9-2007

Turning Metcalfe on his Head: The Multiple Costs of Network Exclusion

Rahul Tongia
Carnegie Mellon University, tongia@cmu.edu

Ernest J. Wilson III
University of California - Los Angeles

Published In
Proceedings of the 35th Telecommunications Policy Research Conference (TPRC 2007).

**TURNING METCALFE ON HIS HEAD:
THE MULTIPLE COSTS OF NETWORK EXCLUSION**

Rahul Tongia[1] & Ernest J. Wilson, III[2]
**Telecommunications Policy Research Conference (TPRC)
September 2007**

Abstract

Most studies of the digital divide and network connectivity begin by asking how many people are in a network, and what are the network effects associated. We flip the framing by using on the excluded instead of the included. We find, looking at Metcalfe's Law, Reed's Law, and other network "Laws" that all of them provide a techno-optimistic view of increasing network value and penetration, when, in reality, the excluded face increasing costs of exclusion. In fact, our simple model shows that the costs of exclusion rise faster than the growth of the network, and are approximately exponential. In addition to such a framing, we introduce ideas of exclusion across multiple networks, which could be considered multiple layers or dimensions of a broader connectivity graph. This initiates the groundwork for further theoretical and empirical analysis on network exclusion, combined with the policy implications of increasing costs of network exclusion.

[1] Senior Systems Scientist, Program on Computation and Society, School of Computer Science (ISR) / Dept. of Engineering & Public Policy, Carnegie Mellon University, Pittsburgh; Senior Fellow, Center for Study of Science, Technology, and Policy (CSTEP), Bangalore, India [tongia at cmu dot edu]
[2] Dean, Annenberg School of Communication, University of Southern California, Los Angeles [ernest dot Wilson at USC dot edu]

Introduction

The once-fierce debate over the definition, extent and consequences of the global digital divide is now all but extinguished. But from the late 1990s through 2002 the question of the 'digital divide' was the central topic of G-8 meetings (e.g., the meeting in Japan, 2000), the subject of World Bank conferences, and occupied the attentions of scholars everywhere. This culminated in the UN sponsored World Summits on the Information Society (WSIS). This focus is no longer the case, especially not from an OECD or "Western" perspective. Instead, the central tenants of more optimistic treatments of digital inclusion dominate scholarly, policy, and (sometimes) business discussions.

For example, Metcalfe's Law states that the value of a network is proportional to the square of the number of members of the network. As important as the formal statement of one of the information and communications technology (ICT) community's most oft-cited 'Laws' is the way the issue of network membership is framed – the benefits of network *inclusion*. This is perfectly appropriate and useful as far as it goes.

In this paper, by contrast, we wish to re-open and enlarge the debate to examine the considerable costs of network *exclusion.* In light of new ways to formulate the interactions of networks, choice and constraint, we believe the matter of inclusion and exclusion – especially of the digital divide – should be revisited by scholars and practitioners alike.

In addition to issues of central framing, we postulate that one difficulty with current network metrics is their focus on inclusion within one particular network, e.g., what is the teledensity in a population. In addition to standard issues of granularity and scale, such metrics ignore the issues of multi-modal access and exclusion, such as telephones vs. Internet, dial-up vs. broadband, etc. We conclude this paper with suggestions for further research along such dimensions, which capture not only issues of exclusion but multi-modal exclusion and access.

Digital Divide – The Decline of a Once-Fierce Debate

There are several reasons for the relative decline of attention to 'the digital divide.' One is the rapid rate at which new kinds of information technologies like the Internet and especially the mobile phone have in recent years diffused into non-user populations. In developed countries (OECD), for example, the higher penetration rates have shifted the Internet user base from a highly educated, mostly male and young demographic, to one that now includes the majority of the population of the G-8 countries (OECD, 2007), and encompasses substantial fractions of women, ethnic minorities (in the US and UK) and those without a high school education.

The empirical evidence seems at first blush to disconfirm the hypothesis that sharp information and communication technology (ICT) discontinuities between the rich and the poor were a permanent feature of modern, post industrial societies. The closing gap between Internet haves and have nots within economically advanced societies found its parallel in some evidence and strong arguments that computer use in some developing countries was accelerating so rapidly that there was hope that the North-South gap might be closing as well. For example, in China the Internet penetration grew from ~2% in 1999 to about 12% in 2007, and analysts expect China to become the single largest country in terms of total Internet users by 2009. In mobile phones, the highest growth is now in India, consistently adding over 6 million new subscribers per month (TRAI, 2007), and even Africa displays exceptional growth.

Furthermore, the near-simultaneous collapse of the computer industry and the telecommunications industry at the start of the decade drowned a great deal of the policy and analytic attention to the topic. To non-experts in government and among funding agencies the subject of the digital divide seemed to become less pressing than before and national funding agencies cut their digital divide budgets, and industry support for such research dried up. Often, the focus shifted from access and overall digital divide to specific uses of ICT in sectors such as healthcare, education, or national security.[3] In addition, cutting edge innovations in industry began to refocus more on wealthier markets or specialized products, while emerging regions were treated as expansion (but still important) markets displaying cutthroat competition.

Finally, additional scholarly work emerged concluding that the original concept of a ‘digital divide’ was much more complicated than originally described, and that the notion of one divide was misleading. Instead of a single one-dimensional divide, scholars found a multi-dimensional phenomenon in which access to hardware and applications played an important role in relative degrees of access, but so did access to financial resources, knowledge and to formal Internet training (Wilson, 2004; Tongia, Subrahmanian et al., 2005),

In light of these shifts, some of the earlier underlying techno-determinist arguments like Metcalfe’s Law began to re-assert themselves, almost implicitly. Much of the popular journalism, and some policy prescriptions, once again assumed that the divide would take care of itself over time.

Over the past several years in the academy and beyond the broader question of global equality and inequality has once again returned. Authors like Stiglitz (2002) point to evidence of growing inequality in the world, both within and between countries. Once-socialist economies like India and China that formerly forced down levels of inequality, have seen tremendous jumps as their economies have sprinted up the capitalist road. Other scholars point to inter-country inequality, as between Africa at one extreme and North America on the other. These claims have been met by counterclaims finding less inequality. The point is that economic asymmetries are again, appropriately, on the intellectual and policy agenda. And as before, it remains an empirical question to the extent to which these gaps are fed in part by ICTs. It has been hypothesized that globalization, which is dependent on the free flow of information and of capital, leads to a “winner takes all” phenomenon.

[3] An example of such a shift was taken by the Markle Foundation, who were pioneers in their focus on ICT and their participation in efforts such as the Global Digital Opportunities TaskForce.

In addition, the meaning of ‘access’ and ‘penetration’ is a constantly moving frontier because of technological innovation. In the early 1900s few believed having a telephone in the home was an essential public service. By 1950 household access to a telephone was considered a necessity for modern living in the US, and the definition of ‘access’ changed accordingly. Public service commissions for cities, states and countries promulgated new regulations to make this once-new technological device commonplace. Today, the meaning of an ‘essential service’ is still defined in part by the technological frontier, and more and more experts and consumer groups now insist that to be connected in the modern world consumers and citizens must have ‘access’ to broadband. Since that is hardly the case in many parts of the world, the matter of network inclusion and exclusion again comes to the fore in a kind of continuing fugue of technology and human ‘necessity’. Even the US laments the lack of broadband in rural areas (with a broadband ranking per household estimated at 25th in the world) (WebSiteOptimization.com, 2007), even through dial-up Internet access is available almost universally.

The purpose of this paper is to encourage more scholars to ‘go back to the future’ and once again give more attention to the distributive elements of the ICT revolution, We attempt to reframe the discussion about the ‘digital divide’ more in terms of network inclusion and exclusion, drawing from the work on Metcalfe’s Law, Reed’s Law and Odlyzko’s Law.

While there has been a welcome flowering of new empirical work in journals like *Information Technology and International Development* (MIT Press) on a wide variety of topics like differential access to telecenters and the multiple societal impacts of mobile telephony, these have not yet been accompanied by a conceptual reprise of where we stand on the subject of inclusion and exclusion. We offer a preliminary framework which we believe does advance the cause, and suggest in our conclusion some implications for scholars and practitioners. We conceive this TPRC paper as the first installment of several essays that will develop the argument for the Law of Digital Inclusion and Exclusion.

While not necessarily the first to point this out, Wilson (2004) made several observations regarding exclusion, including (1) this topic is understudied; (2) the costs of exclusion should be examined at not just an individual but also a societal level; (3) exclusion is a multi-faceted problem; and (4) policy interventions must be sensitive to such issues.

If, as Wilson argued, there are great risks that the disutility of a network will increase over time for the excluded, and that these costs bleed back from the individual into the community and society as a whole, then how can one formalize and extend this argument to make it more rigorous and attentive to other network attributes?

The first step we followed was to examine more carefully the extant claims made by other theorists, and to evaluate the extent to which they address our assumptions about inclusion and exclusion.

Increasing Value of a Network – "Laws" Galore

It is widely acknowledged that access to networks is the hallmark of post industrial society, whether defined as 'knowledge society', 'information society' or 'network society' (Castells, 1996). Extensive scholarly attention has gone to calculating the benefits and utilities of connectivity, combined with simple "Laws" to capture such network effects (Metcalfe's Law, Reed's Law, etc.). Much less attention is devoted to calculating the *disutilities* of network exclusion. We argue that more rigorous conceptualization and analysis are required to better understand the dynamics and implications of network exclusion.

In order to test and provide greater rigor to our assertions we first review relevant network "Laws" that specify the value of a network based on the number of people or nodes connected (Table 1). All display monotonically increasing value, with growth ranging from linear (slowest) to factoral (fastest), but Metcalfe's, Reed's, and Odlyzko's Laws are the most well-known.

Value *(proportional to)*	**Chronology**	**Originator**	**Model**	**Example**
n	1	Sarnoff	Broadcasting	TV
$n*\log(n)$	5	Odlyzko	A practical Metcalfe's Law	Telephone
n^2	2	Metcalfe	Networks	Telephone
n^c	6	Nivi	A practical Reed's Law	Google Groups
2^n	3	Reed	Communities	Google Groups
$n!$	4	Haque	??	??

Adapted from: "Between Metcalfe and Reed" (Nivi, 2005)

Table 1: Network Laws. These "Laws" are based on simplifications and assumptions, and the total network values approximate those shown in the first column.

Metcalfe's law has become synonymous with connectivity, stating that as more people join a network, they add to the value of the network non-linearly. The underlying mathematics for Metcalfe's law is based on pairwise connections (e.g., telephony).[4] If we have a 4 people with telephones, there could be a total of 3 + 2 + 1 = 6 links. The full math for Metcalfe's reasoning leads to the sum of all possible pairings between nodes, so the value of the network of size n is $\frac{(n)(n-1)}{2}$, which is simplified as being proportional to n^2.

[4] We recognize that Metcalfe's original formulation was for the critical mass crossover of device compatibility in a network (a non-linear growth), and not network value per se.

Reed's Law recognizes the value of groups within a network, not just pairs, so our group of four people could not only form pairs, but also groups of 3, or even the superset of all 4 persons. Adding in the 4 groups of 3, plus the entire group of four, all the sets equal $2^n - n - 1$; this approximates as being proportional to 2^n.

Odlyzko and colleagues (Briscoe, Odlyzko et al., 2006) pointed out that these network laws are likely too optimistic in their values, and one can intuitively recognize that the growth rate of the network value growth must decrease as subsequent members join – one would hope the most valuable links are formed first. This led to a formulation based on *n**Log(*n*), where future memberships have decreasing (but positive) growth in value.[5] We can see (Figure 1) that all the network laws show increasing values, but some are faster growth than others.

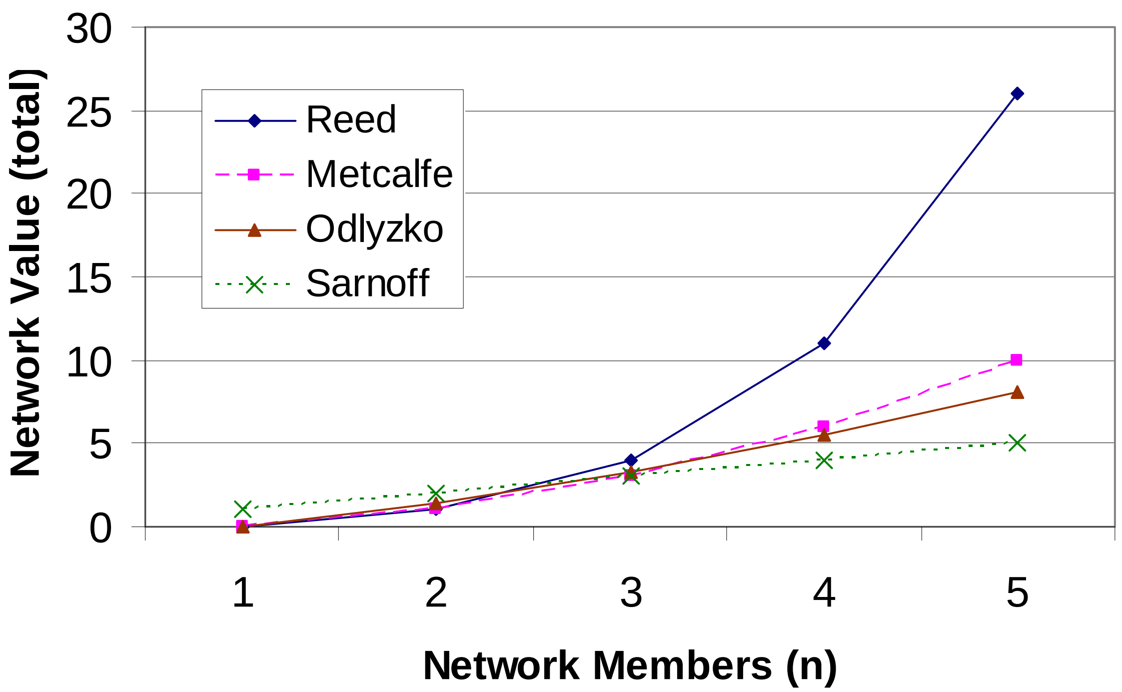

Figure 1: Network Values as per various "Laws". This shows growth from 1 to 5 members. We can see that Reed's law is the fastest growth, followed by Metcalfe's Law, then Odlyzko's, then Sarnoff's. Showing larger networks on the same scale becomes difficult due to the exponential growth of Reed's Law network values. It is important to note all Law formulations show increasing value.

It is beyond the scope of this paper to discuss the pros/cons or even appropriateness of such formulations, but *all* of the present network "Laws" attribute an increasing network value. It becomes self-evident that if we assume that there is a cost of not being in the network equivalent to the gap between the in-network value and those outside, then everyone outside the network faces a growing cost of exclusion.

[5] These would be natural logarithms (base e), but for convenience, these are written as Log(*n*) instead of Ln(*n*) or $Log_e(n)$

How do we value exclusion from a network? If we know the value of a network as per any law or formulation, and assuming each member is equal, we can calculate the value of inclusion per person. One might decide that the cost of exclusion is simply the difference between the outsiders' value (= 0) and the per person value of those included. Thus, for example, if Metcalfe's Law has a value approximating n^2, the per person value of inclusion is simply approaching $(n^2)/n = n$. Thus, exclusion would cost n based on the size of the network, as that is the difference between the values per person of those inside ($=n$) and those outside (=0).

However, what this fails to capture is that any network is of a finite size (if not in theory then in practice). If we state our network size is 19, Metcalfe's Law would indicate the value is proportional to 19*19 = 361, and the per person included value is ~19. Thus, the cost of exclusion for n=19 is also 19 (difference between 19 and zero, the value for those who are not in the network). However, we posit the cost of exclusion would depend on the number of people excluded as well. The previous formulation for exclusion indicates the same cost of exclusion regardless of whether the total population (applicable population universe) is 20 people or 200 people. All the above network laws assign a particular value to the network for a size of 19 in the network, but the exclusion costs are certainly different whether we have only one person excluded or 181!

To capture this effect, we formulate cost of exclusion differently. If value of inclusion per person is simply value of the network divided by size of the network (included persons), then the cost of exclusion is the value of the network distributed (divided) across the remaining population *not* in the network. Figures 2-5 show both the value per person in the network and the exclusion "cost" per person as the networks grow (total population size 20), in addition to the total network value. This is for Sarnoff's Law, Odlyzko's Law, Metcalfe's Law, and Reed's Law, respectively. Also shown is the trendline for a fit to an exponential curve for the exclusion costs, along with R^2 values for the respective curves.

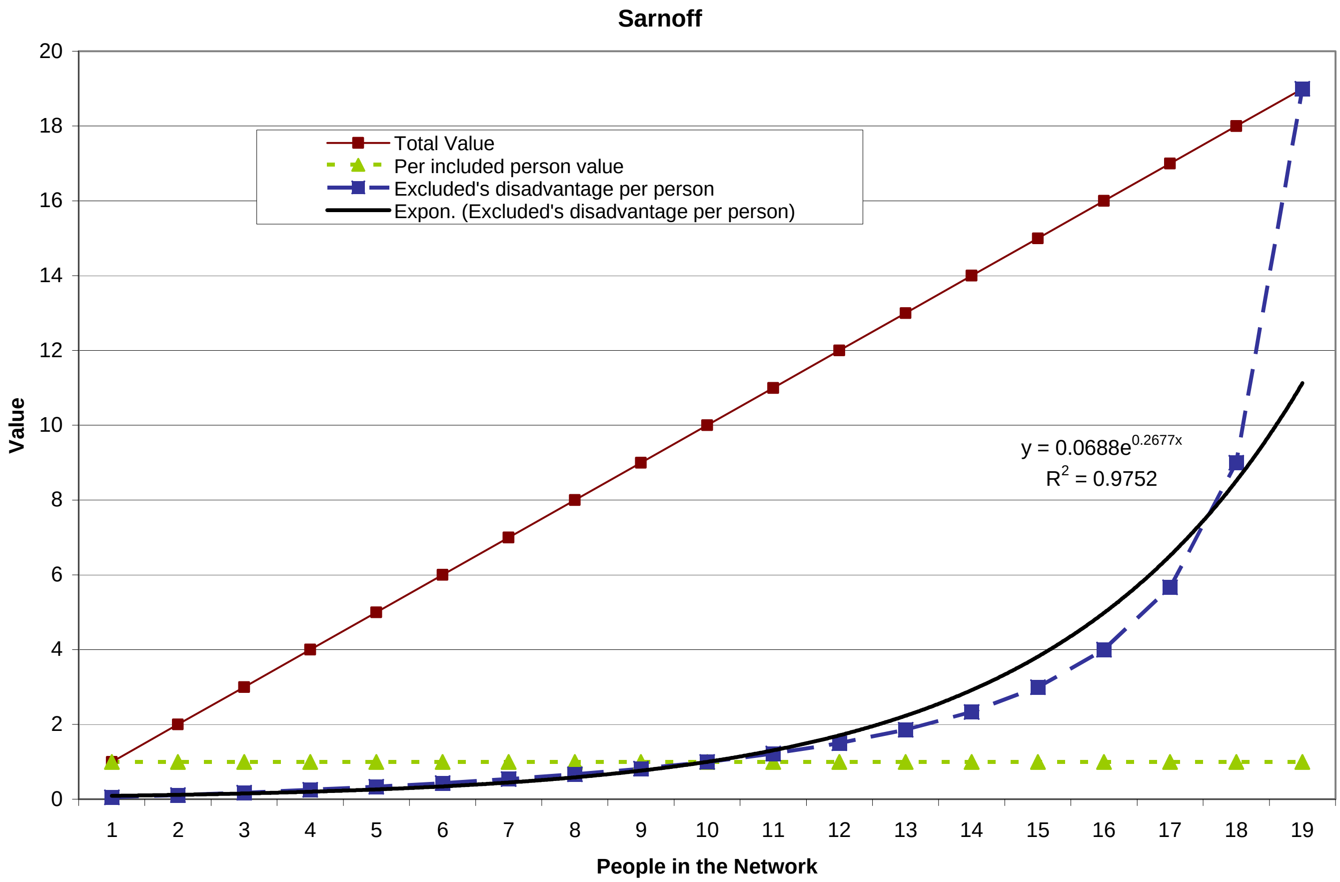

Figure 2: Sarnoff's Law and Network Values. The top line is the total network value, and we see it is linear. The inclusion value per person included (triangles) shows no growth in this linear network formulation, but the cost of exclusion (blue squares) shows rapid growth. In fact, the trendline for exclusion (solid black line) is approximately exponential.

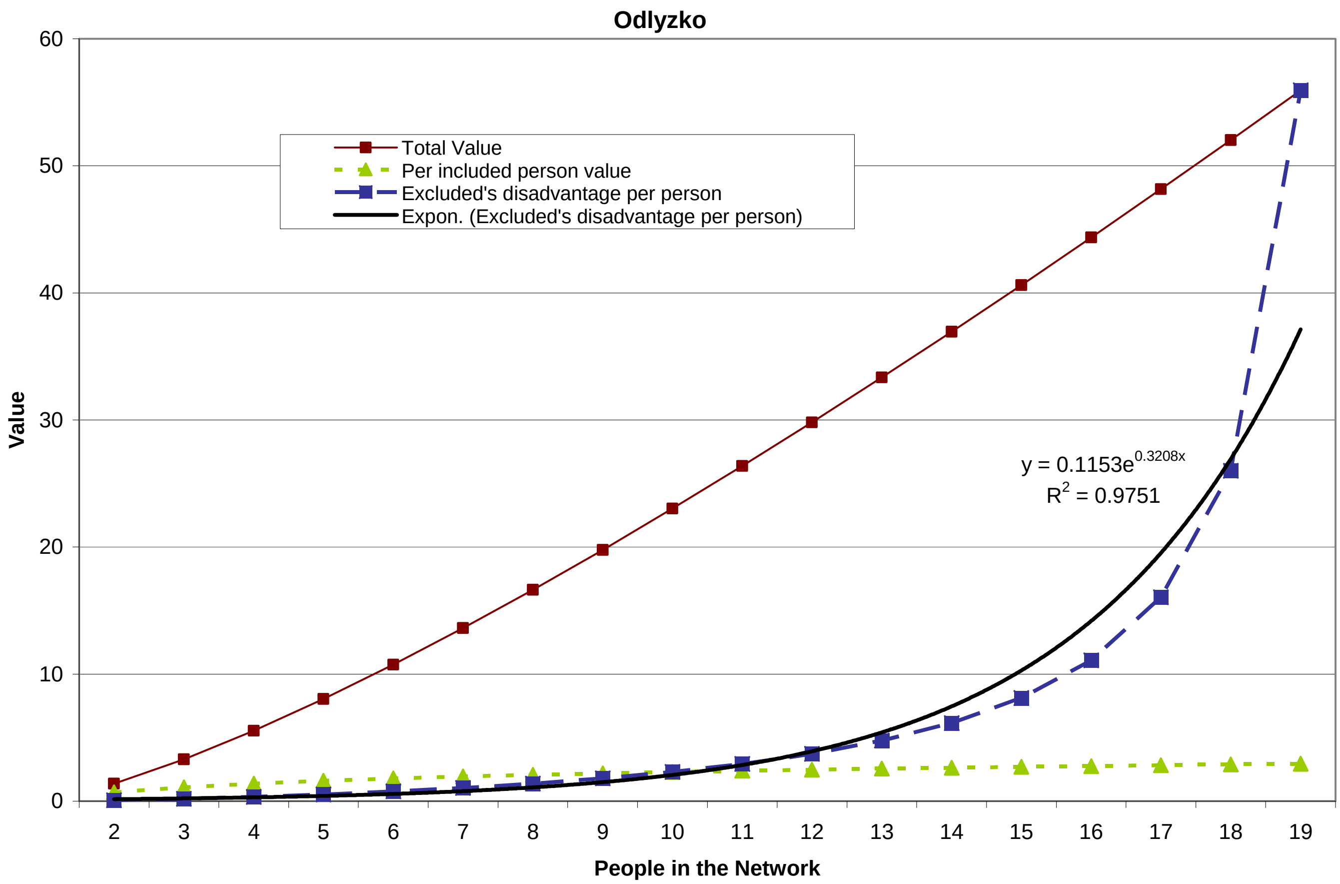

Figure 3: Odlyzko's Law and Network Values. The top line is the total network value, and we see it is slowly growing non-linearly. The inclusion value per person included (triangles) shows logarithmic (slow) growth in this network formulation, but the cost of exclusion (blue squares) shows rapid growth. In fact, the trendline for exclusion (solid black line) is approximately exponential.

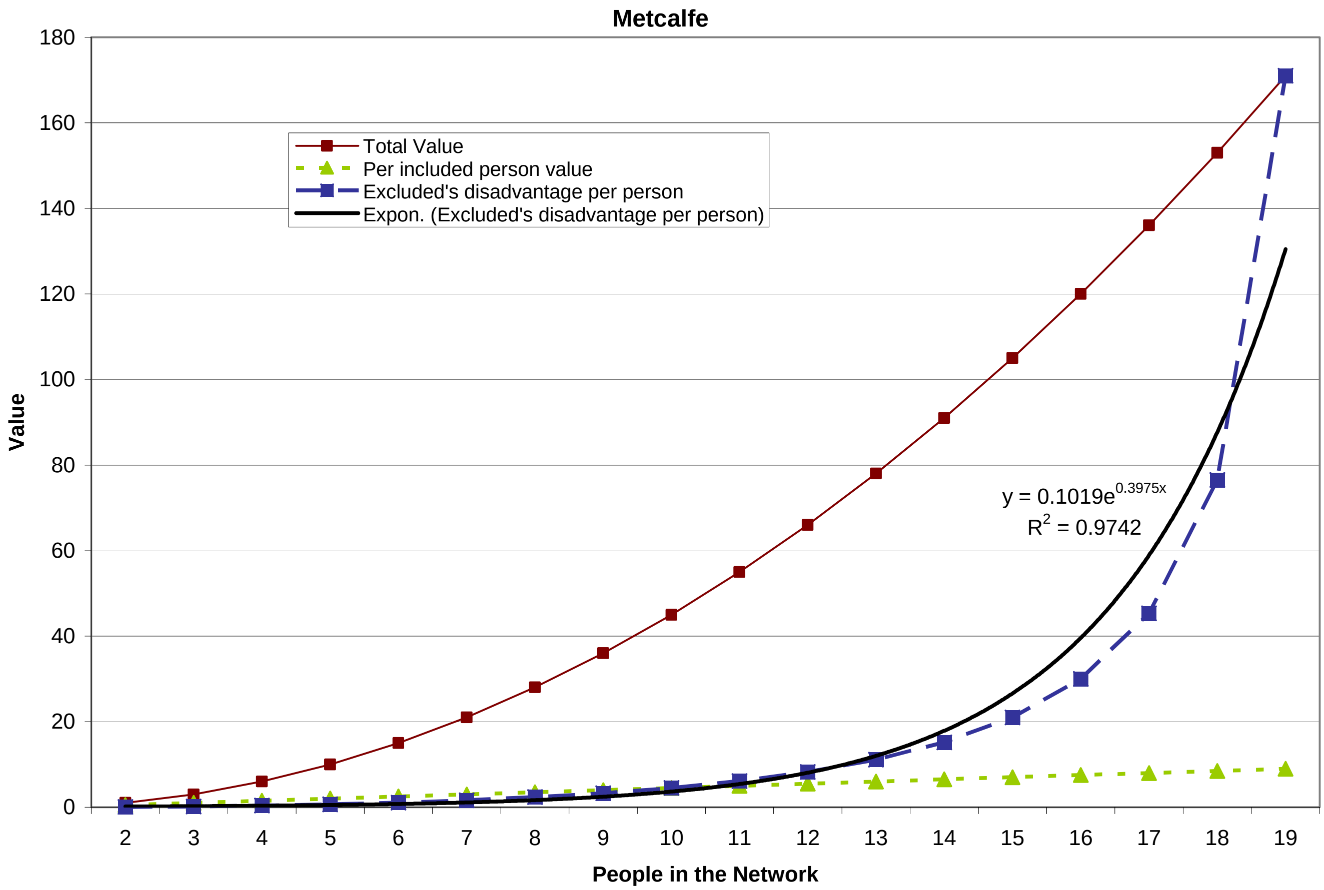

Figure 4: Metcalfe's Law and Network Values. The top line is the total network value, and we see it is growing strongly non-linearly. The inclusion value per person included (triangles) shows linear growth in this network formulation, but the cost of exclusion (blue squares) shows rapid growth. In fact, the trendline for exclusion (solid black line) is approximately exponential.

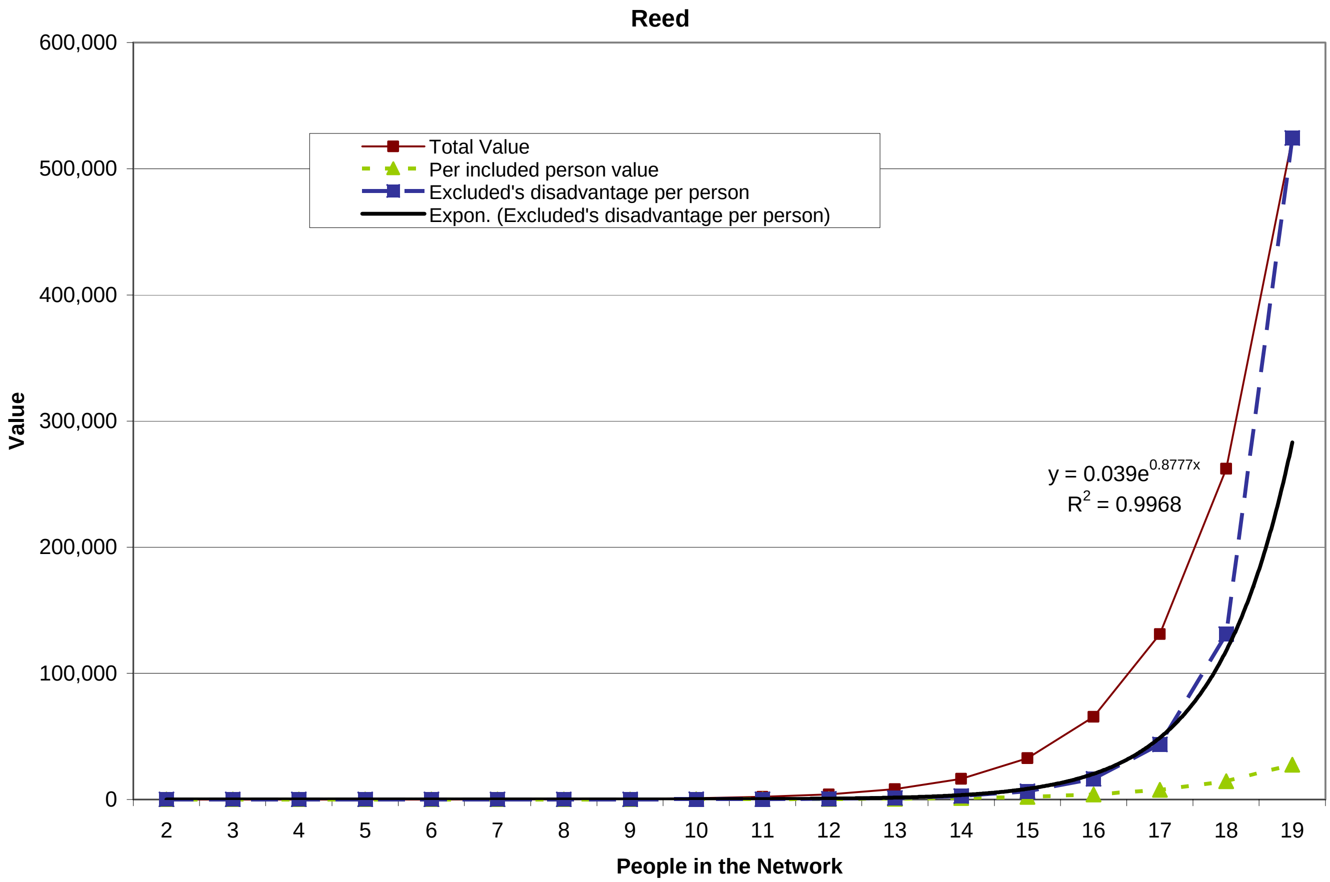

Figure 5: Reed's Law and Network Values. The top line is the total network value, and we see it is growing exponentially. The inclusion value per person included (triangles) shows exponential (rapid) growth in this network formulation, but the cost of exclusion (blue squares) shows even more rapid growth. In fact, the trendline for exclusion (solid black line) is very closely exponential.

We see several interesting and potentially important issues. For *all* the network "Laws" the cost of exclusion rises ~exponentially, and gets worse than exponential as only a few people are left out of the network.[6] . These findings are robust across all the Network Value formulations, evidenced by the very high R^2 for the fits to an exponential trendline. This surprising result has strong implications for how we frame the problem of exclusion, measure it, and attempt remedies.

In fact, if we consider a hypothetical constant value network, where there is a fixed value (say, 20) and this is shared amongst all the members who happen to join, even here (where the per person included finds declining values per person added to the network), the cost of exclusion rises roughly exponentially (Figure 5).

[6] It becomes asymptotic for n = [total population – 1].

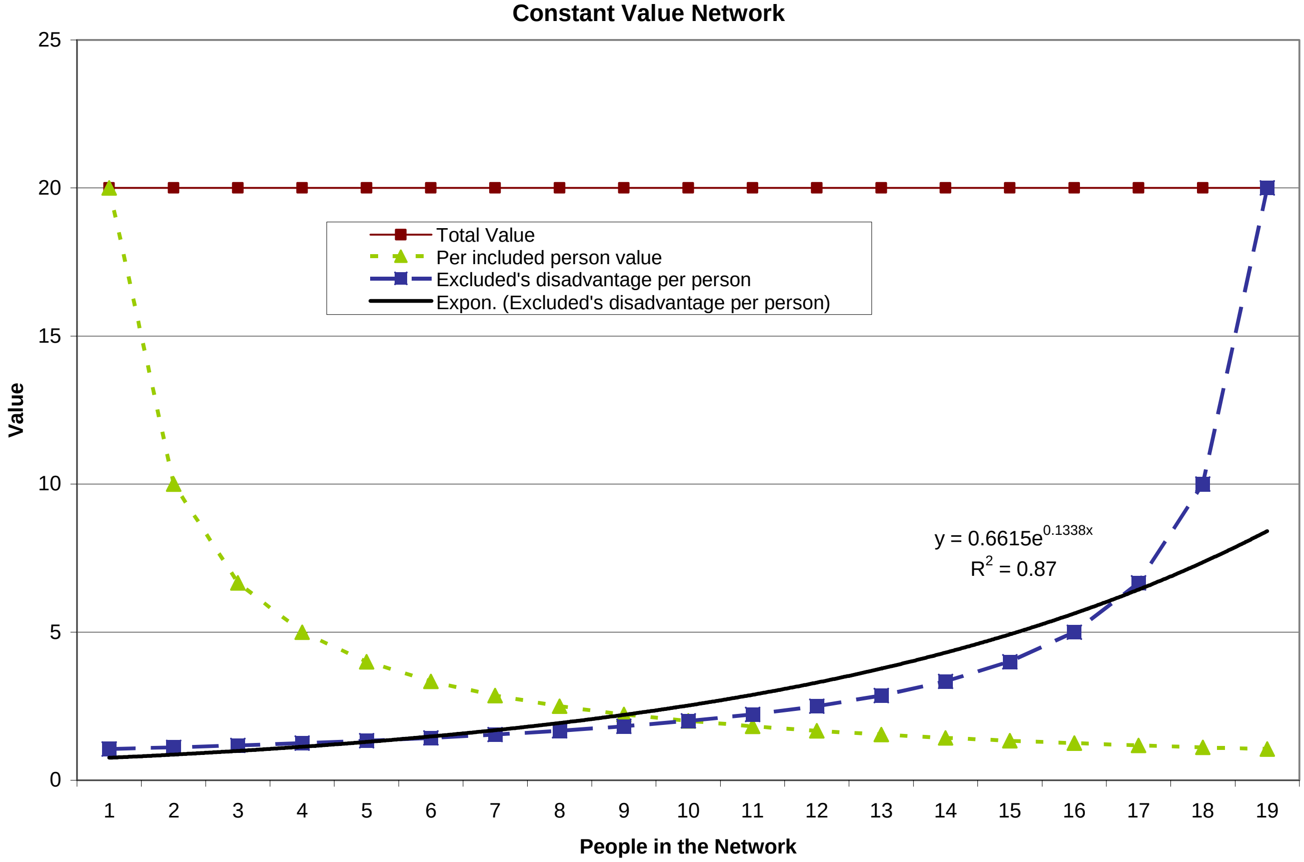

Figure 6: A hypothetical constant total value network. The top line is the total network value, and we see it is flat. The inclusion value per person included (triangles) shows declines in this network formulation, but the cost of exclusion (blue squares) shows non-linear growth. In fact, the trendline for exclusion (solid black line) is close to exponential, though with lower fit than previous Network Laws.

Framing costs of exclusion on the basis of included or excluded population is not merely an issue of semantics, or even standardization. An instance of such a case would be whether we are stating, e.g., A is 50% greater than B, or stating B is 33% less than A (they are both the same). In our framing, basing costs of exclusion on the number of people excluded, which inherently is a smaller base as the network grows, adds additional information that existing network "laws" fail to capture.

Existing Exclusion cost formulations = per person included value

$$\frac{[\textit{Network Value as per any Law}]}{\textit{Members in the Network } (= n)}$$

Proposed Exclusion Cost formulation = total network value divided by number of people excluded

$$\frac{[\textit{Network Value as per any Law}]}{\textit{Members outside the Network } (= N - n)}$$

(Where N = total applicable population)

In fact, if we compare the framing from included to excluded, the ratio of these is the same for any network law, and equal to $\frac{n}{N-n}$, where *n* is the people in the network, and N is the total population size. We can recognize that this ratio is growing, and inclusion and exclusion crossover (are equal) only at *n* = (0.5)(N).

We see from Figures 2-5 the choice of framing costs of exclusion based on traditional measures (the difference between per person in the network and outside, i.e., zero) versus based on distributing the value across those outside the network doesn't matter significantly for network sizes up to roughly half the population. *It's precisely when only a minority of the population is not in the network that the costs of exclusion rise dramatically*. There are many domains where ICT can offer advantages of options, efficiency, and empowerment.[7] Already, the majority of Americans seek health information online, and thus, we are approaching the levels of inclusion in many countries that exacerbate the costs of exclusion.

This is not to say that exclusion costs aren't high when the fraction of population included in a network is low. If we accept Odlyzko and Tilly's (2005) premise that the first few memberships of a network are the most valuable, then the relative advantage the first 10% have is the highest for any decile of the population subsequently joining the network. If we want, we could even frame costs of exclusion to be the higher of either inclusion-based or exclusion-based frameworks. As we've seen, inclusion-based costs are higher upto a point (Figures 2-5). This might be appropriate if we consider when only a few people are members of a network, the exclusion is spread out amongst the majority of the population but the *advantage* is held by only a few. Once a network includes the majority of the population, the *disadvantage* is held only by a few. Mathematically, this translates into saying for $n < 0.5N$, the included have an advantage they share, while for $n > 0.5N$, the excluded have a disadvantage they share. In such a formulation, the lowest disparity between frameworks is when $n = 0.5N$.

If we normalize to remove the network effects as per any "Law", we can easily compare the relative "costs" of exclusion (disparity) based on an inclusion-framing or an exclusion framework (Figure 7). This turns out to be the same, once normalized, for *any* network value Law or framework (constant, linear, lognormal, exponential, etc.) For our 20 person applicable universe (population), for up to 10 people in the network inclusion-based costs (i.e., dividing the total value by included people) shows higher costs, while after the midpoint, dividing by the excluded people shows the higher costs (exclusion-based framing). We posit that such a transition is warranted and analytically appropriate (discussed below). In reality, the costs of exclusion are relatively worse as fewer and fewer people remain outside the network given that all the Network Laws show increasing values to a network (which is hidden in the normalization below).

---

[7] It is widely accepted that those who do electronic commerce and other transactions online or using ICT face lower costs and find more options than those who do not.

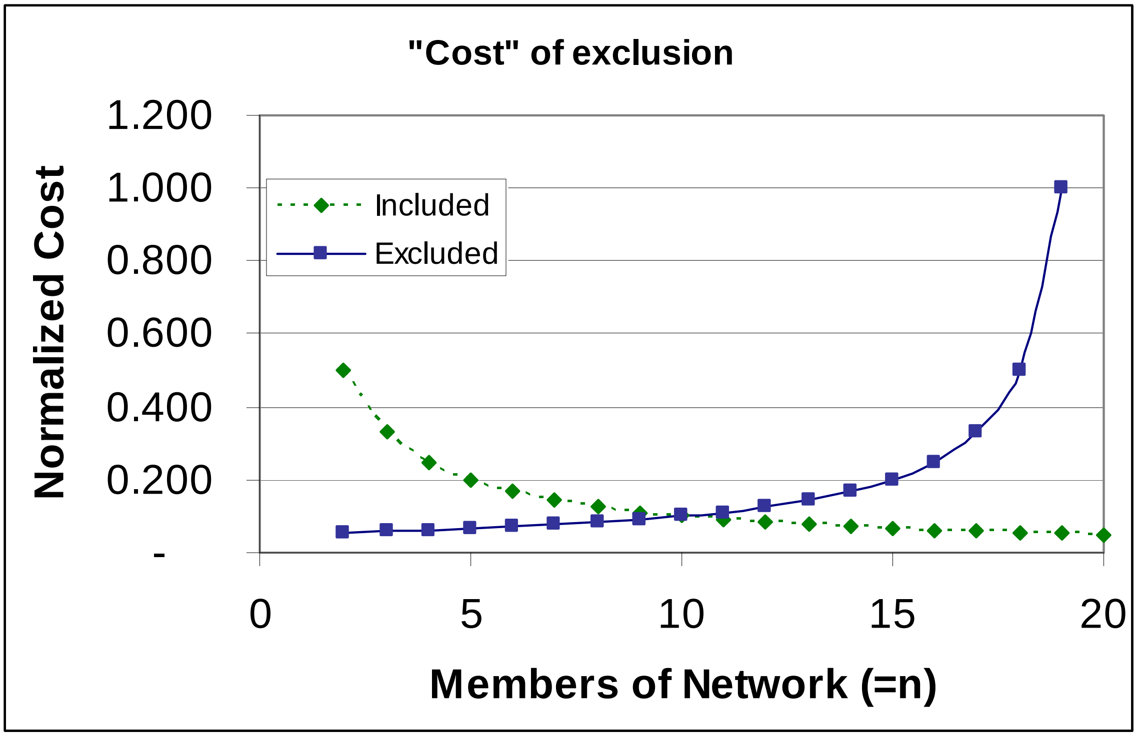

Figure 7: Normalized "Cost" of exclusion based on inclusion and exclusion framings. Inclusion framing divides network value by those in the network; exclusion based framing divides total network value by those excluded from the network (total population N = 20). The effect of normalization is to make inclusion framing look symmetric with exclusion framing, but, given the growth of network values for all the Network Laws, exclusion costs are worse than inclusion-framed costs per deviation from the midpoint (half the population is part of the network).

One might argue this is picking the worst of both worlds, and for consistency only one formulation should be chosen. However, we argue that as the fraction of a population in a network increases, there is a phase shift (occurring perhaps at 50% penetration) where the framing for "costs" of not being in the network should shift from inclusion to exclusion. When only a small fraction of the population is in the network, the median person in the population is excluded. Hence, inclusion is the exception, and not the norm. When the majority of the population is in the network, exclusion is the exception, and not the norm. Hence, it might be appropriate to use inclusion-based disparity costing in the initial growth of the network, and exclusion-based disparity costing as the network grows.

As a thought exercise, we could consider this to be equivalent to comparing two communities where in one, only 10% of the population has cars, and the other where 90% of the population has cars. The least disparity is not when almost everyone has a car (as the few left behind are really left behind) but when half the population has a car and half doesn't. As Figure 6 showed, it's only when the value of the network is constant (e.g., for car ownership) that 10% inclusion is as much disparity as 10% exclusion.[8] In all

[8] The same can be seen when we normalize for the network value (Figure 7).

network laws we have seen, with increasing network values with network size, 10% exclusion is much worse than the disparity caused by 10% inclusion.

Exclusion–Inclusion: Black & White, or Shades of Grey?

The present framings based on inclusion are not unusual, and we suggest a greater emphasis on exclusion is also important. However, this is inherently difficult. There has been an explosion of analysis in social science, networking, communications, biology, and many other disciplines focusing on studying networks (e.g., Monge and Contractor, 2003), including so-termed "small world" network effects. Watts and Strogatz (1998) reignited wide studies on issues of connectivity and networks across disciplines, and today there are entire Ph.D. programs studying such network issues, e.g., the Program in Computation, Organization, and Society (COS) at Carnegie Mellon. Scientists there have developed powerful tools for visualizing and analyzing networks with thousands of nodes, factoring in issues of incentives, dynamics, and even uncertainty. However, how are such tools meant to convey information about those who are not part of the network at hand?

Even one of the most celebrated network experiments has this shortcoming. Stanley Milgram's famous experiment in the late 1960s showing "small worlds" based on letters sent via unknown paths by random individuals in Omaha or Wichita to a pre-addressed person in Boston only through close acquaintances eventually led to the phrase "Six degrees of separation." However, the average of under 6 steps to reach the destination was only for those letters that made it, and a substantial fraction of letters never made it at all!

If we step back and attempt to study issues of inclusion and exclusion, we recognize limitations to the binary model of included vs. excluded. In economics, instead of merely looking at the average per capita GDP, many worry also about the distribution of the wealth.[9] This is often summarized as the Gini Coefficient, which varies between 0 and 1 depending on how equitably wealth (or any other variable) is distributed. We believe not only should network access be measured through similar metrics that capture distribution, there need to be new ideas in how to deal with granularity, differences in technologies, and the fact that, today, there's no easy way to have non-discretized markers of inclusion or exclusion (e.g., do you have a phone or not – it's hard to have half a phone, while incomes, shown in Gini Curves, can take on any continuous value).

In analyzing the digital divide and other networks, we find parallels between framing of network values as per inclusion with those of economic markets. An efficient market is meant to convey stakeholder choices through preferences and pricing signals, but what about those who do not or cannot participate in the market? Not only do they not enjoy the benefits of participation, but they are also left out of overall statistics. E.g., the black market economy is estimated to be a substantial fraction of the official economy in many

[9] A simple example is the failure to separate mean (average) from median in lay press, which might be simple confusion but, some believe, part of an orchestrated effort to mask the growing economic inequality in the US. [The mean will look better than the median because of the presence of a few super-rich, e.g., when we consider tax cuts].

developing regions. Even official energy consumption fails to capture in use of non-monetized energy, .i.e., when a household burns locally available biomass. E.g., the majority of rural Indian households use biomass for cooking. Adding such "non-market" uses of energy would roughly double India's total energy consumption compared to official figures based on commercial fuels. Why formal systems do not capture such uses is a matter of debate, but reasons include data limitations (by definition) and difficulties in conversion to a comparable form.

The same phenomenon is at play in connectivity and information access. If someone doesn't have internet access or have a telephone, are they cut off from information? More likely they would use alternate mechanisms, including borrowing connectivity or, worst case, going by foot or bus to gather the required information (e.g., government forms). Just like we may not know who is cutting down how many trees, we may not be able to properly capture all the alternative networks people may use for gathering information.

Instead of measuring connectedness within a domain, it may be more important to consider a multi-layer or multi-dimensional model of connectivity. Unfortunately, much of the data and methodologies on connectivity and digital divide work with a very binary framework of included vs. excluded. This fails to capture the shades of grey that occur even within a single mode of connectivity. As an example, the presence of public pay phones (including kiosks) dramatically changes the meaning of teledensity (telephones per 100 persons).

Is connectedness a discrete function? If I don't have a phone, am I treating landlines, mobiles, and VoIP as fungible? What if the latter doesn't allow emergency (e.g., 911) calling? What if I don't own the phone, but have access to one? I can use one from a co-worker or friend, or even pay for the service (kiosk or phone lady, ala GrameenPhone). If we allow the latter, within what radius or time (either walking or cycle) should it be available? These are just some of the important questions more sophisticated models need to address.

Parallel systems or networks have existed throughout business history. Often these are meant to be evolutionary but they can even be revolutionary. With the growth of newer technologies, it is increasingly common for a newer system or network to emerge and grow even before the previous system has reached its functional end of life. But comparing across systems becomes problematic when they are not perfectly substitutive in functionality. Consider the example of computer operating systems. While each iteration has an intended life of only a few years, there are still numerous Windows 98 users, not to mention the many more Windows 2000 users. At some point, the user or even the product provider (here, Microsoft) stops supporting the product with security and other updates, this becomes an increasingly parallel system, instead of one that co-exists though backwards compatibility. There are estimates that 80+% of spam comes via “zombie” computers,[10] and it is precisely the older computers that are most likely to unpatched.[11]

What does this have to do with network inclusion or exclusion? This is another example of layered networks, where people in the most current (or dominant) network (e.g., Windows XP or Windows Vista) pay a cost for those outside that particular network, and such costs are often hidden or hard to account for. As Lloyd Benston, then Treasury Secretary observed during a Press Briefing on July 20th, 1994 regarding health insurance inclusion and exclusion:

> *If you have insurance, it's easy to say, well, you know, this doesn't affect me; the uninsured -- that's their problem, not my problem. Don't you believe that; that's not right. It's your problem, too, because insurance costs are then higher. You've got a bed that isn't paid for in that hospital, you've got a doctor that's not paid, you've got a nurse that's not paid -- those fees go up. The hospital cost per bed goes up for those of us that have the insurance. Or if you have a public-owned hospital, like a city hospital, and the bills aren't paid, your taxes go up.*

There is another implication from a “supplier” perspective. At what point is it not worth supporting the older network or system? This may not be driven by marginal cost – marginal benefit economics but also by things like marketing burdens or regulatory requirements. It costs money to inform consumers that they need to shift, as well as provide them the tools (sometimes involving hardware) to shift over. US Broadcasters estimate a cost of over $100 million of advertising just to inform US TV users about the shift to digital TV in 2009, and the corresponding end of analog broadcast television (Steinberg, 2007). There were similar (but order of magnitude larger) costs when London changed its area codes. This highlights how inertia is a major reason parallel networks or systems co-exist.

Even Apple (having dropped the Computing from their name) made network exclusion part of the reason for their shift to Intel-based processors, with power requirements being another. Given the dominance of Windows systems for business and gaming use, an Intel Macintosh could run Windows programs through partitioned booting or through virtualization. Thus, Mac users were able to join the Windows networks.

Most examples that come to mind for parallel networks involve ICT, in part because of the fast pace of change. The majority of telephone instruments are touch tone dialing, but the back end systems in the US all allow pulse dialing, for the very small fraction of users who still have a pulse instrument. This must certainly be an expensive equilibrium.

There are several other examples in other domains where this phenomenon has been seen. In Calcutta, India, till a few years back, a small segment of consumers received DC

[10] Zombies are machines that have been taken over by malicious entities without the knowledge of the owner for nefarious purposes such as sending spam (or worse, such as Trojan or Virus attacks). Computers become zombies usually through a vulnerability in the operating system or a particular application.

[11] Windows Service Pack 2 for XP (SP2) qualifies as a new version from a security point of view, but there were many XP users (note, we didn’t use the term customers!) who didn’t or weren’t able to upgrade.

(direct current) household electric supply, despite the parallel availability of regular AC (alternating current) supply. This was due to legacy reasons, but the entire system and all the uses had to bear the costs of the parallel network for decades.

Implications

One of the implications of properly valuing exclusion, and from figures 2-5, is that exclusion becomes particularly worse when only a subset of the population is excluded. While marketers and innovators may focus on early adopters and expanding markets, policy-makers must worry about society overall. This is one reason a number of basic services were regulated by government, and utilities were often under "must serve" obligations, e.g., electricity or even landlines. The same has not been the case for many other ICT systems, such as mobiles or Internet access (let alone broadband). Presently, if one doesn't have good mobile coverage, in most countries one has little recourse with the carriers.

Certainly economic efficiency arguments would posit that there can be a subset of the population that is too expensive to serve, and forcing service for them raises the costs to everyone, often through cross-subsidies. E.g., residential consumers mostly pay a flat rate for their electricity per kilowatt-hour, regardless of whether they live in a city or a rural shack. The other option is price discrimination, but this has usually been frowned upon by regulators. The fundamental question becomes, as ICT grows in importance, is connectivity (through not just landlines but mobiles, Internet, broadband, etc.) an essential public service, deserving of universal service obligations? If so, how are these to be funded?

Are parallel networks always bad? Certainly we don't advocate hegemonies of "better" networks or systems. Older networks or systems are likely to merge through specifications or simply through market attrition (Cowan, 1991), or there may be inexpensive solutions allowing multiple networks or systems to operate concurrently, e.g., many WiFi devices can operate as 802.11 b/a/g (3 standards concurrently designed into in the device, and configurable through software or automatically).

In spite of the value of multiple networks, there may be cases where the societal costs of running parallel networks justify intervention. This could be compensating people for upgrading their end-user devices (e.g., giving away touch tone phones to pulse phone owners). Related to this is the option of creating a gateway or converter system that can translate between the networks, and one might deem it important to give away such a device as well. As part of the US transition to digital television broadcast, the US Congress has authorized 2 $40 subsidy coupons per household to buy converter boxes to convert the over-the-air digital broadcast into analog signals – this would allow families to continue using their current (analog) televisions.[12] The costs of the coupons are calculated to be lower than the value of the spectrum that would be released through the transition from analog to digital.

[12] At 109 million households, if everyone took advantage of the voucher, this would cost over $8 billion. In reality, there is a cap on the total money available, upto the maximum of $1.5 billion.

One subtle implication of thinking of exclusion as a multi-layered phenomenon is that we might find, with new analysis, that too much emphasis has been placed on intra-modal connectivity instead of inter-modal connectivity. In the Internet world, there are many studies (e.g., Albert, Jeong et al., 1999) that talk about the diameter of the World Wide Web. However, once online, as per the end-to-end design, almost all websites should be reachable, regardless of whether it takes 8, 12, or 16 hops. Admittedly, there are performance implications, but the greater challenges are those of application restrictions (e.g., disallowing VoIP like Skype, which creates gaps between the Internet and Telephony networks) or even the economics of long-distance links (such as international gateways). On the other hand, a large fraction of the world isn't online, but might own or have access to a mobile phone. To what extent is information available to them through such a medium? By this framing, a node is not valuable only because of how much it connects to others in the layer but rather a node is more important if it translates between layers. We believe that many underserved regions are focusing much of their Internet efforts on speeding up their connections (which is certainly important) without enough effort on worrying about those who are excluded. The latter could be benefited by not only their joining the network, but by making information designed for one (e.g., Internet) network and making it available to other networks such as voice telephone, SMS/text message, etc.

This framework for network exclusion, with layers and nodes, could find applicability in a number of real-world scenarios, especially where there is likely to be overlap and issues of compatibility between existing and "replacement" technology. E.g., the transition to the Next Generation Internet (IPv6) has been slow, in part because of issues of transitioning. However, if there were to be a more secure or faster net using IPv6, let alone one that enabled new services or applications, would this not be another "digital divide"?

Similarly, is it still a digital divide when a home only has one (expensive, restrictive) broadband provider, when most others in the same region have multiple providers offering more services, allowing more applications, and lower prices? Not all services can be offered everywhere at the same price – unless one treats the services like a regulated service, or there is simply cross-subsidy between users.[13]

This framing of networks as multiple layers might have applications in other scientific, industrial, and social phenomenon. Of course, in many other networks, such as the power grid, the linkages within a network matter significantly. Parallel networks such as the transportation network or natural gas network might interact, but are not as easily substitutable. However, the information network (across email, web, phones, fax, etc.) is not only relatively more fungible, it is much more easily reproduced and designed for additional capacity and capabilities. After all, it's virtually free to send an extra bit when

[13] This excludes a net subsidy as an option.

one gets lost (unlike electricity); this is how routers operate, re-transmitting packets when some are dropped, sometimes through alternative routes.[14]

Throughout our discussion of parallel networks thus far, we have discussed parallel networks mostly as physical or architectural entities, e.g., IPv4 vs. IPv6, or phones vs. the Internet vs. going to a location in person. Our framework actually posits non-physical aspects of exclusion, which capture factors like literacy, language, cognitive capabilities, etc. These could be modeled either as parallel networks themselves (e.g., the group or cluster of Spanish speaking individuals within the US), or as descriptors for individual nodes based on a graph theory model of connectivity.

In addition to new models as above, all analyses would benefit from greater real-world data, which is often hard to come by. Sadly, at least in the US, state public utility commissions not only don't release most information on DSL consumers, they don't have purview over cable modem subscribers and so such information is very difficult to obtain. Finding reliable and consistent data from regulatory authorities in most countries is a similar challenge, if not worse in developing regions.

Conclusions and Open Questions

We have argued in this paper that some of the leading theoretical formulations that define our thinking about the effects of digital inclusion and exclusion are incomplete, and divert scholarly attention from pursuing alternative lines of investigation. Metcalfe's Law, Reed's Law and Odlyzko's Law are important, but if not sufficiently challenged, risk hiding important trends where technology and social networks intersect, and may hinder the work necessary to explicate and attenuate the individual and social costs of network exclusion.

We believe the lack of consistent, theoretically informed and conceptually ambitious attention to the ***costs*** of distribution, relative to the benefits, reflects not only the impact of these widely-cited claims, but also the influence of techno-optimism more generally in the field of ICT studies. Researchers too often assume that technology trumps societal conditions in explaining diffusion, and that technological impacts are likely to bring positive results to individuals and to society.

Our analysis, while a work in progress, indicates *strongly* increasing costs to exclusion from technological networks. This work also raises a number of questions, many of which would require further analysis if not new research.

Theoretical and Conceptual Implications

If our conceptual and theoretical claims are accurate, then there are implications both for scholars and practitioners in this field. They include the following questions:

- Which of the Network Laws is appropriate for a given scenario or case for determining the value of those *in* the network (assuming our new

[14] Retransmissions by Internet routers at the TCP layer is actually quicker than the time it takes for routers to determine new routing paths at the IP layer.

methodology provides useful figures for *exclusion*)? This has implications for determining how strong the costs of exclusion are, as well.

- What are the precise transmission mechanisms that flow from an individual's low resources, to network exclusion, and back to his or her resource base? What role is played by education, skills, attitude or other intervening factors often used in the literature to analyze ICT adoption?
- Can these aspects spanning literacy, skills, finances, etc. be modeled as multiple layers or dimensions of a more complex graph showing network connectivity?
- How should we conceive of the relationships between individual choices about pursuing greater network connection, to aggregate behaviors as the outcome of those choices?
- What does social network theory tell us about the interactions of communications and information flows among societal networks of family, friends, co-workers and so forth, with access to the technologies that amplify and extend these relations? Within such models, to what extent are issues of directionality, preferential attachment, etc. captured in existing models?
- To what extent are claims made about network inclusion and exclusion within communities and nation states also applicable to inclusion and exclusion between large aggregates, including disparities between the global North and South?
- The most important questions center around our starting points for understanding asymmetrical network participation – should we begin with assumptions of equal access over time, unequal access over time, or do we believe this is essentially an empirical question whose answers are more or less stochastic across space and over time? Stated differently, what percentage of their time should scholars devote to understanding the behavior of the excluded and the under-included, and what share to those heavily involved in using the new networked technologies?

Empirical Research

As these framework–setting analyses proceed, scholars will want to seek ways to confirm or disconfirm the hypotheses offered in this paper. We would expect that while the costs of network exclusion will mount substantially over time everywhere, the exact degree of exclusionary costs will vary upon the nature of the network, the personal resources individuals bring to the network, and the alternatives available to the individual. This will require much more empirical research.

Practical Implications of New Theories of Network Exclusion

Our findings also hold implications for practitioners. At the individual level, what choices should individuals make when they seek enhanced access to networked technologies? What personal steps do they need to take to enhance the likelihood of greater inclusion and less exclusion? Are there particular strategies for accumulating the social and network capital necessary to gain access to and leverage participation in technical networks, whether through greater use of kiosks, or neighborhood or family network nodes? How should they act to gain access to secondary or tertiary networks that provide them with second or third best options, but at least include them in some technology-enhanced network?

There are also implications for strategies, policies and behaviors beyond individuals. Organizations of all types – whether community empowerment groups or local governments – need to develop procedures and programs that will decrease the likelihood of network exclusion and its manifold costs to the community. National governments and international bodies need to return to their earlier concerns about the costs of network exclusion when they design policies. These should include not only stand-alone 'ICT' policies, but also access to networked ICT resources that are embedded in health care delivery, education and business promotion services as well. Such integration is important not only across domains but also different network systems that might operate in parallel.[15]

While the horizontal spread and breadth of the costs of network exclusion may continue their statistical decline, the depth of those costs seem to be on the rise. And at some point the depth of exclusion's costs will begin to spread back into the wider society. We think a fundamental reframing of exclusion is important for not only policy-makers, but also technologists who continue to create technologies and solutions that harness and build on ICT.

References

Albert, R., H. Jeong, et al. (1999). "Diameter of the World Wide Web." Nature **401**(9 September, 1999).

Briscoe, B., A. Odlyzko, et al. (2006). "Metcalfe's Law is Wrong." IEEE Spectrum **43**(7): 34-39.

Castells, M. (1996). The Rise of the Network Society, Blackwell Pub.

Cowan, R. (1991). High Technology and the Economics of Standardization. Intenational Conference on Social and Institutional Factors Shaping Technological Development: Technology at the Outset, Berlin.

Monge, P. R. and N. Contractor (2003). Theories of Communication Networks, Oxford University Press.

Nivi. (2005). " Between Metcalfe's and Reed's Laws." Retrieved July 23, 2007, from http://www.nivi.com/blog/article/between-metcalfes-and-reeds-laws.

Odlyzko, A. and B. Tilly (2005). A refutation of Metcalfe's Law and a better estimate for the value of networks and network interconnections. Digital Technology Center, University of Minnesota.

OECD (2007). OECD Key ICT Indicators.

Steinberg, J. (2007). Converters Signal a New Era for TVs. *New York Times*, June 7, 2007

Stiglitz, J. E. (2002). Globalization and Its Discontents, W. W. Norton & Company.

[15] As an example of the importance of parallel networks and interconnections between these, the success of the ICT initiative e-*Choupal*, which used kiosks to give farmers in India crop price information and transactions, was also driven by improvements in physical supply chain.

Tongia, R., E. Subrahmanian, et al. (2005). Information and Communications Technology for Sustainable Development: Defining a Global Research Agenda. Bangalore, Allied Press.
TRAI (2007). The Indian Telecom Services Performance Indicators October - December 2006. TRAI, Telecom Regulatory Authority of India.
Watts, D. and S. Strogatz (1998). "Collective dynamics of 'small-world' networks." Nature **393**(June 4, 1998): 440-442.
WebSiteOptimization.com. (2007). "April 2007 Bandwidth Report." from http://www.websiteoptimization.com/bw/0704/.
Wilson, E. J. (2004). The Information Revolution and Developing Countries. Cambridge, MIT Press.