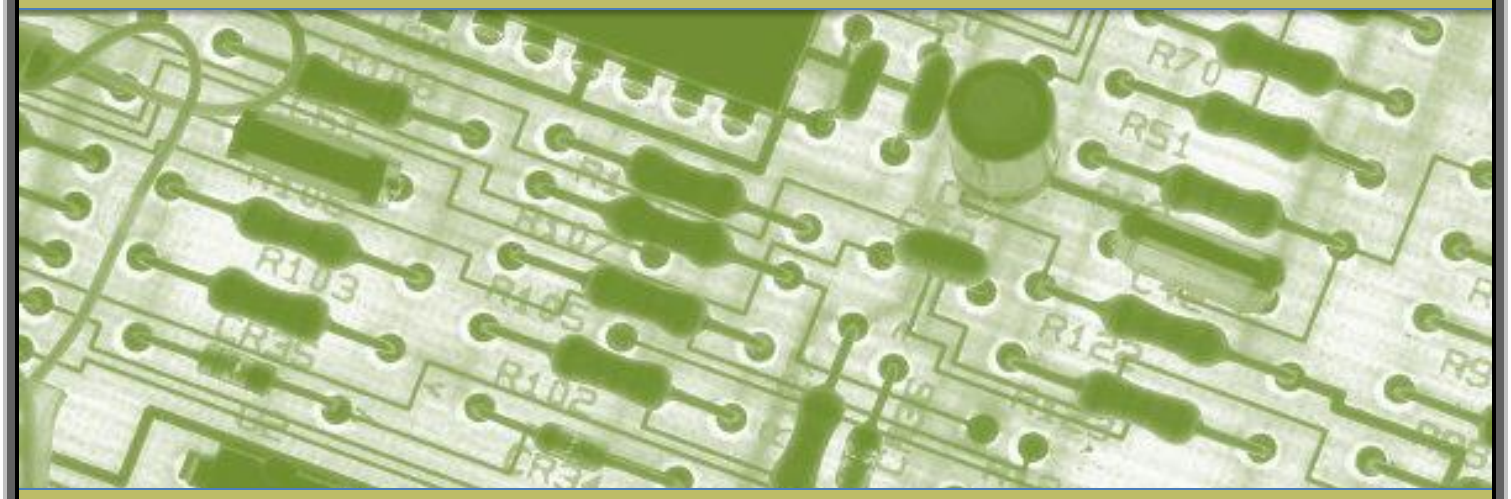
HOOVER IP²

WORKING PAPER SERIES
NO. 19004

ARE "FANGs" MONOPOLIES? A THEORY OF DISEQUILIBRIUM COMPETITION WITH UNCERTAINTY

NICOLAS PETIT
UNIVERSITY OF LIEGE AND HOOVER INSTITUTION

APRIL 16, 2019



Hoover Institution Working Group on
Intellectual Property, Innovation, and Prosperity
Stanford University

www.hooverip2.org

ARE "FANGs" MONOPOLIES?
A THEORY OF DISEQUILIBRIUM COMPETITION WITH UNCERTAINTY

Nicolas Petit*

Introduction

This paper lays down the rudiments of a descriptive theory of competition among the digital tech platforms known as "FANGs" (Facebook, Amazon, Netflix and Google), amidst rising academic and policy polarization over the answer to what seems to be – at least at the formulation level – a simple question: are FANGs monopolies?

To date, two streams of thought pervade the competition policy debate. On the one hand, works in favor of the monopoly motion insist on FANG's control of a large share of output in relevant product(s) or service market(s), high barriers to entry, lateral integration and strong network effects. Some of these works also implicate harder to estimate, and potentially novel, harms like reductions in privacy, labor market monopsony and distortions of the democratic process.

On the other hand, a line of argument skeptical of FANG monopolies argues that traditional monopoly harms are not manifest in FANGs. To the contrary, FANGS would outperform textbook monopolies by observable metrics of prices, output, labor or innovation. In addition, the tech industry is rife with examples of once dominant later irrelevant companies like MySpace, AOL or Yahoo!., inviting caution against anticipative monopoly findings.

Both perspectives carry weight in competition and regulatory decision-making. Yet confidence in the idea that FANGs are monopolies is on the rise, as evidenced by the rapid succession of

* Professor of Law, University of Liege (ULiege) and Liege Competition and Innovation Institute (LCII); Visiting Scholar at Stanford University, Hoover Institution; Research Professor, University of South Australia (UniSA). Nicolas.petit@uliege.be or petitn@stanford.edu. I would like to thank Doug Melamed, Michal Gal, Alex Galetovic, Axel Gautier and Thibault Schrepel for their comments on a previous version of this paper. I am also grateful to Jerome de Cooman for excellent research assistance. This paper complies with ASCOLA Transparency and Disclosure Declaration. I have nothing to disclose.

competition and regulatory findings against them.[1] And because inactive competition and regulatory agencies do not make dispositive findings of absence of monopoly, there seems to be a perception that the FANG monopoly argument is winning.

Against this backdrop, the current structuring of the FANG conversation makes it tempting to pick one side or the other, leading us to neglect one hard underlying question: is the binary monopoly-non monopoly model applied in competition and regulatory decision-making the right framework to diagnose and discipline anticompetitive conduct and transactions in FANG industries, or should it be replaced by an alternative theory?

We shall entertain no certainties about this question. Economic models on which our competition laws are arrayed may lose relevance as a result of industry change.[2] This can be understood with a mythological metaphor. Daedalus assumed that the sun was only a shining star. His model for flying thus underestimated the risk that wax wings would burn as his son Icarus would get closer to it.[3] Sir John Hicks generalized the point: "*since it is a changing world that we are studying, a theory which illumines the right things at one time may illumine the wrong things at another … There is, there can be, no economic theory which will do for us everything we want all the time*".[4]

With an open mind, this paper thus attempts to look at whether the textbook monopoly model is the appropriate framework to analyze digital markets. It suggests that observed average tendencies of FANGs expose the limitations of the textbook monopoly model (I), proposes an

[1] *See* EU Commission decisions against Google of 2017 and 2018, Australian Competition and Consumer Commission report of 2018, German Bundeskartellamt decision against Facebook of 2019, and EU recent opening of an investigation against Amazon.

[2] The problem exists in other areas of competition law. Cartel laws rely on models of competition with and without agreement, making it uneasy to fight welfare reducing tacit collusion or intra group conspiracies.

[3] Though in real life, we know that temperature gets colder as one moves closer to the sun. See Kyle Hill, 'Forget Icarus, Fly As Close To The Sun As You Want!' (*Discover Magazine*, 22 April 2014) <http://blogs.discovermagazine.com/but-not-simpler/2014/04/22/forget-icarus-fly-close-sun-want/>.

[4] John R. Hicks, 'The Scope and Status of Welfare Economics' (1975) Oxford Economic Papers, New Series, Vol. 27, No. 3, 307-326, <https://www.jstor.org/stable/2662172?seq=1#metadata_info_tab_contents>. Hicks however noted that revolutions in economy are changes in "*attention*", not of the science itself as in physics.

alternative theory of disequilibrium competition with uncertainty (II), and considers competition law and policy implications (III).

In our modern times, all papers seem to require a pitch. Ideally the selling point should be counterintuitive and straightforward. Here's one (hoping its simplistic formulation does not exhaust the reader's interest in what I hope is a rigorous analytical treatment of the issue): competition policy should be relaxed when power over price is highest, and strengthened when it is lowest.

Before we turn to the discussion, allow us two more remarks. First, theory formulation and validation require careful observation. Since we ambition to test a model and propose an alternative theory, we use empirical data taken from FANG 10-K reports to the Securities and Exchange Commission ("SEC") aware of their limitations. Second, we use the term monopoly throughout the paper to avoid monotony. We therefore dispense with more legally appropriate qualifications like dominance or significant market power.

I. FANG v Textbook Monopoly Model

As we often do when we talk about patent rights, we associate FANG with monopolies. We are wrong on patents but right for FANGs. Each FANG holds a large share of output in a market where entry is limited. And the fact that high prices, low output and reduced innovation are not manifest in FANG is irrelevant, because absence of evidence is not evidence of absence.

What is, however, critical, is that FANGs display significant motivational differences from the textbook equilibrium model where the monopolist equalizes marginal revenue and costs (A). It is precisely this feature that makes the traditional monopoly model a poor framework to discuss FANG firms (B), and that justifies the search of an alternative theory.

A. TEXTBOOK MONOPOLY MODEL

1. Theory

In the pure world of economic theory, the monopolist is a dictator with absolute powers. No entrant, input seller or buyer can influence its decisions. But how, then, does the lone monopolist set an output and price level combination?

The textbook monopoly model is a response to this decision making mystery. Monopoly output and price setting is a "marginalist" process. Assuming profit maximization, the monopolist grows output and lowers prices up to the level where marginal revenue ("MR") equals marginal cost ("MC"). Put differently, the monopolist decides to produce an extra quantity of output if (and only if) this yields a revenue greater than the costs incurred to produce a marginal unit.[5] Of course, in practice, marginalism is not applied by monopolists. But what matters is that the model emphasizes the high level constraints of falling revenue and increasing (or constant) costs that structure a monopolist's decisional context.

In the textbook monopoly model, MR declines as the quantity of output rises. This is because the monopolist is confronted with a falling demand curve for his product.[6] To sell more output, the monopolist must lower the price to get (all) people to buy more units of output.[7] As served buyers experience satisfaction through consumption, they derive marginally less benefits from extra units, and are thus willing to pay less for them. Moreover, MR is lower than price at each level of output, because all previous units must be sold at a lower price too.

[5] Here is an untechnical example to help the non-economist to understand the monopolist's thinking: do marginal returns on producing 10 additional pages in a long working paper compensate the marginal costs of writing them? In this example, marginal returns are reader's interest, downloads or citations to the paper. And in both the metaphor and the model, marginal returns tend to decrease when more pages are added to the paper, at least when the reader is a journalist or policy maker.

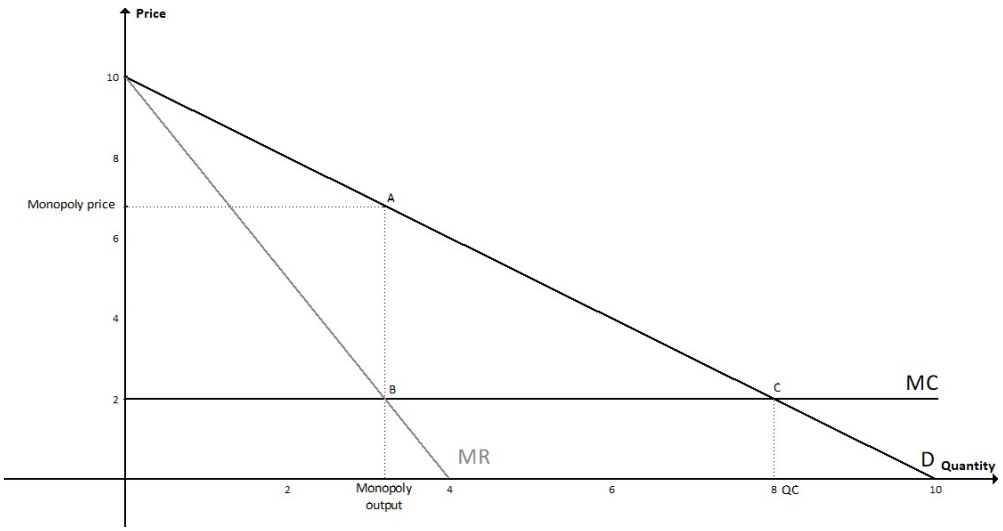
[6] Abba P. Lerner, 'The Concept of Monopoly and the Measurement of Monopoly Power' (1934) The Review of Economic Studies, Vol. 1, No. 3, 157-175 (hereafter Lerner, 'Concept of Monopoly').

[7] John Taylor, *Principles of Microeconomics* (Cengage Learning 2011) 255.

A logical implication is that the monopolist's marginal profit (MP) decreases up to the point of equality between MR and MC. At the output level where MR=MC, the monopolist's total profits are maximized. MP is 0. This is called a stable *equilibrium*, because no other output level can make the monopolist better off: one more unit produced, and the monopolist registers a loss; one less unit produced, and he misses a profit opportunity.

From a social welfare standpoint, the profit maximizing equilibrium level of output leads to a monopoly price level superior to MC. The monopoly equilibrium imposes a loss on society, because some customers ready to pay a price lower than the monopoly price but higher than MC are not served. They must divert their purchases to other, less satisfactory,[8] purchases which must cost more to produce, even though they are relatively less expensive.[9] This allocative inefficiency of the monopoly is coupled with a variety of other harms stemming from insulation from competition. In the familiar parade of horribles come cost inefficiency, low innovation and rent seeking behavior.

Figure 1 – Textbook Monopoly Model



[8] Abba P. Lerner, 'Concept of Monopoly' 157.
[9] Richard A. Posner, *Antitrust Law* (2nd edn University of Chicago Press 2001) 12.

2. Two Monopoly Decision Making Properties that Matter

The textbook monopoly model theorizes an equilibrium tendency. Besides the standard assumption of profit maximization, a critical condition for a monopoly equilibrium is that MR and MC converge, and eventually intersect. In turn, this means that only industries with two properties can move towards the socially inefficient monopoly equilibrium. One is decreasing MR (even if total revenue increases with output).[10] The other is decreasing MP (because MC are deemed constant or increasing).[11] This last condition also seems to hold in the natural monopoly model with increasing returns to scale.[12]

Both properties are linked to what economists call exogenous factors. These are factors on which no firm has no direct control, even a monopolist. Decreasing MR is the consequence of the falling demand curve.[13] And decreasing MP is a consequence of the decreasing marginal product of labor.[14] As output expands, workers are less productive.[15] Monopolists must take these for granted, like dictators must accept to be overthrown by uprisings when citizens are left starving.

B. ARE FANGS DECISION MAKING MONOPOLISTS?

When we discuss FANGs amongst antitrusters, we seldom think about the decision making properties of monopolies. Instead, the discussion invariably focuses on either structural factors or welfare outcomes. Yet, the two properties of decreasing MR and decreasing MP provide

[10] Note that a monopolist may experience an increase in MR if he reduces output after having wrongly set its initial production level too high, so that MR<MC and MP is negative.
[11] Even though there can be convergence when MC decrease slower than MR, though this seem to be a rare scenario.
[12] In a natural monopoly, long run average costs decline with output, because MC is lower than average costs. However, even in a natural monopoly situation, MP decrease because MC and MR intersect at some point. See Richard A. Posner, 'Natural Monopoly and its regulation' [1968] Chicago Unbound <https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=2861&context=journal_articles>.
[13] As seen before, when customers increase their consumption, they derive less marginal benefits from additional units, and are hence willing to pay less.
[14] Note that in an extreme case, MP can be decreasing even with a zero production cost simply because MR is decreasing.
[15] And more costly. Additional workers, overtime plans or additional compensation may be needed. But note that even if wages are constant, there still is decreasing marginal product of labor.

useful testable hypotheses. In particular, we may be tempted to use firm level data to get a better understanding of FANGs' decision making universe, and try to draw some inferences on whether they behave like textbook monopolists.

This is what we try to do in the following sections. We test the properties and results of the textbook monopoly against a dataset covering Facebook, Amazon, Netflix and Google's Securities and Exchange Commission (SEC) 10K filings. The basic finding is that FANG seem to operate in a manner inconsistent with the textbook monopoly model (3). Let us see why this is the case, first by setting out our approach to measurement (1), and then by showing our results (2).

1. Evaluating FANGs Decision Making

Saying that FANGs are monopolists (as many do) should be a way to suggest that as they increase output, they experience a decrease in MR, and move society one step closer to the inefficient equilibrium level where MR intersects with MC, in turn justifying competition policy or regulatory intervention. If the textbook monopoly model appropriately applies to FANGs, we should thus be able to observe decreasing MR and MP at firm level.

To assess this, we need to know FANG's year over year revenues, costs and profits. We can retrieve this data from their 10Ks. However, SEC reporting firms are not required to disclose information about output or quantities sold.[16] While some companies like Amazon like acknowledge "*increased sales*",[17] we have no readily available quantitative measure of output that entitles us to compute revenues, costs and profits on a marginal basis, as required to estimate MR and MP.

[16] This is actually difficult in services industries that supply intangibles services.

[17] Amazon 10K for 2011 <https://www.sec.gov/Archives/edgar/data/1018724/000119312512032846/d269317d10k.htm>.

The Statista premium database provides useful data points that help overcome this difficulty. In particular, we can retrieve Amazon's number of active customer accounts worldwide over 1997 to 2015; Facebook's number of daily active users worldwide 2011-2018; and Netflix streaming subscribers in the United States from 3rd quarter 2011 to 4th quarter 2018. As to Google, its flagship search product has no clear output metric. Search does not require an account or subscription. Yet, the Statista premium database gives data on "*explicit core search queries*" on Google sites in the US over 2008 to 2018. For all firms, except for Google since 2014, output measures have grown.

By crossing SEC data and the Statista premium database, we can estimate "*approximate*" MR, MC and MP. We use the term "*approximate*", because what we actually calculate is the incremental evolution of the average revenue, costs and profits made by each FANG on each customer account, user, subscriber or core search query as output grew on a year over year basis. We can however take those estimates as proxy for marginal measures on the basis of two considerations. The first is not extreme, but deserves explication: it is not our purpose to define what the *level* of MR and MC is, but essentially if both metrics follow a decreasing or increasing *tendency*.[18] In this context, we believe average measures are informative because a property of averaging is that a group average moves in the direction of the contribution of the latest additions to the group.

The second is that we equate reported measures of "*costs of goods sold*" ("COGS" or costs of revenue/sales) to variable costs. COGS include all costs associated or allocated to products sold, and include labor costs, input and material costs, marketing expenses, and other costs required for the sale of products. COGS may be more inclusive than variable costs. We believe this is not a problem, for COGS can only bias our MC estimates upward and our MP estimates

[18] We assume that when average revenue falls, this implies that MR is below average (and that when average variable costs fall, MC is below average).

downwards. Using COGS is a conservative choice because we err on the side of the textbook monopoly hypothesis.

One last remark is in order. We have had to tweak the data for Google. While Google provides indirect data on US revenue from 2008-2017, costs of sales are not broken down for the US. To overcome this problem, we applied Google's reported ratio of US to global revenue to costs figures. For example, when Google reported that 49% of its total revenue was generated in the US, we decided that 49% Google's total costs of sales could be allocated to the US.[19]

2. Results

According to the prediction of the textbook monopoly model, as output grows we should observe a dual decline in FANGs approximate MR and MP. Let us see to what extent this is the case.

With our data, we can estimate, with varying degree of accuracy, FANGs approximate MR and MP on a year over year basis. Figure 2 represents FANGs MR (in blue), MP (in red) and MC (in grey). The resulting diffusion curves show two regularities. First, the MR of all FANGs rises though with heterogeneity at firm level. Second, and perhaps most important, MP is increasing too. FANGs thus seem to violate two essential conditions of the textbook monopoly model.

These observations raise an intriguing question: if both MR and MP increase, is there a possibility of short term equilibrium? A logical implication of a simultaneous increase in MR and MP is indeed that MC remains close to constant (as can be seen from Facebook), decreases or rises more slowly than MR (as can be seen for Google). In all three cases, this means that

[19] A more extreme approach could have been to think of Google producing in the US, and exporting elsewhere. On that basis, we could have allocated all COGS to Google's US revenue. We show the result of this in Appendix 2. The analysis does not change significantly.

there is no convergence between MR and MC. The wedge between both curves increases, and short term equilibrium is unlikely.

Figure 2 – FANGs Approximate MR and MP (moving averages)

3. Discussion

The FANG monopoly hypothesis fails the test of evidence. Unlike a profit maximizing monopoly, the decision making of a FANG cannot be about setting short term output so that MR=MC, absent a foreseeable perspective that both curves will intersect. This has two logical implications. First, assuming that MR=MC remains a valid profit maximization proposition for FANG firms, it must be a long term perspective.[20] Second, FANG firms' short term profit maximization approach must be about something else.

[20] If we consider it a reasonable proposition to think that there are income constraints and declining marginal product of labor.

But what is it? Let us look another time at our data, and think of further differences with the textbook monopoly model. In standard economics, MR reflects the slope of the demand curve. A possible interpretation of rising MR is that FANGs face an upward slopping or shifting demand curve in the short term, denoting higher user willingness to pay as output rises.[21] This could be a fundamental difference with the textbook monopoly model where demand is downward slopping and fixed.[22]

All this suggests that FANGs' short term goal should be to grow output, period. This is consistent with anecdotal observations of early loss making by FANGs like Amazon, who prioritized growth over (net) profits in early years forecasting a long term divergence between rising MR and MC.

True, output also grows in the textbook monopoly model. First, output grows up to the profit maximizing level. And output may also grow strategically, when the monopolist attempts to keep potential rivals out of the market. But a difference with FANGs is that when output increases in a traditional monopoly setting, it is always bounded: intersecting MR and MC prevents the monopolist to serve all quantities demanded. And when the threat of entry disappears, expansion ceases. This is not the case in a FANG setting, where output growth seems unconstrained.

All this cautions against calling FANG monopolies, and drawing early policy implications. But all this leaves also open a critical question: if FANG are not monopolies in the short term, how do they compete?

[21] Indeed, FANG may face a falling demand curve that is shifting upwards year after year. We discuss this distinction in the following sections.

[22] True, output also grows in the textbook monopoly model. It does up to the profit maximizing level. And it may also grow strategically, when the monopolist attempts to keep potential rivals out of the market. But when output expansion occurs in a monopoly setting, it is always bounded: by MC or in time once the threat of entry is gone. This is not the case in a FANG setting, where output growth is unconstrained.

II. How do FANGs Compete?

How do firms compete when the demand curve is upward slopping or shifting? We are not in darkness here. Since the 1970s, economics study firm behavior in markets with "network effects" where users' willingness to pay initially increases with quantity demanded.[23] A much underappreciated tenet of that literature emphasizes the disequilibrium properties of markets with rising demand (A). That finding is critical because it corroborates the uncertainty properties that FANGs and others describe when they discuss the industry environment (B).[24]

A. COMPETITION WITH AN UPWARD DEMAND CURVE

1. Upward Demand Slope or Shift?

Let us start with a simplified presentation of the canonical model of network effects. In a market with network effects, users' willingness to pay (WTP) depends on the number of other users consuming the service. Economists use a variety of concepts to denote this: "*demand side economies of scale*",[25] "*positive consumption externalities*",[26] "*network externalities*",[27] and many others. In all cases, the common idea is that each user's marginal benefit is based not only on the value (v) of the service's functional attributes– but also, and more remarkably, on the number (n) of (expected) users to join the network. This leads to a function $p(x)=v(x)n(z)$,

[23] Other streams of economics studies consider upward slopping demand curve in markets with bandwagon effects. See Harvey Leibenstein, 'Bandwagon, Snob, and Veblen Effects in the Theory of Consumers' Demand' [1950] The Quarterly Journal of Economics, Vol. 64 No 2 183-207 (hereafter Leibenstein, 'Bandwagon'); Gary. S. Becker, 'A Note on Restaurant Pricing and Other Examples of Social Influence on Price' [1991] Journal of Political Economy 99 (5) 1109-1116.

[24] The literature network effects literature is copious, though much work concentrates on topics like lock in, winner takes all, switching costs, price discrimination and bundling, important to decision makers interested in improving social or private gains.

[25] Hal R. Varian and Carl Shapiro, *Economics of Information Technology* (Cambridge University Press 2004) 33.

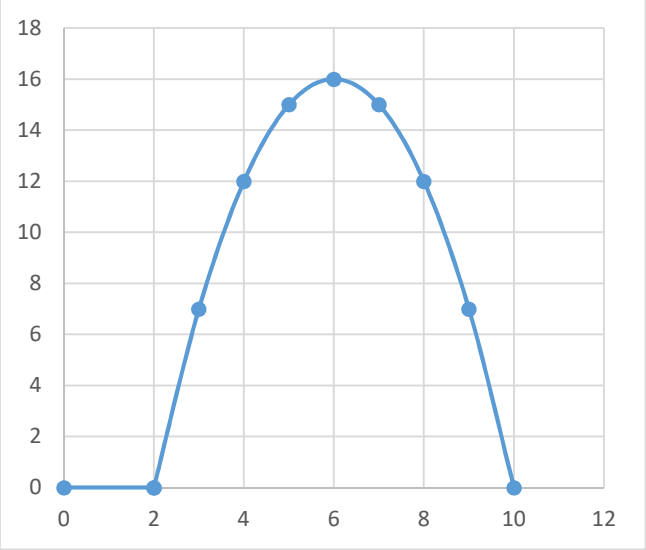
[26] Nicholas Economides, 'The economics of networks' [1996] International Journal of Industrial Organization 14 673 at 678 (hereafter Economides, 'Economics of networks').

[27] Michael L. Katz and Carl Shapiro 'Network externalities, competition, and compatibility' [1985] American economic review, 75(3), 424-440 (hereafter Katz and Shapiro, 'Network externalities'.

where (p) is the reservation price, x is an individual user and z is the population that can use the network.

Two diagrammatical representations of networks effects populate the literature. The first gives an unordinary inverted U curve shape to the demand curve.[28] At low levels of (other) users' adoption of a network service, the network's value is null. The demand curve is flat. No one is willing to join. As more users join, the marginal WTP for the service increases. The demand curve slopes upward. At some point, adding users to the network brings positive, yet lower marginal value. WTP stagnates falls. The demand curve slopes downward.

Figure 3 – Demand Curve with Network Externalities



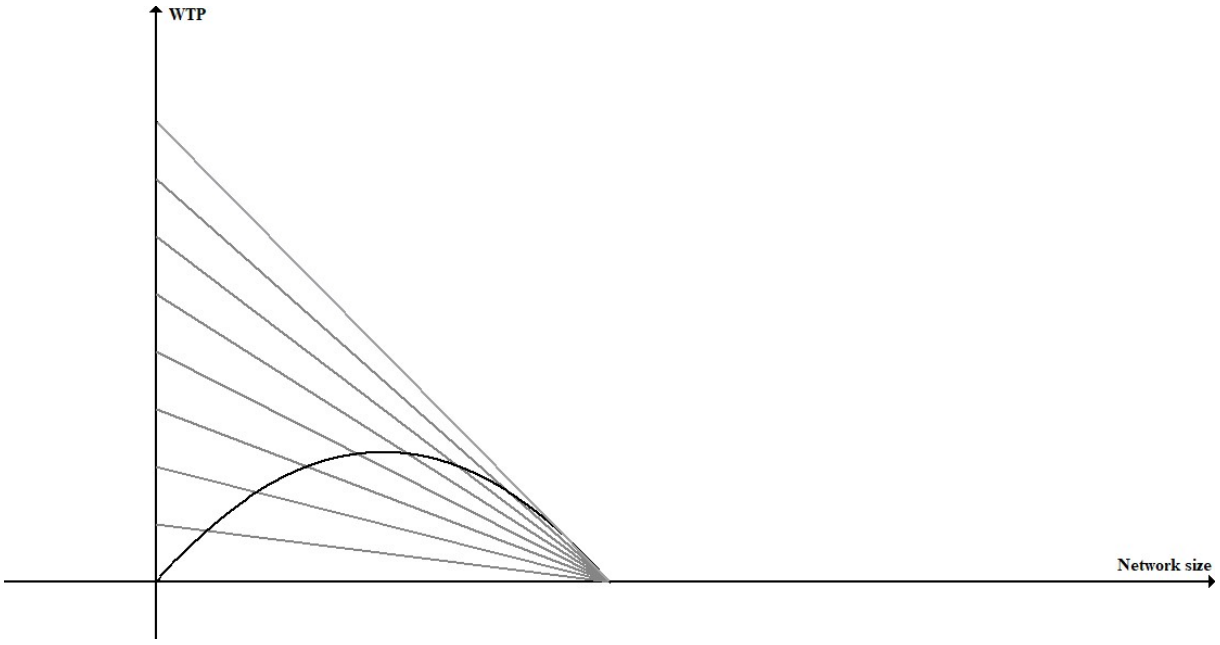
The intuition behind this phenomenon is easy to understand: a marginal increase in my consumption of a network good produces positive externalities on third parties who join the platform, triggering an increase in my own consumption and WTP. This positive feedback loop stops at some point, when the marginal benefits that I derive from the network stem more from its own functionality that other users' adoption.

[28] We discuss below both the individual demand curve and the market demand curve.

The second diagrammatical representation of network effects markets is one in which the demand curve is downward slopping, but shift upwards with increases in users that join the network.[29] In this variant, each individual user derives positive, yet decreasing marginal benefits from network service consumption. The demand curve slopes downward. Yet, because increases in individual consumption produce externalities on third parties, additional users join the network size, increasing marginal benefits, and with it WTP, for the individual user. The demand curve shifts upward.

With an eye on figure 4, we can see that each individual marginal user WTP increases for some time, until it decreases. Again, if we trace out the locus of (virtual) equilibrium for a network of size X (that is, we experience X-1 shifts in the downard slopping demand curve), we find that a typical user's demand curve slopes upward and then downward.[30]

Figure 4 – Alternative Demand Curve with Network Externalities



[29] Economides, 'Economics of networks' 678 ("Network externalities "*signifies the fact that the value of a unit of the good increases with the number of units sold. To economists, this fact seems quite counterintuitive, since they all know that, except for potatoes in Irish famines, market demand slopes downwards. Thus, the earlier statement, "the value of a unit of a good increases with the number of units sold," should be interpreted as "the value of a unit of the good increases with the expected number of units to be sold." Thus, the demand slopes downward but shifts upward with increases in the number of units expected to be sold*.".

[30] The crossing points are virtual equilibrium points. On this, see the early work of Leibenstein, 'Bandwagon' 194.

The demand curve slopes upward as long as there is an increasing marginal external consumption effect. And then it slopes downward when the value derived from increases in n(z) wears out. In that case, v(x) recovers a determinant role, yielding decreasing marginal benefits as under the standard demand curve.

The difference between both representations is a matter of exposition, not ideology. In the first case, the price effect and the consumption externality effects are accounted for together.[31] In the second representation, the price effect and the consumption externality are accounted for separately.[32]

Whichever diagrammatical representation best represents the real world (we could actually invent more of them, depending on how we formulate the problem), what matters more is that in network effects markets, the demand curve may contain an upward-sloping segment.[33]

2. Disequilibrium

The unordinary demand curve met in network effects markets has a key implication. When firms operate in the upward slopping region of the demand curve, there is no stable equilibrium. A perturbation in market conditions – including one due to the firm's own decisions – can precipitate the firm towards success or ruin. By contrast, when the demand curve slopes downward, a perturbation in market conditions – including one due to the firm's own decisions

[31] Paul Belleflamme and Martin Peitz explain: "*[T]he impact of the Law of Demand (according to which the quantity demanded decreases with the fee) is more than compensated by the network effect (which increases the consumers' willingness to pay as demand expands)*". See Paul Belleflamme and Martin Peitz, *The Economics of Platforms* (Cambridge University Press 2018) Chapter 3 mimeo (hereafter Belleflamme and Peitz, The Economics of Platforms).

[32] Traditional economists who consider that "*an upward slopping demand curve is inconsistent with economic theory*" tend to prefer the second representation. See Robert S. Pyndick and Daniel Rubinfeld, *Microeconomics* (9th edn Peason 2017) 737. We ought to note here that Gary Becker, hardly an untraditional economist, used the first method in his work on bandwagon effects. See Gary S. Becker, 'A Note on Restaurant Pricing and Other Examples of Social Influences on Price' [1991] Journal of Political Economy 99 (5): 1109-16.

[33] Belleflamme and Peitz, *The Economics of Platforms*. In appendix 3, we show a diagrammatical representation of a discontinuous demand curve for a network market, where the network effects appear on the left hand side part of the diagram, and where the demand curve recovers its traditional slope on the right hand side.

de – will have trivial influence. Following a period of adjustment, the market will pursue its march towards the high participation equilibrium.[34]

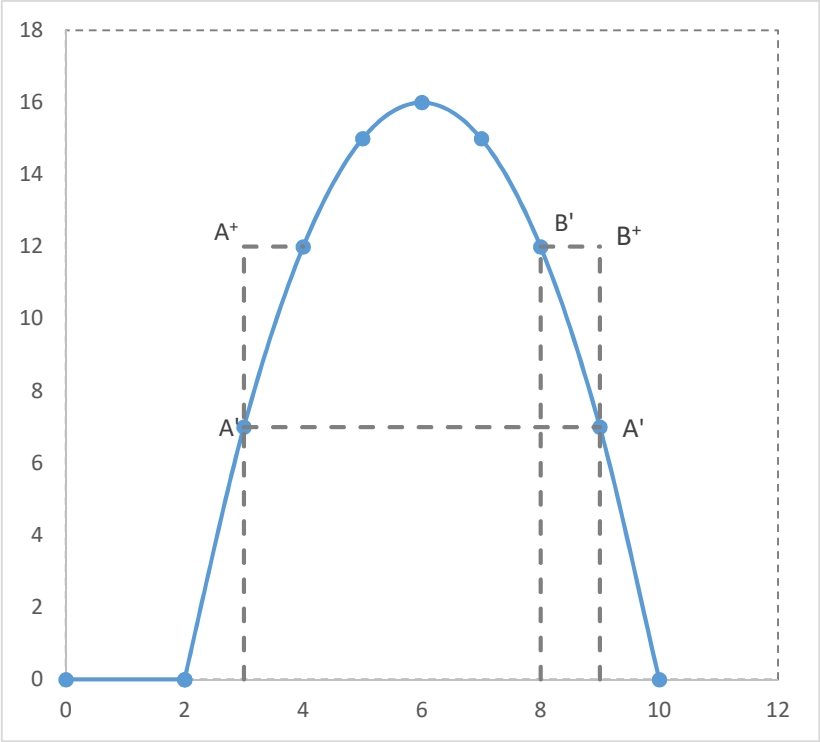
You do not need to blindly trust the model to understand this. A simple diagrammatical illustration drives the point home. Take the conventional representation of network users' WTP in Figure 5 below (WTP relates for instance to a subscription fee, tolerance to ads or willingness to share personal data). A firm launches a social network service with a potential scale of 1 million users. The challenge for the firm is to recover a given level of production costs without undermining users' base growth. The firm knows that the demand curve is upward slopping but it cannot perfectly estimate users' WTP. How to best monetize?[35] Suppose now that the network has developed a small users' base of 3, and that demand is growing. The firm introduces ads at level A+. This exceeds users' WTP. Marginal users leave the platform. This triggers a negative feedback loop of churn, that comes to an end with the collapse of the service at equilibrium point 0. Had ads been set below WTP at level A, the price would have been below the installed base reservation price, and the network would grow progressively until A'.[36]

Figure 5 – Inverted U-Curve Representation of Network Users' WTP

[34] Economists talk of a "*fulfilled expectations demand curve*" (Nicholas Economides, 'Competition Policy in Network Industries', in Dennis W. Jansen (ed.) *The New Economy and Beyond*, (Edward Elgar Pub 2006) 101-103). Behind this complicated term, lies a fairly simple idea: failure to add new users/keep existing users in the growth stage of the market decreases the realized utility of past adopters below their initial expectations. This leads to network churn. By contrast, failure to add new users/keep existing users in the mature stage of the market does not reduce utility below the expected level of quality that has been realized. This leads to network rigidity.

[35] An example of the following approach showing how pricing decisions can place firms below or above equilibrium, and influence success or failure in reaching critical mass, see David Easley and Jon Klinberg, *Networks, Crowds, and Markets: Reasoning about a Highly Connected World* (Cambridge University Press 2010) 17.3.

[36] Or, to put it differently, a discrete decrease in price from A+ could result in a large increase in quantity, consistent with traditional demand curve effects.



Now suppose that the network has a users' base of 9. The firm introduces ads at level B+. Again, this price level exceeds users WTP. The firm's users' base will decrease to 8, and not below (point B'). What we see here is that the costs of wrong business decisions is not qualitatively similar depending on whether a firm operates in the upward or downward slopping region of the market demand curve.

Economists use various expressions to refer to the unstable situation of firms that operate in the upward slopping region of the demand curve: "*disequilibrium*",[37] "*out of equilibrium*",[38] "*multiplicity of equilibria*",[39] or "*unstable equilibria*".[40] To understand this important concept better, let us spend a minute on the meaning of equilibrium and disequilibrium. A standard definition of equilibrium is that of a situation in which there is no room for voluntary improvement. Put differently, as a market moves towards equilibrium, firm behavior is less

[37] Jeffrey Rohlfs, 'Theory of Interdependent Demand for a Communications Service' [1974] Bell Journal of Economics and Management Science, 5 (1), 16-37 (hereafter Rohlfs, 'Theory of Interdependent Demand').
[38] David S. Evans and Richard Schmalensee, 'Failure to Launch: Critical Mass in Platform Businesses' [2010] Review of Network Economics, Vol. 9: Iss. 4, Article 1 (hereafter Evans and Schmalensee, 'Failure to launch')
[39] Belleflamme and Peitz, *The Economics of Platforms*.
[40] Rohlfs, 'Theory of Interdependent Demand' 16-37.

determinant of market outcomes. By contrast, in a disequilibrium market, firms and other economic agents' purposeful choices matter more. In disequilibrium, the universe of business decisions confronting the firm is larger, the impact of decisions is bigger, and the number of decision makers is larger. To capture the uncertainty typical of disequilibrium, Schumpeter talked of a "*a ball that is perched on the top of an inverted bowl*".[41]

The bottom line? As Belleflamme and Peitz write, "*a direct consequence of the existence of multiple equilibria is unpredictability*".[42] Both the firm, its competitors, and external observers face uncertainty as to how the market will behave. Moreover, prices lose relevance, for the market can sustain different network sizes for the same price.[43] And though static economic analysis can typically locate these multiple equilibria, it usually "*cannot tell us which one will be 'selected*".[44] In brief, business seems more risky when demand is increasing than when it is falling.[45]

3. Critical Mass

a) Brief overview of the critical mass principle

As much as disequilibrium seems to generate uncertainty, network effects markets display another property that seems, at least facially, to work in reverse. The key concept here is "*critical mass*". In nuclear engineering, critical mass denotes a chain reaction that becomes self-sustaining.[46] In network effects market, the idea is similar. Firms that reach a critical mass

[41] Joseph A. Schumpeter, *Historic Economy Analysis* (Allen & Unwin 1954).
[42] Belleflamme and Peitz, *The Economics of Platforms*.
[43] Hung-Ken Chien and C. Y. Cyrus Chu, Durable-Goods Monopoly with Network Effects, Marketing Science, Vol. 27, No. 6 (Nov. - Dec., 2008), pp. 1012-1019.
[44] W. Brian Arthur, 'Competing Technologies, Increasing Returns, and Lock-In by Historical Events' [1989] The Economic Journal (n 394) 116-131 (hereafter Arthur, 'Competing Technologies).
[45] Of course, the model assumes reversibility of users' participation choices, but this is a reasonably likely assumption especially when the demand is growing. By contrast, when demand is falling, reversibility is more constrained (for example, people have accumulated many friends), and this strengthens the equilibrium effect. Evans and Schmalensee, 'Failure to launch'.
[46] Thomas Schelling draws a parallel with nuclear engineering ("An example is "critical mass". An atomic pile "goes critical" when a chain reaction of nuclear fission becomes self-sustaining; for an atomic pile, or an atomic bomb, there is some minimum amount of fissionable material that has to be compacted together to keep the reaction

of users can expect a high participation equilibrium. By contrast, firms that do not reach a critical mass of users can expect network collapse.

There is a right and wrong way to think about critical mass in network effects markets. To caricature, the wrong way pitches critical mass as an economic model predicting that as soon as a firm in a network effect market crosses a fixed point of user adoption, it ignites an automatic cycle of self-reinforcing technology diffusion that ends with near market monopoly.

The right way looks at the principle of critical mass not as a theory, but rather as a "*mental model*"[47] apt to characterize complex "*recurrent behavior patterns*" when "*people's behavior depends on how many are behaving a particular way*" in technology markets (but also in other walks of life like sports or university seminars).[48]

Though economists are predominantly on the right side of the discussion, a number of hidden complexities of the critical mass principle leads non-economic savvy readers to cultivate misconceptions about it.[49] To paraphrase Hal Varian talking about network effects, critical mass is one of those ideas "*that you can explain to a regulator in five minutes and they can talk about it for five months*".[50]

More specifically, the simplicity of the intuition behind the critical mass principle obfuscates that it is actually a driver of complexity for firms operating in network effects markets. Let us discuss this in more detail.

---

from petering out."). See Thomas C. Schelling, *Micromotives and macrobehavior* (W. W. Norton & Company 2006) 89 (hereafter Schelling, *Micromotives*). Arthur, 'Competing Technologies'..

[47] Tren Griffin, 'Two Powerful Mental Models: Network Effects and Critical Mass' (*Andreessen Horowitz,* 7 March 2016) <https://a16z.com/2016/03/07/network-effects_critical-mass/ accessed 12 April 2019> (hereafter Griffin, 'Two Powerful Mental Models').

[48] Schelling, *Micromotives* 94.

[49] Schelling observes that the concept of "mass" is inappropriate in physics, and that "number", which is its economic functional equivalent, is also unsatisfactory See Schelling, *Micromotives* 95 ("whether the measure is the number of people engaged, or the number times the frequency or the length of time they engage in it, or the ratio of the number who do to the number who do not, or the amount of such activity per square foot or per day or per telephone extension, we can call it a "critical-mass" activity")

[50] Hal Varian, 'Use and abuse of network effects' (SSRN, 7 Augustus 2018) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3215488> accessed 12 April 2019.

b) Critical mass is not a fixed point

The critical mass threshold does not correspond to a set number of users against which firms can assess network adoption performance.[51] Rather, the critical mass threshold is a *range* of numbers. In empirical works, this range is often expressed as a ratio of market users to market potential.[52]

Most economic works suggest that the critical range is difficult to compute.[53] One difficulty owes to the fact that the critical range is an aggregate of individual users' vision of what constitutes critical mass,[54] and users utility functions vary widely.[55] There is therefore a lot of heterogeneity to account for.[56] This can be best understood with a real life example: a teenager's critical mass of connections on a social network is likely lower than Victoria Beckham's. And yet, to bring the teenager on board, the social network must convince Victoria

[51] Many works that insist on the watershed point or tipping point suggest this, ___.

[52] Michal Grajek and Tobias Kretchsmer 'Identifying critical mass in the global cellular telephony market' (2012) International Journal of Industrial Organization, 30(6), 496-507 (hereafter Grajek and Kretchsmer, 'Identifying critical mass'). ("Empirical work on critical mass focuses on identifying a percentage – typically varying between 10% (Mahler and Rogers 1999) and 25% (Cool et al. 1997) – of market potential as critical mass"); Evans and Schmalensee, 'Failure to launch'.

[53] Virtually all economics works insist on this point. See Bob Briscoe, Andrew Odlyzko and Benjamin Tilly 'Metcalfe's Law is Wrong' (*IEEE Spectrum*, 1 July 2006) <https://spectrum.ieee.org/computing/networks/metcalfes-law-is-wrong> accessed 12 April 2019: "The fundamental flaw underlying both Metcalfe's and Reed's laws is in the assignment of equal value to all connections or all groups"

[54] David Allen 'New telecommunications services: Network externalities and critical mass' (1988) Telecommunications policy, 12(3), 257-271 (hereafter Allen, 'New telecommunications services') ("*each person has an individual vision of what constitutes critical mass*").

[55] Jose Luis Arroyo-Barrigüete, Ricardo Ernst, Jose Ignacio López-Sánchez and Alejandro Orero-Giménez 'On the identification of critical mass in Internet-based services subject to network effects' (2008) The Service Industries Journal Vol. 30 n 5, 643-654. The key problem with determining critical mass is defining the utility functions of the individuals who make up the market. "The fundamental problem is that the definition of each individual's utility function will be much more complex". For instance, if you take the user of a PSN, then its utility function is conditional of at least two types of interaction with different people which are difficult to model (i) sphere of close influence; and (ii) rest of the population.

[56] Schelling, *Micromotives* ("Though perhaps not in physical and chemical reactions, in social reactions it is typically the case that the "critical number" for one person differs from another's"; "When people differ with respect to their cross-over points, there may be a large range of numbers over which, if that number of people were doing it, for a few but only a few among them that number wouldn't be big enough, while the rest would be content. When those few for whom the number is not enough drop out, they lower the number, and some more drop out, and so on all the way. The fact that in the end nobody is doing it does not give us any measure of *how many* satisfied participants were lacking at any point along the way." => add that users act myopically); Allen, 'New telecommunications services' ("each person has an individual vision of what constitutes critical mass").

Beckham to join (and use) the platform.[57] Like other marquee users, however, Victoria Beckham will not sponsor a low subscription platform, raising the critical mass threshold challenge for the social network.

Besides, another difficulty is that the critical range is not fixed. For indeed, market potential keeps changing as firms grow. The critical range is thus a moving target. Consider this: our estimate of Amazon's critical range is not the same depending on whether one looks at it as a firm that seeks to become an online book retailer (in 1997), the world's "*online*" retailer (in 1999), or Earth's most consumer centric company (in 2003, notice the disappearance of any reference to online).[58] Or think about Netflix's entry into streaming content production in 2008. This imposed on Netflix a new critical mass challenge, for it subsequently had to build a library of titles of sufficient "*mass*" to convince viewers from other content networks to join its platform.[59]

One may still ask, however, whether the locus of the critical mass threshold can be expected in the lower or the upper region of the upward slopping demand curve. But again, there is no clear response to this question in the literature. In fact, the critical mass principle is one of these areas where the ratio of theory to evidence is quite large.[60]

[57] In multi sided markets of this kind, this is usually achieved with subsidies or other incentives.

[58] See Amazon 10-K for 2003 <https://www.sec.gov/Archives/edgar/data/1018724/000119312504029488/d10k.htm>.

[59] We could multiply examples. Facebook is one of them. Its critical mass threshold has evolved since the days when it launched as undergrad college service at Harvard to become a worldwide unrestricted social networks. Uber eats is another one. Though Uber's installed base of users in ride sharing services gives certain advantages, its critical mass of drivers is of no assistance when it comes to recruiting new restaurants.

[60] In a thorough literature review from 2015, Ewa Lechman writes: *"it is worth underlining that despite a relatively well-developed theoretical framework and conceptual background aiming to explain the 'critical mass'-like phenomenon, the number of empirical works seeking a quantitative assessment of it is very limited*". See Ewa Lechman, *ICT Diffusion in Developing Countries: Towards a New Concept of Technological Takeoff* (Springer 2015) 54 (hereafter Lechman, *ICT Diffusion*).

What this all means is simple. Firms are living organs. And all growing firms, including large ones, face critical mass challenges.[61] Most successful rock n' roll bands contemplating stadium performance know this.

c) Critical mass is not exogenous

Unlike in nuclear engineering, critical mass effects do not arise in markets due to the laws of nature.[62] Already in 1985, Katz and Shapiro noticed this property. Observing that one would like a theory that tells what leads to the zero adoption or to the high participation equilibrium, they asked "*what can consumers and firms do to influence the market outcome*".[63] In hindsight, Katz and Shapiro early captured the intuition that because network effects vary in strength, more is needed to trigger mass technology adoption.

35 years later, we still miss the theory, but we know a lot more. First, market institutions and firms' strategies influence technology diffusion. This influence is highest when firms are below the critical range, and lowest when firms are above.

Second, below the critical mass, price is only one of the numerous devices available to reach the critical mass.[64] By contrast, when the critical mass is reached, "*installed base effects drive diffusion even in the absence of price decreases*".[65] Crossing the critical mass thus simplifies the set of firms' profit maximizing options, including by plausibly allowing exercise of power over price.

[61] One diagrammatical implication of this point is that network size cannot be normalized, and the demand curve should shift upward, without being tied to a fixed point on the horizontal axis.
[62] Some authors have proposed the following assumption: See, Lechman, *ICT Diffusion* 50 for an overview of the literature.
[63] Katz and Shapiro, 'Network externalities'.
[64] Katz and Shapiro provide a list of firms strategies (pricing commitments on complementary software, "second sourcing", vertical integration, "penetration pricing", etc..
[65] Grajek and Kretchsmer, supra 47 note that the emergence of critical mass is conditioned on "*strength of installed base effects, the size of the installed base, and the current market price*". For this definition, see Grajek and Kretchsmer, 'Identifying critical mass'.

Third, when the demand grows in network effects markets, prices essentially serve as signals.[66] By contrast, prices recover their mediating role of supply and demand above the critical mass.[67]

Fourth, externalities cause higher effects below the critical range. Witt writes "*When critical point is high, innovation by third parties may lead to random fluctuations*".[68] This implies that there is higher uncertainty below the critical range than above.

With all this, the conclusion is straightforward: unlike what often encountered terms like "*snowball*", "*feedback loop*" or "*bandwagon effects*" connote, there is no magic in markets with network effects. And claims about the self-sustaining nature and exponential rate of technology diffusion above the critical mass should be taken with a grain of salt.[69]

d) Critical mass is a two ways street

There is no ratchet in markets subject to critical mass. Firms can fail to sustain market relevance in spite of users' mass adoption of their product and services. In his seminal work on critical mass, Nobel prize winner Thomas Schelling uses the metaphor of the "*dying seminar*" to illustrate the situation of a university gathering that slowly peters out in spite of (sustained) academics' interest.[70]

One specific reason why critical mass effects are not unidirectional in network markets is the following: as a network expands, users' growth yields a tyranny of connections (and more generally of activity). Any Facebook user who once considered removing friends understands this. Firms may therefore occupy what looks like an optimal above critical mass position, yet register network participation decline. Facebook traditional way to deal with this problem has

[66] Allen, 'New telecommunications services'. This in turn, entails the equilibrium possibility of introductory pricing in network effects markets; Luis M. B. Cabral, David J. Salant and Glenn A. Woroch 'Monopoly pricing with network externalities' (1999) International Journal of Industrial Organization 17, 199-214.
[67] Allen, 'New telecommunications services' (the market is mature).
[68] Ulrich Witt, '"Lock-in" vs. "critical masses" – Industrial change under network externalities' (1997) International Journal of Industrial Organisation, 15, 753-773
[69] For they are under-representative of the richness and sophistication of the economics literature.
[70] See Schelling, *Micromotives*, 92.

consisted in tweaking the algorithm of its newsfeed. To make things concrete, Facebook has decided to show top stories or family content, rather than relying on mere chronological order.

e) Critical mass is not firm but market specific

Often, the critical mass story is not told with care. Many works – including this one – discuss if, how and when firms reach critical mass, when they should instead talk about markets. To put things differently, it is not a firm that reaches critical mass, it is the market.

To see what happens when the focus of analysis shifts, let us quote Tren Griffin: "*What happens if a market does "tip"... but it's the competitor that reaps those benefits?*".[71] Griffin offers the following example: MySpace made the early investments in social networks, pushing the market towards critical mass. But MySpace was impatient. Facebook was not. Facebook's slow monetization strategy eventually paid off.

The deeper point made by Griffin is that investments into critical mass are imperfectly appropriable.[72] Competitors can free ride on rivals' network specific investments. This property of network effects markets can work both to the benefit of late entrants, as in the Facebook and MySpace story, and to benefit of incumbents, as evidenced by Facebook development of Instagram in a market initially occupied by Snapchat.

B. UNCERTAINTY

When network effects markets are in a state of disequilibrium, firms behave under uncertainty. If we consider FANGs' disclosures in 10K reports (1) as well as industry specific fact patterns (2), there are sound reasons to believe that the model of disequilibrium competition with uncertainty is apt to characterize their environment.

[71] See Griffin, 'Two Powerful Mental Models'.
[72] The other point is that rival firms have distinct perceptions of critical mass levels.

1. FANGs Declarations

One regularity seen in FANGs 10Ks is a reported fear of disruption. Here is a sample of the most commonplace FANG statements on disruption: "*Our business is characterized by rapid change as well as new and disruptive technologies*"; "*[...] many of the areas in which we compete evolve rapidly with changing and disruptive technologies, shifting user needs, and frequent introductions of new products and services*"; "*Our business is characterized by innovation, rapid change, and disruptive technologies*". Even more graphically, Google CEO Eric Schmidt once declared "*somewhere, someone in a garage is gunning for us*".[73]

Outside of the Silicon Valley, FANGs statements leaves many cold. Critics like to deride FANGs under confidence as PR talk intended to defuse attention from the public opinion, the media and regulatory agencies. They rightly recall that Amazon, Netflix and Google are 20 years' old or so. As years pass, the comparison with "murdered by disruption" giants like AOL, Blockbuster, Kodak, MySpace, Polaroid and SUN Microsystems becomes untenable.

At the same time, the disequilibrium properties of markets with an upward demand curve and the inherent estimation difficulties involved in identifying the critical range bring theoretical backing to FANGs alarmist declarations.

Moreover, it is not all sure that the risk of disruption that burdens FANGs is one of terminal exit. Competitive extinction is perhaps a FANG concern, but competitive irrelevance is certainly a bigger one. To put the point clearly, FANGs do not want to become ghost towns like MapQuest, Nokia or Yahoo!.[74]

---

[73] Eric Schmidt, 'The New Gründergeist' (*Google Europe Blog*, 13 October 2014), <http://googlepolicyeurope.blogspot.com/2014/10/the-new-grundergeist.html>. See also Dominic Rushe, "Jeff Bezos tells employees 'one day Amazon will fail", *The Guardian*, 16 Nov 2018 ("*Amazon is not too big to fail … In fact, I predict one day Amazon will fail. Amazon will go bankrupt. If you look at large companies, their lifespans tend to be 30-plus years, not a hundred-plus years*").

[74] For use of the expression ghost town, see Alex Moazed and Nicholas L. Johnson, Modern Monopolies: What It Takes To Dominate the 21st Century Economy (St Martin's Press 2016).

Last, there is one more reason to give currency to FANGs' expressed under confidence. Several of them today draw a significant share of their profits (and therefore of their ability to grow) from accidental discoveries: Amazon discovered the lucrative cloud services market accidentally and Netflix did not initially believe in online streaming.[75] If this happened to them, why would not this happen to another firm?

2. Industry Facts

When we look at sources of uncertainty in FANG markets, we observe a universe of risk factors all linked to the emerging properties of digital industries. To start, uncertainty stems from demand side factors. Estimating users' WTP is difficult. If this was not the case, the venture capital market would not exist.[76] Many factors enter a users' preference functions. Take privacy. If users of social networks value privacy in absolute terms, the relative weight of privacy is unclear when compared with other factors like convenience, personalization or transaction costs economies. This creates a challenge of setting the right monetization strategy, both level and time wise.

Besides, users' WTP in an upward sloping or shifting demand market is not constant. WTP may decrease due to users' bad behavior. Common examples include MySpace's users' exodus following reports of inappropriate content on the platform or Atari's demise due to failure to lock out unauthorized games.

By contrast, when a network has overcome the critical mass constraint, there is less uncertainty. AOL, once dominant online Internet provider in the US, provides a good example. In 1996, AOL moved from pricing by the hour to a monthly 20 hours' access subscription with

[75] In 2017 Amazon drew 9,81 % of its net sales from AWS. See Amazon 10K for 2017 at https://www.sec.gov/Archives/edgar/data/1018724/000101872418000005/amzn-20171231x10k.htm.
[76] Since it is based on high risk bets with low probability of success. We also would all be as rich as Jeff Bezos or Marc Zuckerberg.

incremental fees.[77] Intense customer backlash did not prevent AOL from further growing its users base in the following years. Arguably, AOL's post 1996 growth took place in the downward slopping region of its demand curve. Closer to us, Netflix brings a possible example of a firm that may have reached critical mass, and knows it. In its 10-K for 2009, Netflix declared: "*We have achieved a level of scale in our business that provides many operational and competitive advantages [...] Such scale economies also have contributed over time to expanding operating margins which has made it possible for Netflix to aggressively price its service offering at levels difficult for competition to meet*".[78]

Uncertainty can arise from the supply side. Recall that in FANG markets, MP increase. When this is understood, the implication is clear: network markets in disequilibrium offer highest profit opportunities. Mario Rizzo goes as far as stating: "*profits exist only in a world of uncertainty and disequilibrium*".[79]

The practical economic consequence is straightforward. One should observe relatively higher competitive entry in the portions of the demand curve that slopes upward. Again, anecdotal evidence of bold FANG strategic moves seems to carry the point. Take, for example, Microsoft's Bing attack against Google's Internet search service or Google's Android attack against Apple's closed smartphone ecosystem or Apple's entry into entertainment to challenge Netflix. Alternatively, disequilibrium profit opportunities may be the business rationales behind FANGs attempt to cream skim profitable market segments. Through that light, Microsoft's purchase of professional social network LinkedIn looks like specialist entry into social networks. The same can be conjectured of the launch of FB's 10 minutes' video uploading IGTC service, which looks a lot like a pointed attack at Google's YouTube.

---

[77] Kara Swisher, *AOL.com* (Three Rivers Press 1999) 160-162.
[78] See Netflix 10K for 2009 at https://last10k.com/sec-filings/nflx/0001193125-10-036181.htm ==
[79] Mario J. Rizzo, 'Disequilibrium and all that' in Mario J. Rizzo (ed.), *Time, Uncertainty and Disequilibrium* (Lexington Books 1979) 10.

Complementors are also a source of uncertainty. To think about this, consider the example of apps and operating systems ("OS"). On the one hand, apps bring added value to an OS, and increase users WTP for the OS. On the other hand, apps capture value from the OS. The underlying economics are straightforward: due to an "income constraint" – economic agents have a finite resources – users faced with a new complementary application B or C, will devote relatively less to A.[80] Of course, A may still benefit from an anchoring effect.[81] Yet, it is readily apparent that complements generate ambiguous effects. Add to this that complements sometime reconfigure the structure of an industry. In his famous Tidal Wave memo of 1995, Bill Gates speculated about how the Internet, a complement to Microsoft's OS and productivity software. Gates correctly conjectured the appearance of a whole host of competing products, like new file formats, browsers and even less expensive devices for Web browsing. Competitive entry in adjacent, neighboring or complement product spaces reinforces uncertainty, and therefore yields pressure on incumbents' products.

Last, firms in disequilibrium markets may not only be victims, but also active agents of uncertainty. Because third party entry dissipates the likelihood of long term equilibrium profits, incumbents in network markets are incentivized to look for emerging or future disequilibrium markets, adding even more uncertainty to the business environment. In the 1990s, AOL spent millions on developing its own movie and media offering, only to discontinue it a few years later. In the 2000s, Microsoft invested successfully in games, and less successfully in online press content. And today, we see many FANGs entering online payments or wearable devices like connected glasses or watches.

[80] If you live in a neighborhood without a Starbucks, then you have a budget for detergent at Walmart that is x; now assume there is a Starbucks and you like it. You will spend y less at Walmart, where y<x.
[81] For admittedly, the loser might be unrelated pizza night.

It is time to move on. Now that we have pointed out to the existence of a competitive constraint due to uncertainty in the disequilibrium region of FANG markets, we must complete our description by specifying normative public policy implications. So what are they?

III. Competition Law & Policy Implications

At this stage, the reader probably wonders: where's the law? To close this paper, we discuss the competition law and policy implications of our previous findings.

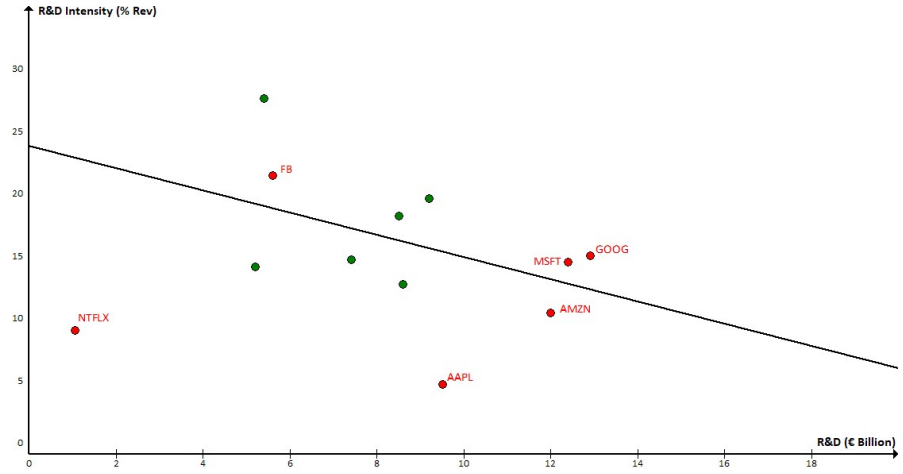
A. ADJUSTMENTS TO COMPETITION LAW AND POLICY FRAMEWORKS?

From a public policy perspective, FANG markets represent a challenge. In disequilibrium, the firm is more fragile, which may be procompetitive. At the same time, because the firm is more fragile, it may have incentives to take steps to remove uncertainty, which may be anticompetitive.

Observed fact patterns in FANG markets lean both in the competition and the monopoly directions. On the one hand, FANGs display patterns of behavior more consistent with well-functioning competitive markets than with monopoly power that leads to consumer harm. Believers in FANG competitive markets, and FANG firms themselves, often point us out to data points as diverse as investments in R&D (see figure 4 below comparing FANG's R&D (red dots) with pharmaceutical companies' (green dots)), slow monetization strategies (*eg*, years of null profits at Amazon) and cost reduction plans (*eg*, Netflix's vertical integration in content).[82]

[82] In his 1974 paper, Rohlfs noted that that attaining the socially optimal equilibrium user set "*may require ruinous (albeit temporary) promotional costs*". See Rohlfs, "Theory of Interdependent Demand" 19.

Figure 6 – FANG R&D Intensity and Expenditure (+Microsoft and Apple)



On the other hand, several patterns of observed behavior are more consistent with the textbook monopoly model, like fee increases, ad cluttering or planned obsolescence.

Besides these polar examples, disequilibrium strategies are difficult to categorize as pro or anticompetitive. Think about cross platform integration of complements through M&A (*eg*, Facebook's acquisition of Instagram), preferential treatment (*eg*, Google's integration of maps on its search engine and mobile OS Android), bundling (*eg*, Netflix's bundle of DVD and streaming subscriptions), imitation (*eg*, Amazon's cloning of merchants' products) or exclusive dealing (*eg*, app stores' bans on third party distribution). In a disequilibrium context, cross platform integration is a well-accepted strategy to grow network effects. At the same time, cross platform integration reduces reversibility, increases switching costs and exacerbates lock-in. This may deprive competing firms in disequilibrium markets from profit maximizing network externalities … or strengthen their incentives to invest into disruptive innovation and inter platform competition. As if this was not all, history is of little help. More than 20 years after the Microsoft antitrust saga, we still entertain doubts as to whether Microsoft's anticompetitive strategies towards rivals were not a self-inflicted wound. After all, a credible argument can be built that Microsoft "take-no-prisoner" approach to complement software applications led the Redmond firm to underestimate the commercial potential of nascent

disruptive technologies, and incentivized the computer industry to move elsewhere leading to the emergence of Google, Facebook and myriad other firms.

Overall, characterizing firms' strategies in disequilibrium markets probably requires adjustments of competition law and policy's frameworks.[83] Hereafter, we discuss several options for law and policy reform. All are essentially first principles.

B. POLICY PRIORITIES: EQUILIBRIUM, NOT DISEQUILIBRIUM?

From a comparative statics perspective – the perspective usually taken in competition policy – markets in equilibrating tendency represent higher risks of higher social costs. This is due to convergence between marginal benefits that decrease (the demand curve slopes downward) and/or marginal costs that increase (the supply curve slopes upward). While above cost pricing in a disequilibrium market is a moot problem (an overpriced service will instantly collapse to the 0 equilibrium), it is a possible source of deadweight loss in an equilibrium market.

The interesting consequence for a competition policy framework committed to consumer welfare (CW) is that it should place priority emphasis on equilibrium markets. Counterintuitively, this also means that some competition policy instruments should be relaxed when power over price is relatively highest (when marginal profits increase during disequilibrium) and strengthened when it is relatively lowest (when marginal profits decreased during equilibrium).

But can we refocus competition law and policy towards certain types of markets where levels of competitive harm exceed certain threshold levels? At a high level, nothing pleads against this. Competition law proscriptions embody many threshold rules like the concept of

[83] See already in 1998 (with reference to predatory pricing), Lemley, Mark A. and McGowan, David, Legal Implications of Network Economic Effects. 86 Cal. L. Rev. 479, 1998 ("*Arguments based on network effects may suggest that the law must rethink the rationality of behavior considered un-likely under neoclassical theory, such as predation in antitrust jurisprudence, and address new risks not considered under models based on declining returns*").

dominance in unilateral conduct law; of appreciability in coordinated conduct law; or the screen of harm to interbrand competition in vertical agreements cases.

Make no mistake. Our suggestion is not to introduce a rule of *per se* legality for disequilibrium markets with uncertainty. In disequilibrium markets, firms may have ability and incentives to reduce uncertainty in ways that are anticompetitive. Moreover, firms that operate at a higher point of the upward slopping demand curve may try to protect relative competitive advantages by recourse to anticompetitive means. Put simply, we do not exclude that anticompetitive conduct can occur on the road to equilibrium. There should remain regulatory ability to enforce applicable competition laws in such markets.

C. ANTITRUST LAW

There are two elements to a violation of antitrust law: bad conduct and market power (and a causal link between them). Let us investigate how the concept of disequilibrium competition challenges both.

1. Significant Market Power: Structure v Pressure?

In competition law, a great deal of work is spent assessing market structures. Structural analysis ultimately seeks to determine whether a firm is dominant or hold significant market power.

Under a structural analysis, many firms in network effects markets are likely to be deemed dominant on the ground that they control a large share of output even though their environment is one of disequilibrium where uncertainty is highest.

The problem with this is both economic and legal. The economic issue is the well-known type I error cost that arises when a firm in a competitive market type is deemed in a dominant position, and instantly subject to a specific set of legal constraints under the doctrinal concept

of "*special responsibility*".[84] The legal problem is that the conventional definition of dominance makes no sense in a disequilibrium market with uncertainty. Recall that in the seminal *United Brands* and *Hoffmann-La Roche* cases, the Court of Justice of the European Union ("EU") defined dominance as:

> "*a position of economic strength enjoyed by an undertaking which enables it to prevent effective competition being maintained on the relevant market by affording it the power to behave to an appreciable extent independently of its competitors, customers and ultimately of its consumers*"[85]

As we have seen, a firm confronted with a network effects markets in growth phase is to some extent *dependent* on marginal users' adoption choices (by contrast, the firm's existence is not at stake when the demand curve slopes downward). If we believe that unilateral anticompetitive conduct is possible in disequilibrium markets, the definition does not help and creates a type II error.[86]

Unless we are ready to tolerate an error prone structural method to measure market power as well as of an empty definition, we may try to formulate a more suitable concept, definition or measurement of significant market power that appropriately distinguishes the properties of procompetitive and anticompetitive disequilibrium markets.

Again, a deep dive in the economics literature can help here. An idea that transcends many works on monopolies is that dominant firms are subject to low *pressure* to change. In 1935, Sir John Hicks captured this idea in the famous line "*the best of all monopoly profits is a quiet life*".[87] Hicks' suggested that monopolists being likely "*people with sharply rising subjective*

[84] *Michelin v Commission* [1983] ECLI:EU:C:1983:313, para 57.
[85] *United Brands v Commission* [1978] ECLI:EU:C:1978:22, 65 and *Hoffmann-La Roche v Commission* [1979] ECLI:EU:C:1979:36, 38.
[86] Because antitrust decision makers will not be able to observe the independence characteristic of dominance.
[87] Hicks' suggested that monopolists being likely "*people with sharply rising subjective costs*" may actually "*not [be] bothering to get very near the position of maximum profit*", in violation of the venerable profit maximization condition of economic theory.

costs" may actually "*not [be] bothering to get very near the position of maximum profit*", in violation of the standard profit maximization condition of economic theory.

Decades later, micro economist Harvey Leibenstein expressed the same idea. According to Leibenstein, the real social harm of monopoly is not as much seen in supracompetitive prices, than it is in a lack of motivational efficiency. Firms subject to low competitive pressure display "*inertial behavior*", and occupy "*non optimal effort positions which persist over time*". Leibenstein writes:

> "*In situations where competitive pressure is light, many people will trade the disutility of greater effort, of search, and the control of other peoples' activities for the utility of feeling less pressure and of better interpersonal relations. But in situations where competitive pressures are high, and hence the costs of such trades are also high, they will exchange less of the disutility of effort for the utility of freedom from pressure*".[88]

Without much reinterpretation, one can read in Leibenstein's work an invitation to use a firm's "*degree of effort*", rather than its control of a large share of output in a marker, as a proxy for significant market power. When competitive pressure is weak, firms do not work as hard to reduce costs, utilize or search new information.[89]

Now, to make things concrete one must specify a test of competitive pressure. Several options exist. In a subsequent paper, Hicks had stressed the relevance of "*subjective factors*" to firm behavior like the willingness to bear risks or the rate of time preference. The downside of Hicks' proposal, however, is that subjective tests are also error prone. A more objective test could consist in shifting focus from market level data towards firm level evidence, and assess whether the firm under investigation is subject to a *change* constraint. Put more simply, the question is: does the firm hustle to move away from its current position? A range of metrics

[88] Harvey Leibenstein, 'Allocative efficiency vs "X-Efficiency"' [1966] American Economic Review Vol. 56 No. 3, 413 (hereafter Leibenstein, 'Allocative Efficiency').

[89] By contrast, monopoly or oligopoly positions are less informative. Competitive pressure is compatible with industry concentration, because as costs declines, "some firms are forced out and fewer firms exist". See Leibenstein, 'Allocative Efficiency' 411.

may be considered like entry and exit choices, R&D expenditures and intensity, rate of product introduction, change in business methods and strategy, capital allocation choices, etc.

2. Bad Conduct: Competing under, against and after Uncertainty?

Antitrust laws proscribe "*bad conduct*" that by purpose or effect harms the competitive process. We discussed previously the ambiguity inherent in categorizing FANG conduct as anti or procompetitive. Again, a substitute concept of "*bad conduct*" seems necessary in disequilibrium markets. What should it be?

In addressing this question, it is useful to recall that the main source of hardship for firms in disequilibrium markets is uncertainty. And it is also useful to observe that uncertainty is in the social interest because it incentivizes firms to compete and innovate.

The normative implication of this is obvious: private incentives may not be aligned with the public interest, and there is a market failure justification for competition enforcement if firm conduct in disequilibrium markets reduces uncertainty. Interestingly, existing competition doctrine accepts to catch uncertainty reducing behavior as a form of "*bad conduct*". For example, established theories of liability in coordinated conduct cases affirm violations of the law when firms jointly reduce market uncertainty.[90]

Figure 7 – Liability Regime for Competition under Uncertainty

From this, if one imagines a continuum with procompetitive conduct on one end and anticompetitive conduct on the other, three types of conduct can be distinguished. First, firms

[90] See Commission Decision, IV/31.370 and 31.446, UK Agricultural Registration Exchange, 17 February 1992, OJ L 68 of 13 March 1992, 19 para 43 and on a low degree see T-35/92 *John Deere Ltd v Commission* [1994] ECLI:EU:T:1994:259, 47–9).

that compete *under* uncertainty take the environment as given and accept failure. They do not "*stick*" to a failing business concept. Instead, they follow *exploration*, *innovation*, and *repositioning* strategies.[91]

Second, firms that compete *against* uncertainty reject failure. This leads them to implement *insulation*, *imitation* and *rent seeking* strategies. Insulation means limiting reversibility of users' participation choices to restrict migration towards better platforms (for example, by placing restrictions on data portability). Imitation means that the firm replicates other firms' successful business concepts, products and services, instead of reinventing itself. Rent seeking denotes attempts to secure regulatory privileges that protect the core business at existing scale. In essence, firms that compete against uncertainty either "*stick*" to their core business or move to proven business concepts from competing firms.

Last, some firms compete *after* uncertainty. Firms in this situation approach equilibrium. As uncertainty fades, these firms enjoy higher ability and incentives to exploit consumers through extractive practices. This concept provides a framework to catch rapacious monetization strategies that exploit locked in users, including ad cluttering, fee introduction in zero price markets or even planned obsolescence.

D. MERGER LAW: KILLER V CATCHUP V RAISER?

Merger control frameworks have been criticized for their failure to prohibit "*killer mergers*". Killer mergers are M&A transactions by platforms that purport to nip in the bud disruptive startups. In his best-selling book *The Master Switch*, Tim Wu talks of a "*Kronos effect*" – referring to "*efforts undertaken by a dominant company to consume its potential successors in*

[91] For more on this, see Petit, 2019 (forthcoming, Oxford University Press).

their infancy" – congenial to information communications industries.[92] Facebook's acquisitions of Instagram and WhatsApp are often described in terms of killer mergers.

Though it is seldom fully articulated, the "*killer merger*" critique rests on the (i) necessary conjecture of an optimal "*but for*" merger world, where the startup becomes a credible competitor to the platform; and on the (ii) ancillary assumption that the buyer has correctly anticipated the "*but for*" world. From this, "*killer merger*" critics advocate investigation of startup acquisitions by dominant digital platforms under a default presumption of anticompetitive harm.[93]

We believe the presumption is inappropriate. For a start, it assumes the rate of occurrence of the "*but for*" merger world as more likely than not. This cannot be right. Our best possible assumption can be that the startup operated in a disequilibrium market (assuming that it has passed the stage of product launch, which is not a given). What this suggests, then, is that little can be predicted of its odds of success.

Besides this highly important point, the presumption mischaracterizes the universe of motives for M&A in disequilibrium markets. In his "*High Growth Handbook*", venture capitalist Elad Gil singles out three types of reasons underpinning M&A transactions with startups: team buys (*eg*, Facebook's acquisition of Drop.io), product buys (*eg*, Google's purchase of Android) and strategic buys (eg, Facebook's acquisition of Instagram).[94] Gil's book is based on qualitative interviews with dozen venture capitalists with stakes in both FANG and FANG disruptive firms, hence it is non-susceptible of pro-FANG bias.[95]

[92] Though reference to Kronos is not very apposite, because in most cases the acquired successor was not home grown.
[93] Tommaso Valletti, 'Après moi, le déluge!, Tech Giants in the Digital Age' (CRA, 5 December 2018) <https://ecp.crai.com/wp-content/uploads/2018/12/Tommaso-Valletti-2018.pdf>.
[94] Elad Gil, *High Growth Handbook: scaling startups from 10 to 10,000 people* (Stripe Press 2018).
[95] Gil, himself, has been a forceful critic of large tech firms. See Eric Johnson, 'Startup adviser Elad Gil is worried that we're losing our ability to be optimistic about tech' (*Recode*, 20 Augustus 2018)

While the concept of strategic buys seems to best match the "*killer mergers*" scenario, Gil discusses the four motives behind such transactions: "[1] *Change the overall market structure;* [2] *Provide [the buyer] with a key, non-reproducible/defensible asset;* [3] *block a competitor from a major action, market position, or the like;* [4] *dramatically change some aspect of your business (cost structure, distribution channel, etc.).*".

Again, harm to competition is only one of them. And in fact, at least two other reasons are compatible with our previously explained theory whereby the acquiring firm is under a change constraint.

The upshot is simple. It would mark a radical introduction of a precautionary principle to subject M&A with startups to a presumption of anticompetitive harm. A more incremental evolution of existing competition law and policy seems instead to formulate a standard of review that tests: (i) whether the target is a competitive force of disruption in a relevant market or in adjacent, neighboring, or complementary markets; and (ii) whether the acquiring firm's incentives are to discard the product or service. As part of this second test, competition law and policy should consider alternative scenarios, including "*raiser*" and "*catchup*" mergers. Raiser mergers occur when the purported merged entity's strategy is to grow the target's product. With benefit of hindsight, we know that this is what happened to Instagram following integration with Facebook. Catchup mergers occur when the purported merged entity's strategy is to keep iron in the fire in competition with rival companies. Again, looking in the rearview mirror, we know that Microsoft's $6.3 billion acquisition aQuantive in 2007 was a (failed) attempt to compete with Google in the online advertisement market.

Of course, competition policy makers can claim that we have been historically wrong, and that should be more cautious in the face of uncertainty. A valid justification could be the existence

<https://www.recode.net/2018/8/20/17757412/elad-gil-startup-handbook-scaling-growth-optimism-pessimism-culture-kara-swisher-decode-podcast> accessed 12 April 2019.

of tail risks of anticompetitive effects due to killer mergers with startups. But any policy adjustment of this kind would need to be based on evidence (i) that startup independent exit represents a class of low probability, high social benefit event; (ii) that the social benefits worthy of merger control protection outweigh the private benefits to startup founders and the social costs of a decrease in exit opportunities for other entrepreneurs; and (iii) that *ex ante* merger control intervention is necessary, short a less costly *ex post* alternative that would consist in remedying past mergers on the basis of subsequent market developments. Are these conditions fulfilled? Until now, we have little evidence that this is the case. A paper that is often cited in competition policy circles claims that 6.4% of all acquisitions in its dataset are killer acquisitions.[96] Perhaps, this meets the low probability prong of the first condition. But this leaves all others unanswered.

Conclusion

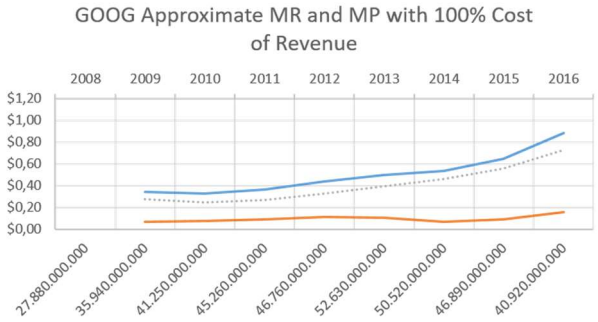
This paper has presented the rudiments of a theory of disequilibrium competition with uncertainty. This framework may assist antitrust and regulatory decision makers in assessing whether FANG firms and many others' strategies are procompetitive or anticompetitive. True, our framework can be criticized as incomplete, abstract and perhaps even unpractical. And yet: because knowledge is cumulative in nature, we hope that our preliminary theory, with both its strengths and weaknesses, will assist competition and regulatory agencies in their much necessary ongoing efforts to rethink the application of competition law to digital industries.

[96] Colleen Cunningham, Florian Ederer and Song Ma, 'Killer Acquisitions' (SSRN, 28 August 2018). <https://ssrn.com/abstract=3241707>.

Appendix 1

FANG	Year	Output (Q)	Price (P) ie R/Q or Marginal Benefit	Marginal Cost ie "Cost of Revenue"/Q	Marginal Profit
AMZN	1997	1.500.000	$98,51	$79,30	$19,21
	1998	6.200.000	$98,39	$76,80	$21,59
	1999	14.000.000	$117,13	$96,37	$20,76
	2000	20.000.000	$138,10	$105,31	$32,79
	2001	25.000.000	$124,90	$92,96	$31,94
	2002	n/a	n/a	n/a	n/a
	2003	40.000.000	$131,59	$100,16	$31,43
	2004	n/a	n/a	n/a	n/a
	2005	n/a	n/a	n/a	n/a
	2006	n/a	n/a	n/a	n/a
	2007	76.000.000	$195,20	$151,08	$44,12
	2008	88.000.000	$217,80	$169,27	$48,52
	2009	105.000.000	$233,42	$180,74	$52,68
	2010	130.000.000	$263,11	$204,32	$58,79
	2011	164.000.000	$293,15	$227,37	$65,79
	2012	200.000.000	$305,47	$229,86	$75,61
	2013	237.000.000	$314,14	$228,61	$85,53
	2014	270.000.000	$312,39	$232,41	$79,97
	2015	304.000.000	$326,27	$235,69	$90,58
GOOG	2008	27.880.000.000	$0,38	$0,15	$0,23
	2009	35.940.000.000	$0,31	$0,12	$0,19
	2010	41.250.000.000	$0,34	$0,12	$0,22
	2011	45.260.000.000	$0,39	$0,13	$0,25
	2012	46.760.000.000	$0,49	$0,17	$0,32
	2013	52.630.000.000	$0,51	$0,19	$0,32
	2014	50.520.000.000	$0,56	$0,22	$0,34
	2015	46.890.000.000	$0,74	$0,28	$0,46
	2016	40.920.000.000	$1,04	$0,40	$0,63
NTFL	2012	99.600.000	$36,24	$26,63	$9,61
	2013	123.490.000	$35,42	$25,24	$10,18
	2014	148.240.000	$37,13	$25,32	$11,82
	2015	171.620.000	$39,50	$26,75	$12,75
	2016	191.030.000	$46,23	$31,57	$14,66
	2017	210.290.000	$55,60	$36,42	$19,18
FB	2012	2.280.000.000	$2,23	$0,60	$1,63
	2013	2.849.000.000	$2,76	$0,66	$2,10
	2014	3.385.000.000	$3,68	$0,64	$3,05
	2015	3.949.000.000	$4,54	$0,73	$3,81
	2016	4.624.000.000	$5,98	$0,82	$5,16
	2017	5.378.000.000	$7,56	$1,01	$6,54

Appendix 2


GOOG Approximate MR and MP with 100% Cost of Revenue
2008
2009
2010
2011
2012
2013
2014
2015
2016
$1,20
$1,00
$0,80
$0,60
$0,40
$0,20
$0,00
27.880.000.000
35.940.000.000
41.250.000.000
45.260.000.000
46.760.000.000
52.630.000.000
50.520.000.000
46.890.000.000
40.920.000.000

Appendix 3

Figure 8 – Alternative Discontinuous Demand Curve for Network Good