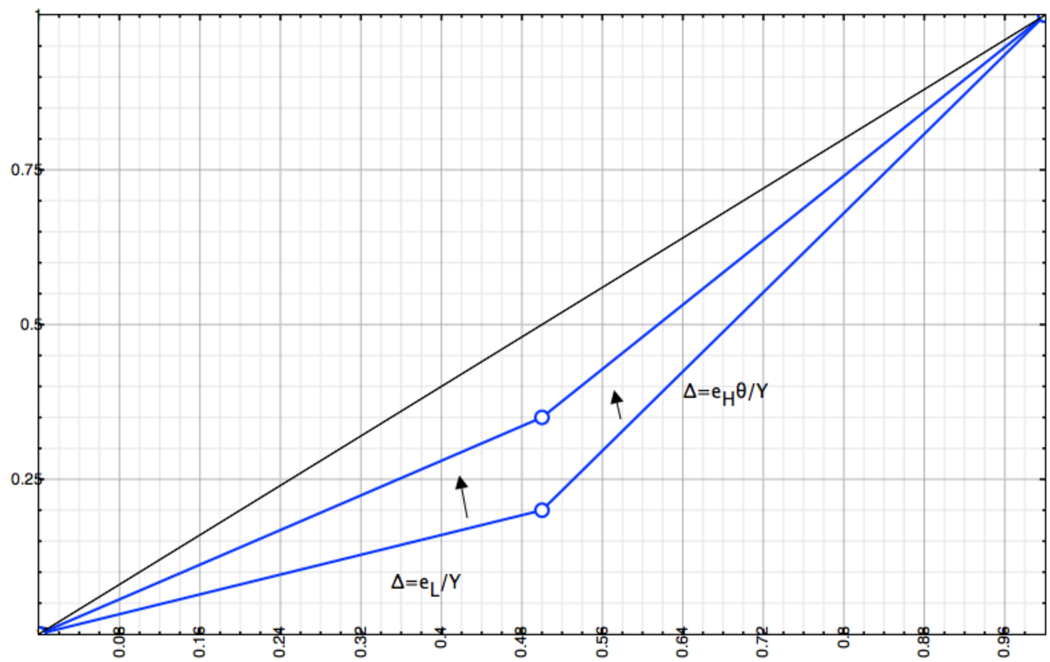

Figure 4: Effect on Gini Coefficient

Analytics suggest several hypotheses for empirical testing. The main prediction of this model is that an increase in usage that is caused by an exogenous improvement in infrastructure and education should be associated with lower inequality. Additionally, a complementary relationship should exist between internet access and education as well as internet access and income. The following section of this paper sets out an econometric framework which can be used to test these hypotheses.

4 Econometrics

4.1 Model

The objective of this study to to estimate the impact of internet access on inequality. I use two approaches that are common in the literature to estimate this relationship.

The first relies on a single equation approach in which inequality is specified (similarly to Barro, 2003) as:

$$
\begin{aligned}
INEQ_i =&\beta_0 + \beta_1 USERS_i + \beta_2 GDP_i + \beta_3 GDP^2 + \beta_4 ENROL_i \\
&+ \beta_5 USERS_i * GDP_i + \beta_6 USERS_i * ENROL_i + \beta_7 ROL_i + \epsilon_i
\end{aligned} \tag{8}
$$

where,

$USERS_i$ - Number of Regular Internet Users per 100 people in country i

GDP_i - ln(GDP) of country i

$ENROL_i$- Secondary school enrolment in country i

$INEQ_i$- Inequality in country i

ROL_i- Rule of Law Index for country i

Inserting internet usage as an additional explanatory variable in the standard Kuznets equation is the simplest possible approach. Because internet usage is highly correlated with other covariates such as GDP and education, estimating this model should yield the unique impact of internet access on inequality. This can be interpreted as the the contribution of information in the convergence or divergence of incomes. This framework isolates internet use as having an independent contribution outside of education or income in the determination of inequality. This particular relationship was first posited by Kuznets and has been widely tested, most notably by Barro (2000).

An alternative approach is to estimate the relationship between internet use and inequality as part of an overarching macroeconomic system. The rationale for this is that many theoretical macro models specify technology as being determined endogenously. There is no glaring reason that internet usage be an exception. As such, I estimate a system of three equations in which inequality and GDP are determined endogenously and both are affected by internet use. Additionally, internet usage is determined endogenously. Each of the three equations is derived from either theory or a specific literature within economics. Equation (10) is related to a standard production function. Equation (9) is the same as the sole inequality equation above. Equation (11) is specified similarly to technological adoption equations found in Billon et. al (2010) The model is specified as follows:

$$\begin{aligned} INEQ_i = &\beta_0 + \beta_1 USERS_i + \beta_2 GDP_i + \beta_3 GDP^2 + \beta_4 ENROL_i \\ &+ \beta_5 USERS_i * GDP_i + \beta_6 USERS_i * ENROL_i + \beta_7 ROL_i + \epsilon_i \end{aligned} \tag{9}$$

$$GDP_i = \alpha_0 + \alpha_1 USERS_i + \alpha_2 SLE_i + \alpha_3 GE_i + u_i \tag{10}$$

$$USERS_i = \gamma_0 + \gamma_1 GDP_i + \gamma_2 IIB_i + \gamma_3 SLE_i + \eta_i \tag{11}$$

where,

SLE_i- School Life Expectancy[7] for Country i

GE_i- Government Effectiveness Index for Country i

IIB_i- International Internet Bandwidth for Country i

Each of these models is subject to some caveats, which are mainly related to the potential for endogeneity bias. In both models, there is an unclear relationship between inequality and internet use. At low mean incomes, internet use is likely to have a negative effect on inequality, as demand for internet use is characterized by a threshold income at which an individual adopts the technology in a binary decision process. The impact of inequality on internet use is less certain at higher mean incomes. Notably, the work of Barro (2000) and Noh and Yoo (2008) supports a "rich get richer" hypothesis. Martin and Robinson extend this to analyze internet diffusion in this context. Conversely, it is clear that as internet use rises, individuals are exposed to new information, which aides in decisions concerning the labour market,

[7]School Life expectancy is the same as Average Years Schooling

personal finance healthcare and education. Each of these is a proven determinant of economic progress. This issue surfaces in both the single and multiple equation case.

Additionally, the sign of the coefficient on internet usage is uncertain a priori. In general, the information benefits of internet access could have a homogeneous effect across income levels, which could lead to a zero net change in inequality. If the returns to access to the poor are larger, then the sign of the internet use coefficient would be negative. This would imply that internet access largely benefits the poor. Conversely, a positive coefficient would suggest that the benefits of internet access are more pronounced at the upper end of the income spectrum.

4.2 Methodology

In the single equation case, the relationship described above must be estimated using instrumental variables. In choosing appropriate instruments, I exploit the fact that there is no theoretical reason that inequality is related to the current level of internet infrastructure and the analytical model above suggests the exogenous movement of infrastructure will affect usage. As such, the use of internet infrastructure or capacity is a reasonable instrument, as it is also strongly correlated with internet use. For a given year, internet infrastructure is predetermined, due to the lengthy process involved in construction. Using a standard two-stage least squares approach with infrastructure as an instrument will yield unbiased estimates.

In the case where the relationship of interest is estimated as part of a system of equations, the interdependence between GDP, inequality and internet usage is captured with $Cov(\epsilon_i, u_i, \eta_i) \neq 0$. This can be estimated using a seemingly unrelated regression. Equation (9) is subject to the same potential endogeneity bias as equation (8). There is additional potential for endogeneity bias in equation (10) as it is unclear whether internet use is a driver of growth or whether countries with higher incomes have more productive telecommunications industries, better infrastructure and higher personal incomes. Equation (11) has the same problem as it is unclear whether national income drives technological development or technological development drives growth. As such, this model is estimated using both seemingly unrelated regressions and three-stage least squares, to account for the interdependence of errors and potential endogeneity bias, respectively.

4.3 Data

Summary statistics are shown for key variables in Table 1. Data on GDP, internet use, school life expectancy and secondary school enrolment is from the World Bank Database. The gini coefficient series is comprised from this database as well as the United Nations University World Income Inequality Database (WIID). Rule of law and government effectiveness indexes are from the World Bank's Worldwide Governance (WGI) Indicators project. The chosen measure of internet infrastructure is international internet bandwidth, data on which was taken from the World Telecommunications Union database.

Table 1: Descriptive Statistics

Variable	Name	Obs	Mean	Std. Dev.	Min	Max
Gini	Gini Coefficient (2000-2005)	59	39.712	9.8826	23	60.1
GDP	log of GDP per Capita	59	8.784	1.2486	6.2295	10.890
Users	Internet Users per 100 People	58	12.251	14.780	.03626	47.888
Enroll	Secondary School Enrolment	42	69.978	23.989	8.65092	95.608
SLE	School Life Expectancy	56	12.518	3.8957	2.84	20.36
ROL	Rule of Law Index	59	.314	.97396	-1.07	1.9
GE	Government Effectiveness Index	59	.43847	1.0345	-1.16	2.08
IIB	International Internet Bandwidth	59	7854.1	21965.62	.13	111307.3

The limited number of observations on secondary school enrolment is restrictive in most regressions. Secondary school enrolment is not replaced by school life expectancy or an alternative education variable because it is the most significantly different measure of educational attainment across countries. In the case of inequality, secondary school is the most significant education level for labour market participation and a such, is the most relevant regressor in equation (1).

5 Results

5.1 Single-Equation Results

The results of the single equation regressions are shown in Table 2. Results are generally robust, but inconsistent with theory and the findings of other authors such as Barro (2000). Depending on the specification, income variables are insignificant and institutional variables are strongly significant. The coefficient on internet use is generally insignificant, though an inverse relationship is persistent.

Internet usage is statistically significant in models (1) and (5) only. The interaction term with education is significant in model (3) only. The same can be said for the interaction term with income. In this case, the coefficients are similar in sign and magnitude. The individual internet usage coefficients differ significantly across specifications, yielding mixed results in terms of direction. The addition of income and education variables yields significant and similar coefficients.. The coefficients on GDP are consistent with the Kuznets hypothesis and Barro (2000), with the exception of model (7). The addition of the institutional variable measuring the effectiveness of law yields a significant coefficient with a negative sign. This result is consistent with theory as institutional factors such as the enforcement of property rights should have a negative impact on inequality. The addition of this variable does not change the coefficient on internet use significantly and reduces the significance of GDP. Model (7) is the instrumented equation in which all variables are included. Surprisingly, there is little significance though the model tests well for overall significance. This indicates that multicollinearity may be a concern, which is consistent with the system approach taken below. The fits of these models are satisfactory, as the degree of inequality in a country is likely influenced by numerous country-specific factors such as culture and religion, on which data is not directly available.

Table 2: Single Equation Results

	OLS					2SLS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Users	-.355***	-.124	-2.456	-.623	-.421***	4.364	5.623
	(.085)	(.117)	(1.869)	(1.828)	(.073)	(6.969)	(6.912)
GDP		40.913***	70.242***	33.344		31.669	-14.359
		(12.876)	(17.800)	(20.413)		(39.233)	(53.718)
GDP^2		-2.503***	-4.027***	-1.598		-1.673	1.396
		(.748)	(1.064)	(1.253)		(2.337)	(3.339)
Enroll			-.257**	-.244*		-.299***	-.270**
			(.114)	(.127)		(.101)	(.116)
Enroll*Users			-.016*	-.010		-.018	-.008
			(.010)	(.009)		(.016)	(.013)
Users*GDP			.382*	.152		-.283	-.481
			(.228)	(.215)		(.594)	(.634)
ROL				-8.345**			-12.237**
				(3.456)			(5.602)
Obs.	42	42	42	42	42	42	42
R_a^2	.265	.334	.556	.609	.255	.415	.504
F statistic	17.328	8.266	13.97	14.894	33.427	13.37	17.561

5.2 System Results

The results from estimating the equation of interest as part of a system are shown in Table 3. Table 4 shows the other regressions in the model. In addition to greater explanatory power, the SUR results are consistent with the single equation results. While the sign and magnitudes of the income coefficients are reasonable, GDP does not appear as significant in the model, the the relationships of internet usage and its interaction terms are consistent with the single equation results. Only secondary school enrolment and the rule-of-law index are statistically significant. Chi-squared statistics show each equation is significant overall.

Table 3: Simultaneous Equation Results: Inequality

Gini	SUR	3SLS
GDP	27.502	102.491**
	(20.201)	(46.373)
GDP^2	-1.231	-5.837**
	(1.260)	(2.903)
Enroll	-.247***	-.288**
	(.070)	(.114)
Users	-.546	-4.695
	(1.759)	(7.675)
Enroll*Users	-.010	-.015
	(.008)	(.010)
Users*GDP	.149	.604
	(.193)	(.759)
ROL	-9.318***	-2.324
	(2.983)	(4.962)
Constant	-87.596	-381.936**
	(78.738)	(179.927)
Obs.	41	41
R_a^2	.677	.568
χ^2	91.842	75.627

The coefficients from the 3SLS regression are noticeably different in magnitude from those of the the SUR regression, though the direction of each relationship is maintained. In terms of significance, this model fares better that the SUR, with income variables being insignificant. The Kuznets hypothesis is validated and internet usage reduces inequality. The interaction between information and income is shown to increase inequality. Surprisingly, the institutional variable is insignificant, despite being strongly significant in the SUR regression.

Table 4: Simultaneous Equation Results: GDP and Internet Usage

GDP	SUR	3SLS
SLE	.135***	.135***
	(.033)	(.033)
Users	-.001	-.003
	(.009)	(.009)
GE	.750***	.767***
	(.144)	(.145)
Constant	6.794***	6.805***
	(.387)	(.387)
Users		
GDP	.497	.417
	(3.553)	(3.600)
IIB	3.277***	3.329***
	(1.260)	(1.274)
SLE	.723	.684
	(.765)	(.769)
Constant	35.032	36.818
	(41.368)	(41.899)
Obs.	41	41
R_1^2	.868	.868
χ_1^2	281.901	282.043
R_1^2	.659	.659
χ_2^2	85.535	84.754

Given the potential for multicollinearity and endogeneity bias, the 3SLS results are theoretically appealing. An alternative single equation approach is suggested in the following section

5.3 Robustness Check

The insignificance of most variables in single equation model (7) despite overall significance suggests multicollinearity is a concern. This occurs once the rule-of-law index is added to the model and as such, the rule-of-law index is suspect as the primary source of multicollinearity. This is a reasonable assumption, as institutional quality is somewhat fundamental to society and could be seen as a driver of incomes, education and technology. As such, an alternative to model (7) is to estimate the model using the rule-of-law index as an instrument for GDP. The results are shown in table 5.

Table 5: Alternative 2SLS Approach

Variable	**Coefficient**	**(Std. Err.)**
Users	6.029**	(1.928)
GDP	-49.925	(37.123)
GDP^2	3.060	(2.296)
Enroll	-0.283**	(0.079)
Enroll*Users	0.000	(0.002)
GDP*Users	-0.639**	(0.227)
Intercept	259.283†	(150.575)
N	42	
R^2	0.6	
$F_{(6,35)}$	8.763	

The estimates in table 5 were obtained by manually running the first-stage regressions and predicting GDP and internet usage. These predicted variables are then used to generate interaction terms. The second stage regression is then run[8].

The results of this alternative approach are generally consistent with the other single-

[8] This approach is taken because when the interaction terms for secondary school enrolment and education are generated, the standard approach will instrument the interaction term, despite the assumed exogeneity of secondary school enrolment

equation models, despite the income variables being estimated with the wrong sign. While secondary school enrolment shows statistical significance, its interaction term with internet usage does not. The interaction term for internet usage and GDP is significant, with the correct sign. Additionally, internet usage itself is significant, but with a positive sign. In the presence of interaction terms, this confirms the hypothesis from the analytical model which suggests that technology adoption can further polarize incomes when education levels vary or adoption costs are significant. Some of the insignificance in this regression is likely attributable to persistent multicollinearity amongst other regressors. As such, this strengthens the case that the system of equations approach is valid.

6 Discussion

The results of this study generally indicate that the internet usage on inequality is complicated and is dependant upon several complementarities. The single equation results indicate a significant impact only when internet usage is the sole regressor and this vanishes once both socioeconomic and institutional variables are added. This is probably because of the integration between technology, income and education. When additional regressors are added to account for this interdependence, internet usage becomes less significant. Interaction terms to capture the integration of technology, income and education are inconsistently significant. The interaction term for internet usage and income is significant more often than the interaction term for internet use and education. Income and education alone are strongly significant, with the coefficients on income supporting the Kuznets hypothesis. This result is similar to Barro (2000). The results differ, however, in that once the institutional variable measuring the effectiveness of law is added, income, variables become insignificant. Barro (2000) finds a weaker impact from institutional variables and a stronger impact from income. This can be accounted for by noting that income and institutional quality are related (Acemoglu et al., 2001). As such, the impact of income on inequality proposed by Kuznets (1955) could be primarily related to the development of institutions, the quality of which would likely increase alongside economic growth and technological change. This relationship has been explored in detail by Sokoloff and Engerman (2000). This could also be accounted for by the particular selection of countries in this sample.

While the unique impact of internet use on inequality is not generally significant in the above regressions, there is some evidence in favour of increased access to information having a polarizing effect on incomes found by Martin and Robinson (2004) and Acemaglu (2002). The inconsistent signs of the coefficients on internet use indicate that access to information can affect incomes homogeneously. That is, both the poor and the wealthy could potentially see income gains from internet access, ceteris paribus, and as such, internet usage appears to have no impact on inequality. In this case, there is no income/education advantage or disadvantage in benefiting from additional information, contrary to analytical predictions. As such, low-income individuals may not be proportionately disadvantaged in acquiring knowledge/skills via the internet because of income/education constraints. Similarly, there is no evidence to suggest that high-income individuals have an advantage related to higher income/education levels. In terms of the productivity debate, this suggests that the productivity of ICT and non-ICT workers are affected similarly, and as such, income-inequality does not change, which is consistent with the findings of both Acemaglu (2002) and Lloyd-Ellis (1999).

Additionally, if the insignificance of internet usage in the majority of the above regressions is related to the income gains related to worker productivity offsetting each other, then the significance of GDP and education variables and some interaction terms suggests that these factors do influence the incomes of the rich and poor unequally. This is the micro-analogue to findings of Baddeley (2006). For example, the marginal effect of education on the income of a poor individual is greater than that of a wealthy individual, while the information benefits of internet access provide similar knowledge and skills to both income classes. In the context of the digital divide, this result suggests that internet access should have a positive impact on living standards, and that the complementarities of education and income are independent of the informational component of internet access. As such, much-publicized campaigns to increase connectivity such as the Connect-to-Learn Initiative in the developing world accrue legitimacy as a means of alleviating poverty through reducing economic isolation.

7 Conclusions

In effect, the results of this study are inconclusive. The analyses conducted in this study indicate that internet usage does not have a significant impact on inequality outside of its relationship to education and income. This result is inconsistent with theory, however, and begs that further studies be conducted. The internet is perhaps the most important technology of the past two decades. It has become an essential force in the integration of national economies, changed the way agents access information, and has reshaped virtually every production process in the industrialized world. The effect of internet usage on growth has been documented. As such, it is surprising that internet use does not have a unique effect on other macroeconomic variables such as inequality. It is unlikely that the increased access to information that comes with internet usage does not aide in individual/household decision-making. With this in mind, this study should serve as a benchmark for future work in this area.

8 Further Research

There are several possibilities for future research on this topic. The principle flaw with this study in the use of inequality as a measure of social well-being. Instead, a similar analysis could be conducted that evaluates the impact of internet access on the incomes of high-income and low-income individuals separately. The insignificance of internet usage in this study is suspect and likely results from internet access having a homogeneous effect on personal incomes. As such, inequality will not necessarily change with increased internet access. Using various income shares as a set of dependant variables could also increase the sample size of such a study, as gini coefficients are scarcely available in abundance for any given year. Another potential measure of social well-being could be poverty rates.

An additional flaw of this research is that there is a tradeoff between sample size and choosing a relevant time period. The year 2000 is chosen as a benchmark for internet use, infrastructure, income and education because the follow decade has seen remarkable innovation concerning the internet. As such, it is assumed that this constitutes a critical period in which internet access can impact individuals. If data were available more consistently through the mid-1990s, a cross-sectional model in differences could be estimated. Additionally, much of empirical macroeconomics is conducted using panel data. If data were available (particularly infrastructure and internet use) further back, this problem could be addressed with a temporal dimension. This would be an important contribution, as technological adoption is an inter-temporal phenomenon. As such, the internet usage "demand" equation would be estimated as a logistic curve. As noted throughout this paper, the logistic curve is the standard specification for technological adoption.

A final extension of this paper would be to extend the multiple-equation framework used in section 3.1/4.2 to have a stronger grounding in theory. This could be done as previously suggested by specifying internet use as a logistic curve, as well as using a richer specification for an aggregate production function. Though GDP is estimated using standard variables from the production function, it is not specified as a neo-classical production function, for example. Re-developing the analytical framework to including both initial wealth and education differences together would provide a more comprehensive analysis of the complementarities between technology, growth and inequality. Additionally, the model could be extended to include saving, which could be used to endogenize the progression of technology and consequently, the degree to which technological progress interacts with education and income in determining inequality.

References

Acemoglu, D. 2002. "Technical change, inequality, and the labor market." Journal of Economic Literature, 40(1), 772.

Acemoglu, D and Simon Johnson and James A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." American Economic Review 91(5): 1369-1401.

Aker, Jenny C and Isaac M. Mbiti. 2010. "Mobile Phones and Economic Development in Africa" Journal of Economic Perspectives 24(3): 207-32

Anwar, Sajid and Lan Phi Nguyen. 2010. "Absorptive Capacity, Foreign Direct Investment-Linked Spillovers and Economic Growth in Vietnam" Asian Business and Management 9(4): 553-70.

Baptista, Rui. 1999. "The Diffusion of Process Innovations: A Selective Review" International Journal of the Economics of Business 6(1): 107-29

Barro, Robert J. 2001. "Inequality and Growth in a Panel of Countries" Journal of Economic Growth 5(1): 5-32

Billon Margarita, Fernando Lera-Lopez and Rocío Marco. 2010. "Differences in Digitalization Levels: A Multivariate Analysis studying the Global Digital Divide" Review of World Economics 146: 39-73.

Billon Margarita, Fernando Lera-Lopez and Rocío Marco. 2009. "Disparities in ICT Adoption: A Multidimensional Approach to Study the Cross-Country Digital Divide " Review of World Economics 146: 39-73.

Chinn, Menzie D. Chinn, and Robert W. Fairlie. 2010. "ICT Use in the Developing World: An Analysis of Differences in Computer and Internet Penetration" Review of International Economics, 18(1), 153167.

Chinn, Menzie D. and Robert W. Fairlie. 2009. "The Determinants of the Global Digital Divide: A Cross-Country Analysis of Computer and Internet Penetration"

Oxford Economic Papers 59(1): 16-44.

Dewan, S., and Riggins, F. J. 2005. "The Digital Divide: Current and Future Research Directions." Journal of the Association for Information Systems, 6(12), 153.

Fusaro, Marc Anthony. 2009. "The Rank, Stock, Order and Epidemic Effects of Technology Adoption: An Empirical Study of Bounce Protection Programs" Journal of Technology Transfer 34(1): 24-42

Griliches, Zvi. 2008. "Hybrid Corn: An Exploration in the Economics of Technological Change" The Economic Theory of Invention and Innovation: 355-76

Haile, Daniel, Abdolkarim Verbon Sadrieh, and AA Harrie. 2006. "Cross-Racial Envy and Underinvestment in South Africa" CESifo Working Paper Series: CESifo Working Paper No. 1657

Hultberg, Patrik T, M. Ishaq Nadiri and Robin C. Sickles. 1999. "An International Comparison of Technology Adoption and Efficiency: A Dynamic Panel Model." Annales Dconomie et de Statistique 55-56: 450-475

Lacovone, Leonardo and Gustavo A. Crespi. 2010. "Catching Up on the Technological Frontier: Micro-Level Evidence on Growth and Convergence" Industrial nd Corporate Change 19(6): 2073-2096.

Lloyd-Ellis, H. 1999. "Endogenous Technological Change and Wage inequality." American Economic Review, 89(1), 4777.

Krammer, Sorin M S. 2010 "International R&D Spillovers in Emerging Markets: The Impact of Trade and Foreign Direct Investment" Journal of International Trade and Economic Development 19(4): 591-623.

Kremer, Michael. 1993 "The O-Ring Theory of Economic Development "Quarterly Journal of Economics 108(3): 551-75.

Kristjansdottir, Helga. "Foreign Direct Investment: The Knowledge-Capital Model and a Small Country Case" Scottish Journal of Political Economy 57(5): 591-614.

Mansfield, E. 1968. Industrial Research and Technological Innovation. New York: Norton.

Martin, S. P., and Robinson, J. P. 2004. “The Income Digital Divide: An International Perspective.” IT and Society, 1(7), 120. OECD. (2001). Understanding the Digital Divide, Paris.

Murphy, Kevin M; Shleifer, Andrei; Vishny, Robert W. 1989. ”Income Distribution, Market Size, and Industrialization” Quarterly Journal of Economics 104(3): 537-64.

Nelson, Richard and Edmund Phelps. 1966. “Investment in Humans, Technological Diffusion and Economic Growth” American Economic Review: Papers and Proceedings 61: 69-75.

Odozi, John Chiwuzulum, Timothy Taiwo Awoyemi, and Bolarin Titus Omonona. 2010. “Household Poverty and Inequality: The Implication of Migrants’ Remittances in Nigeria” Journal of Economic Policy Reform 13(2): 191-99

Romer, P.M. 1990 Endogenous Technological Change, Journal of Political Economy 98: 71-102

Sandberg, Mikael. 2007. “The Evolution of IT Innovations in Swedish Organizations: A Darwinian Critique of ‘Lamarckian’ Institutional Economics” Journal of Evolutionary Economics 17: 1-23.

Schleife, Katrin. 2010. “What Really Matters: Regional Versus Individual Determinants of the Digital Divide in Germany.” Research Policy: 39: 173-185.

Schumpeter, J. 1934 The Theory of Economic Development, Cambridge, Mass: Harvard University Press

Sokoloff Kenneth L. and Stanley L. Engerman. 2000. “History Lessons: Institutions, Factor Endowments, and Paths of Development in the New World.” Journal of Economic Perspectives 14(3): 217-232.

Solow. Robert M. 1956. “A Contribution to the Theory of Economic Growth” The Quarterly Journal of Economics 70 (1): 65-94.