Netnomics
DOI 10.1007/s11066-014-9084-1

Metcalfe's law: not so wrong after all

Leo Van Hove

Accepted: 19 May 2014

Abstract Briscoe et al. IEEE Spectrum, *43*(7), 34–39 (2006) claim that Metcalfe's law is "wrong". One of their arguments is that "if Metcalfe's Law were true, then two networks ought to interconnect regardless of their relative sizes". This paper shows that this argument is flawed.

Keywords Metcalfe's law · Zipf's law · Networks · Telecommunications · Network effects

1 Overly optimistic and strategically unsound?

In a high-profile article in the July 2006 issue of *IEEE Spectrum*, Briscoe et al. [1] argue that Metcalfe's law – which states that the value of a communications network is proportional to the square of the number of its users – is wrong. Their main argument is that Metcalfe's law is substantially overoptimistic. Briscoe et al. propose instead that the value of a network of size n grows in proportion to $n\,log(n)$; that is, in accordance to Zipf's law.

This short paper refrains from taking a stance on which growth 'law' is the more realistic, as this is ultimately an empirical question.[1] Rather, the paper concentrates

[1] In a recent article, Metcalfe [15] himself fits his law to the annual revenue of Facebook over the last 10 years, as a proxy for the value created by the Facebook network. Madureira et al. [13], for their part, exploit Eurostat data to show that certain types of Internet usage indeed grow quadratically with the size of the network. However, Madureira et al.'s results are "qualitatively the same irrespectively of using Metcalfe's law or Briscoe's adaptation of it" (o.c., p. 255).

L. Van Hove (✉)
Department of Applied Economics (APEC), Vrije Universiteit Brussel (Free University of Brussels), Pleinlaan 2, 1050 Brussels, Belgium
e-mail: Leo.Van.Hove@vub.ac.be

Published online: 12 June 2014

on a second, more theoretical argument put forward by Briscoe et al. in their refutation of Metcalfe's law. Briscoe et al. argue that the law clashes with observed behaviour of networking companies, in that in the real world only companies of roughly equal size are eager to interconnect without side payments, whereas "if Metcalfe's Law were true, then two networks ought to interconnect regardless of their relative sizes". However, this inference is flawed. Briscoe et al. make two related mistakes. For one, they reason in static terms and apparently fail to realise that interconnection alters the networks' competitive positions and may, over time, very well impact their market shares. Second, Briscoe et al. mainly reason in aggregate terms, whereas the strategic implications are best analysed by looking at the utility of individual users and how this utility would be affected by interconnection. In what follows, these points are first made in a largely non-formal way; that is, in terms of the example used by Briscoe et al. themselves; see Section 2. Subsequently, Section 3 restates our criticism more formally, by relying on the theoretical literature on network externalities.

2 Static and aggregate vs. dynamic and individual

In markets that are subject to network effects, size matters. Hence, networks can create value for their members – and for themselves – by opting for interoperability and, in this way, in effect turning two or more networks into one bigger network. However, intuition would dictate (1) that such interconnection is most likely to materialise when networks are of roughly equal size (and thus have similar gains), and (2) that, in the other case, an established, dominant network has less reason to interconnect than a new, fledgling competitor – unless the latter can compensate the former by means of so-called side payments.

In their article, Briscoe et al. argue that Metcalfe's law does not correctly capture these strategic incentives. In particular, they argue that it flies "in the face of a great deal of the history of telecommunications: if Metcalfe's Law were true, it would create overwhelming incentives for *all* networks relying on the same technology to merge, or at least to interconnect. [...] Yet historically there have been many cases of networks that resisted interconnection for a long time" (emphasis added). More specifically, in a sidebar, Briscoe et al. claim that "if Metcalfe's Law were true, *no matter what the relative sizes of two networks*, both would gain *the same amount* by uniting, making the observed behavior [of resistance to interconnection] seem irrational" (emphasis added). This argument is summarised in a table; an earlier working paper by Odlyzko and Tilly [16] provides more detail.

While there is nothing wrong with Briscoe et al.'s mathematics, reasoning in terms of aggregate gains in value obscures what happens on the level of individual network members, ultimately leading Briscoe et al. to misjudge the strategic implications of Metcalfe's law. In particular, the present paper will contend that even under Metcalfe's law bigger networks are not irrational in resisting interoperability. To demonstrate this, Table 1 below builds on the example used by both Odlyzko and Tilly [16, p. 5–6] and Briscoe et al. [1]. Table 1 is in fact simply Briscoe et al.'s table with additional columns for the utility of individual users; see (a), (c), and (e).

Table 1 Individual utility and aggregate network value

	No interoperability		Interoperability		Gains	
	Per member (a)	Aggregate (b)	Per member (c)	Aggregate (d)	Per member (e)	Aggregate (f)
Network *A*, size *m*	m	m^2	$m + n = N$	$m(m+n) = m^2+mn = mN$	n	mn
Network *B*, size *n*	n	n^2	$n + m = N$	$n(n+m) = n^2+nm = nN$	m	nm

So, consider two communications networks *A* and *B* with *m* and *n* customers, respectively. Let us assume that Metcalfe's law holds; that is, let us maintain Metcalfe's assumption – the one criticised by Briscoe et al. in their first argument – that all connections are of equal value. Following Odlyzko and Tilly [16, p. 5–6)], let us also assume, for simplicity, that the constant of proportionality in Metcalfe's law is 1. Prior to interconnection, a customer of network *A* can obviously only send messages within that network, so that her utility is (proportional to) *m*; see column (a). Because a member cannot connect to herself, the correct number of communication possibilities is in fact *m1*. But for large numbers this difference does not really matter. Hence, before interconnecting the aggregate value of network *A* is $m(m\text{-}1) \approx m^2$; see column (b). After interconnecting, customers of network *A* can also send messages to the *n* persons in network *B*. In the words of Odlyzko and Tilly [16, p. 5–6)] "interconnection would provide each of the *m* customers of *A* with additional value *n* (...), or a total added value of *mn* for all the customers of *A*. Similarly, each member of *B* would gain *m* in value, so all the customers of *B* would gain total value of *nm* from interconnection. Thus aggregate gains to customers of *A* and *B* would be equal, and the two [networks] should peer,[2] if they are rational."

There is not much wrong with the above reasoning – apart from the second part of the final sentence. Indeed, by focusing on columns (d) and (f), Briscoe et al. overlook what happens in columns (c) and (e). The *aggregate* gains of *A* and *B* may be equal, but – unless $m = n$, which we will call the 'symmetric case' in Section 3 – the picture is different when it comes to the value proposition for individual users; see column (e). To highlight the strategic implications, let us focus on a setting where network *A* is substantially larger than network *B*; say, twice as large: $m = 2n$. Table 2 spells out how, in such a setting, the value created by the two networks compares (both on the individual and aggregate level), and what the implications are in terms of value capturing. Note that the framework behind Table 2 has been kept very simple in order not to stray away from Briscoe et al.'s example. In particular, abstraction is made of differences in quality and economies of scale. Moreover, switching costs are not modeled explicitly.

[2]Note that no mention is made of side payments.

Table 2 Value creation and value capturing

	Network A	Ratio A/B	Network B
No interoperability			
Value creation			
Utility, individual	$2n$	>>	n
Network value, aggregate	$m2n = 4n^2$	>>>>	n^2
Value capturing			
Consumer surplus, indiv.	$2n - p_A$		$n - p_B$
p^{max}	$p_B + 2n - n$		$p_A + n - 2n$
Interoperability			
Value creation			
Utility, individual	$3n$	=	$3n$
Network value, aggregate	$m3n = 6n^2$	>>	$3n^2$
Value capturing			
Consumer surplus, indiv.	$3n - p_A$		$3n - p_B$
p^{max}	p_B		p_A

[a] Recall our simplifying assumption that the constant of proportionality in Metcalfe's law is 1. That is, a more general expression for the individual utility of belonging to, say, network A in the absence of interoperability would be: $km = k2n$, with $k =$ the constant of proportionality. This constant would have to be estimated empirically – as in [13] – and is instrumental in turning the (non-monetary) number of possible connections into a (monetary) maximum willingness to pay, or reservation price. When computing the consumer surplus, the latter can then be compared with the price charged by the network, p_A. Similar remarks hold for the interoperability scenario.

Let us then first focus on the pre-interconnection situation. Clearly, in the absence of interoperability, and assuming that network A does not price too high, members of A are unlikely to contemplate a switch to network B, as the latter is just too small. Indeed, the utility offered by network A is twice as high (as indicated by the >> signs in the Table). Hence, as can be seen by comparing consumer surpluses, it is uncertain whether network B is able to lower its price[3] sufficiently to lure away members from network A. And if it is able to do so, network A can easily counter the move by lowering its own price. In other words, the difference in size in and of itself creates lock-in – quite apart from any switching costs involved. Conversely, members of B can far more easily be tempted to switch, provided that the costs of doing so are not prohibitively high. In short, prior to interconnecting, network A has an enviable competitive position. This is reflected in a higher p^{max}, the maximum price that a network can charge without having to be afraid that its subscribers will defect. As can be seen in the Table, the advantage in size allows network A to charge a substantial premium over the price charged by network B: $p_A^{\max} = p_B + 2n - n$. All else equal, this will result in a higher margin and, ultimately, higher profits.

[3] For simplicity, this price can be thought of as a flat-rate tariff (say a monthly fixed fee).

In the interoperability scenario, the playing field is completely different: despite the substantial difference in size, consumers now value the two networks identically. Indeed, after interconnecting, all of a sudden both networks are of size $m + n = 3n$. Any difference in valuation would thus have to come from differences in quality. Obviously, network A still has the higher aggregate network value, but this is only because of its higher number of subscribers. Most importantly, the competitive situation is now entirely different. Whereas B was initially no viable alternative to A, after interconnecting some of the customers of A may well be tempted to switch – contingent on intrinsic quality characteristics, relative price levels, and the extent of the switching costs. Network B can now also compete on an equal footing for new adopters (who are not yet locked in). Its bigger size now no longer confers a competitive advantage upon network A. This can be seen by looking at the values for p^{max}: neither of the two networks can now charge more than its competitor, everything else equal.

3 Network effect vs. substitution effect

This Section shows that the point made in Section 2 is (obviously) not new and has already been established in the network externalities literature. As a matter of fact, this literature pays a great deal of attention to the incentives that induce companies to make their networks compatible or, conversely, block such compatibility. This is not surprising because, as Katz and Shapiro [9, p. 434] put it, "[w]hen the network externalities are large, the choice of whether to make the products compatible will be one of the most important dimensions of market performance".

This said, not all models in the literature are equally relevant for the case at hand. For one, Briscoe et al. (implicitly) assume away cost advantages, differences in quality as well as switching costs. In order to stay as close as possible to the situation analysed by Briscoe et al., we should thus eschew models that analyse so-called intergenerational rivalry [5, 11, 12] – where the new technology has a clear cost or quality advantage over the incumbent technology – or models that focus on the role of switching costs [6]. Second, Briscoe et al. would seem to concentrate on markets where compatibility can only be attained via a joint decision – as opposed to a setting where a company can unilaterally make its product compatible (by means of an adapter or converter). Hence, we should look for models where technologies are sponsored and we should steer clear of models that study 'natural', *de facto* standardisation [10]. Third, we need a model with a linear externality function, so that Metcalfe's law is effectively assumed. This explains why Katz and Shapiro [9, p. 434] model, for example, was not selected. It meets all the other criteria mentioned so far but assumes marginally diminishing network effects (o.c., p. 426). Fourth, there are models that are unsuitable because the products/networks they study have features that differ too much from the communications networks that Briscoe et al. have in mind. Economides [3], for example, studies 'composite goods' or 'systems' and He et al. [8] study the within-network connectivity, or "intraconnectivity", of video game consoles – for online, multiplayer gameplay. Finally, a dynamic, multi-period model analysing duopolistic competition would be preferable.

A model that meets just about all of the above criteria is the model of Xie and Sirbu [17]. A small caveat is that it focuses on durable goods, so that there are no repeat purchases.[4] Reassuringly, Xie and Sirbu's key results are in line with our analysis in Section 2. "Under very general assumptions about the nature of the demand function" (o.c., p. 923), Xie and Sirbu find that "a compatible entrant has two opposing effects on the incumbent's profitability. On the one hand, the incumbent's product will be valued based on the total installed base, not on its installed base alone, which makes the incumbent's product more attractive. On the other hand, the entrant, by providing an alternative for the potential customers of the incumbent, may attract customers away from the incumbent or force it to lower prices to maintain sales" (o.c., p. 916). Interestingly, Xie and Sirbu also study the specific situation of a symmetric duopoly – with identical initial conditions and symmetric product costs. They find that the present value of future profits of each firm is greater with compatibility than without compatibility. In other words, in the symmetric case, "both firms will always be better off to be compatible than incompatible, no matter what the installed base at $t = t_0$ (*provided that they are equal*)" (ibidem; emphasis added). However, in the asymmetric case, Xie and Sirbu find that when network externalities are strong, "the larger the installed base of the incumbent the more relative advantage it gives up by agreeing on compatibility" and "the less is the incentive to support compatibility. The incumbent will suffer a pure loss from compatibility when its installed base is large enough" (o.c., p. 923). Conversely, the later entrant realises a higher profit when its product is compatible with that of the incumbent than when it is not, and the more so the larger the installed base of the incumbent.

In the terminology of Matutes and Padilla [14, p. 1114], who analyse not duopolistic competition but a situation with three players,[5] the above results can be summarised in terms of a trade-off between a (positive) network effect and a (negative) substitution effect. Which of the two effects dominates, and whether compatibility will effectively prevail, depends on the relative strength of the companies, the strength of the network externalities, and the modus operandi and costs involved in making the networks compatible. In the case that interests us most in this paper – the asymmetric case – the two firms have conflicting interests: the firm with the largest network will oppose compatibility, whereas the weaker firm will be in favour. If compatibility requires a joint decision as would appear to be the case in the Briscoe et al. example – the outcome is that the networks will remain incompatible.[6]

[4]This said, Xie and Sirbu (o.c., p. 924) themselves claim that "the dynamic demand function [they] derived can be applied to both durable goods and network service markets".

[5] In particular, Matutes and Padilla consider three banks with Automated Teller Machine (ATM) networks of equal size that need to decide whether to share their networks. A striking result, which is in line with the results of Xie and Sirbu is that the banks – despite the *ex ante* symmetry of their networks – will never agree on full compatibility. In equilibrium, either partial compatibility or total incompatibility will prevail. The intuition behind this result is straightforward: "If all banks share their networks none of them obtains a network advantage so that the only effect of compatibility is to reduce the effective degree of horizontal differentiation among banks. As a result, competition gets tougher and banks are unable to internalise the positive network externality, [...]. It follows that two compatible banks would always veto the entry of a third bank into their common network" [14, p. 1122].

[6]Always assuming no side payments can be made.

The underlying intuition is self-evident: the dominant network is against compatibility because interconnection with the smaller network would only have a limited impact on consumers' willingness to pay for its product. As a result, the network effect would be outweighed by the adverse effect of more intense price competition due to the increased substitutability.[7] The problem with Briscoe et al.'s argument is that they reason purely in terms of the network effect and overlook the substitution effect.

4 Overly optimistic? Quite possibly. Strategically unsound? No

The upshot of the analysis in both Sections 2 and 3 is that, unless the two networks are of roughly equal size one of the networks has every reason to resist interoperability. The competitive advantage of large networks is precisely that they can offer access to a large network whereas smaller networks cannot. Hence, Metcalfe's law does not need "obtuseness on the part of the management" ([16, p. 4] or "stunningly inefficient financial markets" [1] to explain the refusal to interconnect without payment. While it is more obvious under the alternative 'law' put forward by Briscoe et al.,[8] Metcalfe's law is no less able to convey "that there may be sound economic reasons for larger networks to demand payment for interconnection from smaller ones" [16, p. 6].

All this also explains why in an industry characterised by technological compatibility – think, for example, of the Global System for Mobile communications (GSM) standard in the mobile communications industry in Europe – larger operators are often quick to create some degree of artificial or 'economic' incompatibility by making use of price discrimination between on-net and off-net calls. Unsurprisingly, smaller operators are more likely to offer similar conditions to their users regardless of the destination of their calls; see [7, p. 6] for illustrations.

To sum up, Metcalfe's law may well prove 'wrong' in real life, but the theoretical argument that Briscoe et al. advance in their refutation is flawed. Metcalfe's law does *not* imply that "two networks ought to interconnect regardless of their relative sizes" [1] and is no less consistent with real-world behaviour of networking companies than Briscoe et al.'s alternative. Hence, to the best of our knowledge, there is no valid theoretical reason to reject Metcalfe's law. The empirical data will have to do the talking.

Acknowledgments I would like to thank my colleague Marc Jegers for the stimulating discussions, and an anonymous referee and Julian Wright for their comments. Special thanks go to the Editor for his guidance and encouragement.

[7] These conclusions may continue to hold when one explicitly allows, as Esser and Leruth [4] do, for a difference in 'basic' quality between the goods of the two firms. Esser and Leruth show that when network externalities are strong, the dominant network will prefer incompatibility because "[it] benefits so much from the externality that the (large) size of its own market is sufficient to insure high profits" (o.c., p. 265).

[8] It is more obvious because under Zipf's law the aggregate network effect will be smaller for the bigger network than for its smaller competitor.

References

1. Briscoe, B., Odlyzko, A., Tilly, B. (2006). Metcalfe's law is wrong. *IEEE Spectrum*, *43*(7), 34–39. http://spectrum.ieee.org/jul06/4109.
2. Briscoe, B., Odlyzko, A., Tilly, B. (2010). When a "law" isn't a law at all. *International Commerce Review*, *8*(2), 146–149.
3. Economides, N. (1991). Compatibility and the creation of shared networks. In Guerin-Calvert, M., & Wildman, S. (Eds.) *Electronic services networks: A business and public policy challenge*. New York: Praeger.
4. Esser, P., & Leruth, L. (1988). Marketing compatible, yet differentiated products: in search of competitive equilibria when network externalities are at work. *International Journal of Research in Marketing*, *5*(4), 251–270.
5. Farrell, J., & Saloner, G. (1986). Installed base and compatibility: innovation, product preannouncements, and predation. *American Economic Review*, *76*(5), 940–955.
6. Farrell, J., & Shapiro, C. (1988). Dynamic competition with switching costs. *RAND Journal of Economics*, *19*(1), 123–137.
7. Fuentelsaz, L., Garrido, E., Maicas, J.P. (2012). A strategic approach to network value in network industries. Forthcoming in *Journal of Management*. Published online 31 May 2012. doi:10.1177/0149206312448399.
8. He, T., Kuksov, D., Narasimhan, C. (2012). Intraconnectivity and interconnectivity: when value creation may reduce profits. *Marketing Science*, *31*(4), 587–602. doi:10.1287/mksc.1120.0705.
9. Katz, M.L., & Shapiro, C. (1985). Network externalities, competition, and compatibility. *American Economic Review*, *75*(3), 424–440.
10. Katz, M.L., & Shapiro, C. (1986a). Technology adoption in the presence of network externalities. *Journal of Political Economy*, *94*(4), 822–841.
11. Katz, M.L., & Shapiro, C. (1986b). Product compatibility choice in a market with technological progress. *Oxford Economic Papers*, *38*(0), 146–165.
12. Katz, M.L., & Shapiro, C. (1992). Product introduction and network externalities. *Journal of Industrial Economics*, *40*(1), 55–83.
13. Madureira, A., den Hartog, F., Bouwman, H., Baken, N. (2013). Empirical validation of Metcalfe's law: how Internet usage patterns have changed over time. *Information Economics Policy*, *25*(4), 246–256. doi:10.1016/j.infoecopol.2013.07.002.
14. Matutes, C., & Padilla, A.J. (1994). Shared ATM networks and banking competition. *European Economic Review*, *38*(6), 113–1138.
15. Metcalfe, B. (2013). Metcalfe's law after 40 years of ethernet. *Computer*, *46*(12), 26–31. doi:10.1109/MC.2013.374.
16. Odlyzko, A., & Tilly, B. (2005). A refutation of Metcalfe's law and a better estimate for the value of networks and network interconnections, mimeo. Digital Technology Center, University of Minnesota. http://www.dtc.umn.edu/~odlyzko/doc/metcalfe.pdf.
17. Xie, J., & Sirbu, M.M. (1995). Price competition and compatibility in the presence of positive demand externalities. *Management Science*, *41*(5), 909–926.