

Economic Aspects of Biometrics

Jonathan Cave
University of Warwick
December 2004

DRAFT

Economic Aspects of Biometrics.....	i
1 Introduction: the economics of identity	1
1.1 Identity	1
1.2 Biometrics	3
1.3 Additional comments	6
2 Framework	8
2.1 Logical structure	9
2.2 Stakeholders and processes.....	9
3 General economic aspects of biometrics.....	11
3.1 The ‘Value Chain’	11
3.2 Errors and uncertainties	13
3.3 Costs and benefits	17
4 Market situation	21
4.1 Market shares by region.....	21
4.2 Market shares by application area.....	22
4.3 Market shares by technology	22
4.4 Barriers to growth	24
4.5 Structure.....	24
4.6 Sectors.....	27
Government and other public sector	27
Physical access control	28
Retail and other payments.....	28
Telecommunications	29
Financial services.....	29
Health.....	30
IT, etc.....	30
Transportation.....	30
5 Economic outcomes	32
5.1 Evolution Scenario Issues	33
Network effects.....	36
Public goods.....	37
Evolution and intellectual property.....	37
Security	37
Trust.....	39
Privacy	39
Commerce	40
Work	40
Societal discourse.....	40
5.2 Competitiveness and efficiency	41
5.3 Equity and distribution.....	41
6 The role of policy	41
6.1 Potential areas for policy intervention	42
6.2 Policy levers.....	43
Regulation.....	43
7 The economic future of biometrics	45
7.1 The market for identity	45
7.2 The impact of identity on market outcomes	47
8 References.....	48
9 Annex: A simple model of biometric adoption.....	50

9.1	Description of the one-sided model	50
9.2	Equilibrium behaviour	50
9.3	A two-sided version	51
9.4	An incomplete-information model.....	51

DRAFT

1 Introduction: the economics of identity

1.1 Identity

Economics is concerned with transactions. In many cases, these transactions rely on a degree of *trust* among the parties. This is particularly true when payment or fulfilment occur ‘remotely’ – at other times, in other places or through goods and services whose authenticity and quality cannot immediately and reliably be verified. One way of providing the needed assurance is through the identity of the transacting parties – they need to know with whom they are dealing. This establishes a fixed base for liability – the individual making a commitment accepts future liability for its delivery and can seek recourse from the other party. In addition to providing assurance that the other party can trust the identified party, providing identification information can act as a signal or token of goodwill.

A second role of verified individual identity is to serve as a token for specific characteristics of individual demand or supply. The matching of demands and supplies for individualised (not mass) goods and services relies on possibly rich and hard-to-verify data. *Personalisation* of these data provides convenient summaries that allow markets to operate more efficiently.

A third role for identity is as an *indexing* device – shorthand for e.g. a transaction history or other bodies of data.

A fourth economic aspect of identity takes the form of a *capital* asset. For instance, an individual’s credit rating is a durable asset that controls the extent and cost of access to financial capital. It is formed through investment and is subject to depreciation. Unlike financial capital, however, ownership may be split or diffused – credit rating agencies may maintain separate accounts and different amounts of information or control.

In this context identity can be established (to the satisfaction of the parties) by a range of data. These include¹:

Non-biometric	identifying knowledge	name, id code	tokens
Biometric	natural characteristics	appearance	behaviour
	acquired characteristics	bio-dynamics	

In practice, each method offers different degrees of exactness (probability of false positives or false negatives) and specialisation, depending on context, intended purpose, etc. For instance, some names are more common than others; some behaviour can be copied easily; and some identifying knowledge can be confused or forgotten. For this reason, identification (the linking of a person to data relating to them) typically relies on a combination of the above.

The benefits and disbenefits of identity in economic interactions depend on the balance among the uses and the means of identification chosen. For this purpose, it is appropriate to distinguish the transactional status of an assertion of identity. For

¹ Further discussion of the characteristics of these different forms of identification can be found at: <http://www.anu.edu.au/people/Roger.Clarke/DV/HumanID.html>

many economic and other purposes, it is neither necessary nor appropriate to completely identify a party to a transaction with a specific living individual. For instance, in a cash transaction a party need only be identified as having the right to exchange goods and services for money – the seller must have title to the goods and the buyer must have the cash. If either the purchase or the price are uncertain or contingent, greater identity may be needed: for instance, the buyer may need to identify himself as someone with assets sufficient to ‘cover’ a non-cash transaction, may need to demonstrate his eligibility to buy the object in question (e.g. through proof of age, etc.) or may need to certify that the purchase will or will not be used in certain ways. By the same token, the seller may need to establish the provenance of the goods offered for sale or to certify quality, origin, etc. This may involve retrospective identity (e.g. that a device was made by a competent mechanic or that medical advice comes from a qualified professional) or prospective identity (e.g. that the individual can be sought out in future if the goods prove unsatisfactory). These require different degrees of identity: in some cases it is sufficient to prove that the party belongs to a specified class (e.g. physicians) or does not belong to a specified class (e.g. bankrupts or felons). In other cases (such as prospective assurance) it is necessary to identify a specific individual or their legally designated representative.

In addition to the role of identity in economic transactions (the demand for identity), it is increasingly necessary to consider the economics of what has come to be called identity management (the supply of identity). The scope is not limited to economic transactions: it applies equally to the supply of identity technologies, data and other goods and services in e.g. citizen-government interactions (voting, benefits claims, travel, etc.).

Both the supply and demand sides are strongly affected by characteristics that have problematic implications for ‘conventional’ economic analysis. One is *complementarity* – identity is complementary to a range of economic transactions and is thus naturally bundled with them – but complementary goods and services may be associated with economic instability or even non-existence of equilibrium. A second group of features are network externalities: put simply, verifiable identity establishes links between a person and others. The benefits of this linkage are not wholly captured by the provider of identity – whether the person themselves or others who create or demand identification – so identity may be under-supplied from a societal point of view. By the same token, the costs or disbenefits may not fall on those with choices to make, leading to oversupply.

Economies of scale and interoperability – each individual may benefit from having a single approach to identity – fewer cards to carry or lose, fewer passwords to remember, etc. A single identification system carries risks as well. In particular, it prevents compartmentalisation of identity that may be useful in limiting risks (e.g. of fraud or identity theft). There are also collective advantages in having identification systems that span many people. Some are purely cost-driven; to the extent that such systems have large fixed costs, it is efficient to spread them as widely as possible. Others are concerned with information security and risk-management – it is expensive to protect a system or a data repository; these fixed costs may lead to lower levels of protection for small or disaggregated systems.

Identity offers private benefits, but the greater interaction allowed by assured identity and the resulting higher levels of trust are public goods. Thus any consideration of biometrics or other forms of identity should consider the wider

societal impact and the twin difficulties of measuring the appropriate level of identification and of ensuring a feasible allocation of costs and benefits.

Identity is tied to the individual, and not necessarily to the individual in the context of a particular transaction. There may thus be some adverse consequences of ‘creep’ whereby identification for one purpose (e.g. voter registration, health care provision or food purchases) is reused for other purposes (e.g. credit or insurance rating). In addition, where the data affect the quality of the ‘match’ between an individual and various transactions, the other party may have incentives to prevent the reuse of the data by potential competitors – or even by the individual themselves. Among the consequences of such one-sided incentives to limit the transferability of identity may be an emphasis on relation-specific identity data and a proliferation of separated identity systems. While this can enhance privacy by compartmentalising data, it can at the same time restrain competition, and limit the benefits of ‘reputation effects.’

1.2 Biometrics

Biometric technologies are means of identification based on one or more physical and difficult to forge characteristics. They offer the promise of enhanced confidence in identification - connections between an individual and data that represent that individual’s identity. However, individuals in modern society no longer necessarily have just one identity and the relevant identity may not be tied directly to a biological person.

Biological identification is very old – personal recognition – either directly or through witness identification – is part of the root structure of commerce and law. But these techniques are informal and/or subjective and are difficult to ‘store’ and transmit. Biometrics involves the use of recorded measurements in lieu of direct physical evidence or subjective matching. The range of relevant measurements includes:

- personal appearance: descriptions of height, weight, gender, race, hair and eye colour, scars and other physical characteristics; glasses; etc. These measurements are often accompanied by templates in the form of photographs for easy comparison to the individual
- behaviour: voice characteristics, speech patterns, handicaps of movement, etc. sometimes accompanied by motion picture templates
- bio-dynamics: signatures, keystroke dynamics, statistical voiceprint analysis, etc.
- natural physical characteristics: dental and skeletal records; fingerprints; hand geometry; retinal and iris images; DNA; etc.
- acquired physical characteristics: tattoos; barcodes; implanted devices (micro-chips, RFID tags); rings; brands, etc.

These may change over time and may be more or less easy to change ‘on purpose.’

The biometric identification of an individual may occur as a result of a claim by them or a demand by the provider of a service or the operator of a facility. It may also be more or less ‘conscious’ on either part: the individual may ask to be identified and be aware of the scrutiny and its result, or may be passively identified (e.g. through the use of facial recognition software on surveillance devices). Similarly, an identification system may be more or less automated and records of identification may not be

retained, or retained only in attenuated form (e.g. recording authorised access but not identity in the case of voting registers).

The evaluation of biometrics and the consideration of the likely future of the industry and economic effects should thus take account of the continued existence of alternative means of identification.

Biometrics offers higher assurance, but it is not uniformly and perfectly accurate, and the identification comes at a price. As discussed in Section 3 below, biometric technologies provide different combinations of a range of costs and benefits. These fall on different stakeholders and are differently perceived by them. The actual level and distribution of benefits and costs reflects dependencies among different parts of the value chain and the responses of individuals. In particular, adoption and implementation decisions at any moment are influenced both by prior decisions of others and by expectations of the future. This introduces a degree of path dependence.

From the standpoint of overall economic development and impact, then, it is worth recording that biometric technologies increase the accuracy with which an individual claiming identity can be matched to a pre-recorded template. As a result, the accuracy of the system relies heavily on accurate enrolment and the integrity with which templates are stored and retrieved. This creates vulnerabilities in terms of both accident and attack.

Even if biometrics is 'better' than alternatives, it does not follow that the certainty offered by a biometric solution is cost-effective or 'optimal.' Increasing the certainty of biometric identification is costly, whether achieved by more intensive template encoding or by combined biometric identification. From the standpoint of users, the quality of a particular implementation may thus be too high to be cost-effective. Indeed, since not all applications require exact identification of the individual or a secure link to extensive historical or other data, some implementations – regardless of monetary cost – may be too strong for the purpose for which they are employed. This may come about because of e.g. privacy concerns that reduce the acceptance of biometrics by those whose identity is verified, or through contingent liability or legal restraints on the verifier's collection of information. It may even arise because the permissible accuracy of identification is bounded above as well as below – an example would be voter identification, in which it is essential to establish that the individual is eligible and has not already voted, but equally essential not to further identify him or her.

This leads on to the question of which (combination of) biometric solutions will (and which should) emerge from market forces, planned adoption decisions and the lessons of experience. If there were no questions of interoperability, it would be possible to 'fit horses to courses' and thus to adopt for each application area a biometric (or other) identity management solution that offered an efficient balance among costs, accuracy and other aspects of quality. But identity management systems do need to interoperate and the burden for individuals and systems of multiple identity systems is not negligible. Thus, even though the 'optimal' biometric solution for one application area may differ from that for another, there are strong pressures to adopt 'compromise' or second-best solutions that imperfectly span a number of areas.

A further consideration is that – to the extent that biometric solutions provide cheaper, stronger and/or faster means of establishing identity, they may 'tilt the

playing field' against those who cannot or will not participate. This may impair equity – for instance if the vast majority of users migrate to a biometric solution, provision of alternative, non-biometric channels may become uneconomic and be discontinued, excluding the minority or increasing their costs of participation. Alternatively, those who might have e.g. privacy concerns about a particular implementation may not be able freely to opt out without losing access to goods and services (or societal interactions) to which they are entitled. The very comprehensiveness of biometric identification and the perceived possibility of extensive data mining may cause certain groups consciously to opt out – not only from biometrics but also from the interactions to which they hold the key. This may even have aggregate efficiency consequences, disadvantaging those on the 'inside' as well as those on the 'outside' as the example of eParticipation (electronic voting) suggests. As a general matter, a system whose benefits depend on user interactions will be damaged by changes that raise barriers among users.

Because there are various biometric technologies offering different characteristics, there is competition among these technologies reflecting such things as lifetime cost, reliability, acceptance, accuracy, etc. The prevailing technology or combination of technologies, and the uses to which they are put, will depend on these characteristics. But it is not a head-to-head competition among different technologies for accomplishing a single task. Not only are there many aspects of performance (so that one is unlikely to dominate all others) but the technologies are embedded in a value chain and a range of different devices in ways that complicate the competition.

For example, in some regions fingerprints or iris scans are deemed invasive, while in others they are well accepted; some technologies (e.g. retinal scans) require active and careful user cooperation, while others (iris scans or fingerprints) are more fault tolerant; some are easier to 'spoof' than others, and so on. Thus it is unlikely that a single solution will work effectively for all applications – or even for a single application applied on a wide enough scale. Some parts of an integrated identity management solution can be modularised to avoid technological dependence – for instance, with a suitable interface standard PCs can be made to work with fingerprint, hand geometry, voice recognition and/or iris software. Various types of sensor can be made to work with an identification algorithm for a given type of biometric information. But other parts of the system are more difficult to 'open' to a range of technologies.

If a biometric solution is being adopted by a single large organisation, it is reasonable to suppose that a period of acceptance testing (which may involve competition among a range of approaches) will be followed by widespread adoption of a single or harmonised system. But in competitive market environments this need not happen; even when it does it may not be desirable.

For an example, consider two implementations of a given identification algorithm for use in a retail environment. In the first, templates are stored and matching performed centrally. This offers the possibility of configuration control, reduced vulnerability to spoofing, relatively low 'point of sale' costs to the seller (since the retail outlet need only be equipped with sensors and communications equipment) and the buyer (since it is unnecessary to carry a copy of one's personal template). But it may be vulnerable to interception or to loss of functionality in the event of systems or communications failure – at the least; the seller would need some form of encryption and decryption capability. By contrast, a solution that relied on local matching

between a tamper-proof template copy and the individual has complementary strengths and weaknesses. But the main point is that different solutions would appeal to different sellers (e.g. large vs. small outlets) and different groups of buyers (e.g. children vs. adults). So the adoption of a particular solution would necessarily ‘tilt the playing field.’ If sellers could choose which to adopt, they might make different choices; if they could not, they would enjoy different degrees of functionality.

There are other important differences in functionality that will affect both costs and demand – and thus the uptake and use of different technologies.

- One advantage of certification or password identity systems is that they can be revoked or reissued if lost, forgotten or compromised. Not all biometrics offer the possibility of revocation and reenrolment – or may only offer them at some costs in terms of the security provided. For instance, a relatively secure biometric identification can be based on a random sample of an iris scan – the sample can be stored in compact form even though the full template and the original are quite complex. If the compact version is compromised, a new sample could be generated. This might be much harder with voice, face or fingerprint.
- For some biometrics, the individual might diverge from their template through aging or other changes. Template aging may add significantly to the lifetime cost of the system.
- Some biometrics (notably retinal scans and possibly DNA) may carry additional information (e.g. related to health) whose reuse might raise economic as well as privacy and equity concerns.

1.3 Additional comments

The development of the sector has been deeply affected by the heightened security consciousness of the past few years, by heightened economic concerns related to globalisation and intellectual property rights (especially software patent monopolies, by the increasing strategic importance of standardisation and by the rapid proliferation of eGovernment initiatives (including taxation, health, education, benefits entitlement, crime control and migration. At the same time, the expanded identification possibilities are very attractive in addressing a range of other potential policy issues: The evolution of policy and practice is driven by a wide range of stakeholders. The public sector is represented by regional, national and supranational governments, by a range of ministerial players and by specific policy areas from service delivery to security to procurement to R&D. The private sector is represented by specific (networks of) firms who develop, supply, integrate and demand biometrics in a range of sectors: transportation, security, telecommunications, health, retail commerce, IT hardware and software, etc. The civil sector is represented by those who must use and/or rely on enhanced identification: people acting as customers, workers, citizens, taxpayers and participants in societal discourse. Not all – indeed, not most – of them are motivated by economic considerations alone. But economic considerations are important to all of them and economic forces will influence the success or otherwise of a range of policy options.

One area of particular concern relates to knowledge about technology. Biometrics is a technology that rests on both formal (R&D) and informal (learning-by-doing) innovation. All such ‘knowledge capital’ produces spillover effects. In some sense, the use of knowledge by one party does not reduce the amount available for use by

others – it is ‘non-rival.’ The traditional way of measuring the ‘right amount and kind’ of knowledge and of ensuring that creative effort is rewarded, is to create and enforce a property right by which the ‘owners’ of a technology may exclude others from using it or charge them for the privilege. This, in effect, goes down a private goods route by creating a (limited and temporary) monopoly in exchange for open publication of the R&D. The inefficiency caused by this distortion of competition is – in principle – justified by the increased flow of innovations and the competition among ideas. But there is no guarantee that the two effects will balance or that the use (or otherwise) of the technology will not produce other, external, impacts (e.g. on privacy) in areas where it is not practicable to define property rights. An alternative is to pursue a private goods route by pursuing an open source policy whereby the rights to the research are publicly held. In this model, exemplified by the General Public Licence or ‘copyleft’ provisions, use rights are granted freely and even derivative innovations may be bound to the public domain. Economic returns may be sought in selling related goods and services or in selling enhanced versions of the compiled (not source) code bundled with other applications.

This has been thought to raise security concerns, since it is not obvious who (if anyone) ‘owns’ the liability for flaws in the technology or its implementation. On the other hand, the broad involvement of an active development community competing on the merits could produce a diversity of approaches to security challenges and could at least expose potential problems as easily to those interested in overcoming them as to those bent on exploiting them. On the basis of empirical evidence, open-source systems seem to be at least as secure as proprietary systems and sometimes much more secure².

A related point is that biometrics is rarely used as a stand-alone application in a single context. Instead, it is almost always used in conjunction with other applications and between or among a number of parties. As a result, interoperability is key, but not suitable for proprietary ownership. Instead, public good – and the health of economic competition – are enhanced by the adoption of common, non-proprietary and non-discriminatory standards. These are not software algorithms or middleware *per se*, but set the ground rules for their interoperation. It is worth noting that adoption of a common standard can enormously stimulate market competitiveness, as the GSM example from mobile telephony shows. In the area of biometrics, a common interface standard (BioAPI) was developed by a range of industry stakeholders. Interestingly (and perhaps characteristically), it was shortly followed by a proprietary standard tied to the dominant personal computer operating system (BAPI). This strategy resembled the ‘standards wars’ in Internal HTML protocols, and many observers were braced for a similarly monopoly-friendly result. However, it appears that international pressure has led to the development of a translation layer between the two. To the extent that the differences are justified by differences in application (there is precedent for this) this is a reasonable solution – but the enormous installed base associated with Windows means that a degree of regulatory scrutiny will probably be necessary – especially if governments are to remain free to pursue non-proprietary software whilst at the same time using biometrics in eGovernment applications.

A final prefatory remark concerns the relation between biometrics and border controls. This review concentrates on economic impacts, so the security and civil

² Compulsory licensing provides a limited form of ‘third way’ but is costly and legally complex to operate.

liberties aspects are beyond its scope. However, it does note three main areas in which biometric border control applications are relevant. The first is labour mobility; the competitiveness of the European economy depends on factor mobility. Until recently, labour mobility had been fairly limited, but recent key shortages of specific skills (or of workers in specific age ranges) have led to a reconsideration of the need for economic migrants. To the extent that these workers earn – or expect to earn – real wages in excess of those in their countries of origin, migration is an inevitable (and a desirable) result. But such migration may not be (for practical and ethical reasons) finely tuned to labour needs and may produce its own external effects. It is significant in this respect that ‘people-smuggling’ enterprises have arisen to capture the differential between origin and destination income *expectations*. If biometrics can reduce the gap, it can make a contribution to reducing the misery this trade creates. But it also enables other, less-altruistic policies and is not a panacea. The second economic aspect, which follows from the first, is the use of biometrics to control participation in the public economics of the country: entitlement to public goods and benefits, payment of taxes and having a say in how expenditures and transfers should be targeted. The relation to border controls is two-fold. First, no border control is impermeable, and biometrics inside the country not only strengthens border controls but reduces the ‘draw’ of illegal immigration and goes some way towards eliminating an adverse disparity between the treatment of legal and illegal immigrants. Second, much of the (perceived) pressure on specific classes of public service and much leakage in tax collection are associated with illegal immigration. The third aspect is the sheer size of border control, passport and identity card schemes. These are massive public procurement projects and, like all such (especially in the IT area), carry grave risk of wasting money and/or stifling innovation – thus they must be carefully considered in light of what is known about value for money in public procurement and about the connection between public procurement and innovation.

- Everyday life:
 - Individuals: inclusion and exclusion in (tele) work, consumption, financial transactions, benefits, transportation, utilities
 - Networking aspects (identity as a network linkage, neighbourhoods, indirect effects)
 - Acceptance: intrusiveness, reliability, data-reuse, etc. – levels and dynamics of uptake, efficiency of use, risk allocation, shifting and management, bypass, convergence and creep, divergence and divides, opting-out vs. shutting-out
 - Exclusion/inclusion in markets: small firms, complementarity and lock-in, security levels inside and outside (esp. substitution of one form of security for another or ‘adhesive contracts’ to share/transfer information, rents and power)
 - Consumer protection issues.

2 Framework

The analytic structure begins with a logical framework summary of the system and influences linking technology and economic impacts. It then considers: a ‘process map’ of principal actors; constraints and possibilities determining their activity; a

market overview (including the influence of uncertainty); impacts on key issues and the stakeholders affected by them; and broader, longer-term outcomes in terms of cross-cutting themes. The conclusions will make links to other domains of analysis (e.g. technology, law, etc.) and draws appropriate policy conclusions.

2.1 Logical structure

- Biometrics represents (part of) a general-purpose (disruptive) technology.
- Its adoption follows a nonlinear path (perhaps even catastrophic)
- It is subject to costly diffusion and network externalities
- It has spillovers – the creation of detailed data records, tighter linkages among activities, etc.
- There are a range of logical processes and choices behind its deployment:
 - Technology development, IP ownership and deployment
 - Demand-pull v. supply push
 - Administration, civil (customer, citizen, etc.) and business players.
 - Who chooses what, and subject to what constraints? (businesses or governments can ‘force’ use of BM, but constrain each other. Citizens can (?) choose which form of identity to provide, but obtain different benefits from different forms.)
- Various technologies
- Ownership: of technologies, data, etc.
- Technological convergence can cross ‘functional’ lines (digital signatures as ‘more than’ holographic ones; Unified ID cards (e.g. DL + Credit card)
- Mission creep mediated by acceptance and uptake, and by possibility of abuse.
- Decentralisation changes cost incidence, perceived and actual vulnerability, path-dependent growth.
- Information collection and integrity.

2.2 Stakeholders and processes

The ‘process map’ section provides *context*: an overview of recent developments, trends and bottlenecks to the development and deployment of biometrics from an economic perspective, including economically-relevant policy developments and industry initiatives (esp. self-regulation and standardisation). It will then identify *actors*: those in (or linked to) the biometrics sector, complements (esp. information processing, encryption, storage, etc.); substitute forms of identification; and the rest of the supply chain linking biometrics to economic activity. The process map will further describe the parties’ *interactions*, including impacts on regional and macro-economic development, albeit at a fairly simple level. It will also describe outside drivers including non-economic concerns: security, the need for higher levels of identity verification in eGovernment, eDemocracy, eHealth services, etc. In addition the role of biometrics in managing issues related to labour and personal mobility will be considered.

The major stakeholders can be divided among four major societal groupings.

- Administrative sector
 - Providers of government services or benefits linked to identity (demand)
 - Law enforcement, tax etc. authorities (regulation)
 - Data protection, privacy, fraud, etc. regulators (regulation, template supply)
 - Competition and sectoral regulators (regulation)
- Private sector
 - Suppliers of biometrics technologies
 - Commercial (retail) sellers (demand)
 - Financial services (supply)
- Civil/public sector
 - Customers for private sector goods and services (payment, authorisation)
 - Taxpayers, etc.
 - Benefits claimants (possibly including migrants)
 - Borrowers, lenders, savers, etc. (authentication)
 - Members of civil society
- Defectors
 - Economic criminals
 - Those who 'opt out'
 - Those who wish to disrupt the system

Figure 1 shows a somewhat simplified 'top-level model' showing the main stakeholders and their interactions. Such a model is useful to provide a 'sanity check' to ensure that a scenario is not missing any potentially important logical connections. It is not an economic model *per se*.

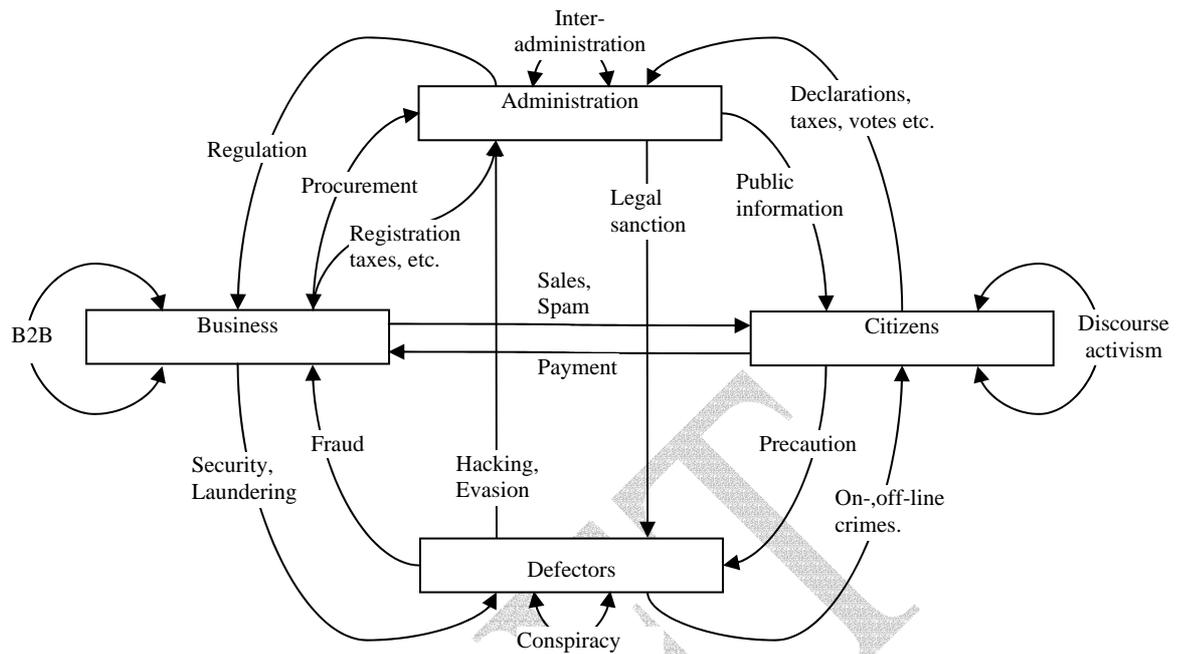


Figure 1: overall schematic model

3 General economic aspects of biometrics

3.1 The 'Value Chain'

A biometrics solution involves a number of players along a 'value chain.' Such chains differ slightly with application area and other factors, but a fairly generic version is shown in Figure 2. The top part shows the internal application structure and the bottom part the market players. Some obvious links between the parts are not shown.

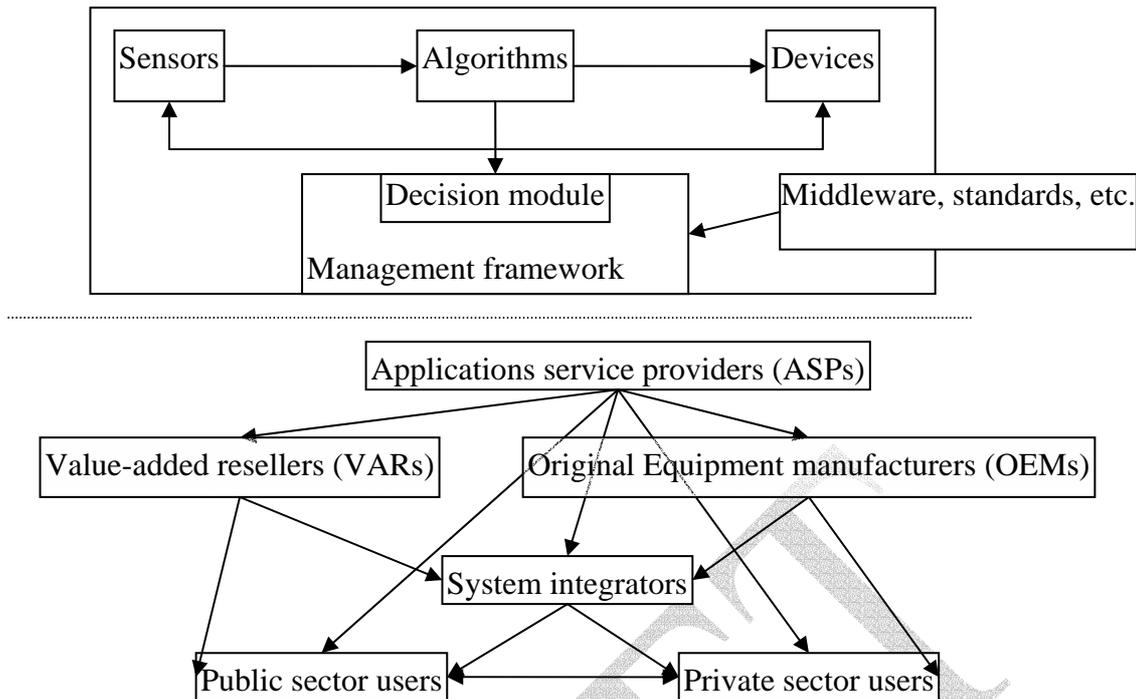


Figure 2: supply-side market components and players

As the application matures, the hardware (sensors and devices) will become increasingly cheap, interoperable and commoditised³. Algorithms will remain proprietary and distinctive and will continue to improve as use intensifies and spreads. Thus IPR in this area will remain profitable. By contrast, the software involved in middleware, which mediates both functionality and interoperability, is likely to be convergent, less profitable and ultimately to domain of open-source solutions and/or compatible free software.

ASPs will be the dominant firms in the supply side during the growth phase – initially by providing solutions but ultimately through providing support to users and inputs to the ‘intermediary’ layers below. As the market matures, they may be bought out by integrators and other intermediaries. The VAR and OEMs, as in other ICT markets, will provide important transitional competition to the ASPs and integrators. Ultimately, the market will probably belong to integrators, whether specialised to security (e.g. RSA) or diversified across ICT solution provision (e.g. EDS). Their relationships – with each other and with clients - are likely to be characterised by strategic and/or collusive partnerships. They are likely to acquire successful ASP and others after market shakeout has concluded. Ultimately, biometrics may cease to be a separate sector and become wholly subsumed in markets for associated technology (e.g. PCs) and integrated ICT and/or security solutions.

The diagram also distinguishes public and private sector users because they differ in terms of risk aversion, motivations, sensitivity to competitive forces and ability to await developments. Typically, experimental government contracts drive the first phase of many new technologies. This is particularly true of biometrics as its advent has coincided with heightened security provisions and with rapid development of eGovernment and eParticipation initiatives. However, there are limitations to the

³ Acuity Market Intelligence, http://acuity-mi.com/Industry_Evolution.html

public sector's role as a 'launching customer.' These include the embedded risk aversion (and even deeper uncertainty aversion) of public sector procurement officers, the well-known difficulties in defining appropriate requirements for public procurement, the confounding effects of departmental stovepipes and auditing practices that are less than friendly to innovation and the impact of procurement regulations – though this is in the process of change in the European context as the new procurement directives expand the scope of procurement as a tool for driving innovation and stimulating private sector R&D investment.

Private sector clients may be more willing to take risks provided they are well understood. Constraining factors to date has been the legal and regulatory status of biometric solutions and competition among different technologies. They interact with the public sector both as followers of the launching customer and as suppliers of an increasing range of goods and services to the public sector.

Ultimately, however, both governments and private sector clients who adopt these technologies depend on end users – those whose identity the biometrics authenticates. The cooperation and willingness of end users to accept the technologies will determine their eventual impact. If the applications are not accepted or not well understood, progress can be slowed or reversed. For instance, if users do not make appropriate use of the technologies, error rates and associated costs may rise to the point where the application is no longer cost effective. If the users have a choice they may vote with their feet and the initial investment may be lost.

3.2 Errors and uncertainties

No biometric identification system is perfect. The different types of errors and their interaction with the ways biometrics are applied, give rise to a set of criteria that in turn influence market uptake and impact.

Biometric identification (to date) has been used in one of three ways:

- To verify a user's claim of his or her identity. This involves an identifier that connects the claim to a stored template. The template is designed to facilitate comparisons with measured characteristics – this means highlighting aspects that are easy to measure, store and/or transmit. This saves on computation (and possibly communication and/or storage) cost for the system owner.
- To identify a user. This involves taking a measurement and scanning for a match against a large database of stored templates. This does not involve making a claim or producing a local template, so involves less user cost.
- To check for multiple identities. In cases of identity fraud, it may be useful to maintain multiple identities. To prevent fraud (or misidentification), multiple identity screens check the data of a user wishing to enrol against stored templates (as above) to ensure that there is no match.

In each case, it is necessary to measure and codify user data and to match these against prior codified measurements. There are a variety of accidental, contextual or deliberate ways in which such measurements may vary, so the matching algorithms have to be fault-tolerant. More precisely, the error bounds of individual identity need to be set appropriately. If they are set too tightly, the result will be a large number of 'false negatives' where an individual's measurements differ from their own stored template and they will be wrongly rejected. If they are set too wide, the result will be

too many false positives, in which an individual is matched to someone else's template. Figure 3 shows the variation.

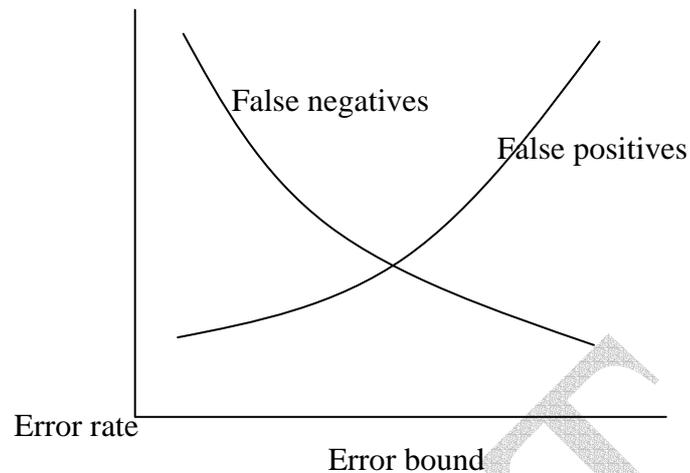


Figure 3: Identification errors

The trade-off between the two determines the usefulness and security of the system. It depends in part on the mode of identification. Even a small false positive rate for a system used for identification (rather than verification) will produce large numbers of false matches. If the identification system is used to screen for exclusion (e.g. terrorists, hooligans or fraudsters), the consequences may include compensation claims or even a tendency to discount the match results – and thus relax vigilance.

Where both parties have common interests it can be a relatively simple to determine the optimal level of security; this could even apply to choices among systems. Indeed, in some applications the error bound can be personalised to the individual or the system – for example by varying the precision with which the template is considered or the biometric recorded.

Further optimisation can be sought through a ‘mixed strategy’ approach to identification. In any application, there are costs to both parties attached to false positives and false negatives, and the optimal decision balances the *expected* costs on both sides. For example, it is certainly possible in identifying criminals to minimise false positives (e.g. by requiring the highest level of identification – for example, DNA evidence combined with positive witness identification) or to minimise false negatives (e.g. by requiring the individual to prove they were not present or culpable). But the first of these weakens punishment and the second punishes too many innocents. Any system, even one apparently based on absolutes of identification, in practice probabilistic. If the risks have a known relationship to the characteristics of the system, the error bounds and other implementation details can be set through consideration of the *ex post* probability that the person is who they claim to be⁴. This can already be seen in practice in e.g. IT applications; high-security applications for remote access to firewalled systems have relatively tight error bounds because the perceived cost of repeat attempts is less than the perceived risk of penetration.

It therefore follows that at the root of the biometric choice and implementation question lies a decision problem under incomplete information. In situations where the interests of the parties coincide (e.g. in providing access to corporate IT systems),

⁴ Sequential equilibrium!

the choice then rests on the feasible risk combinations. Because the risks are not known in advance, the choice must take account of or adapt to context, experience and human performance factors to which the error ratios are very sensitive. Empirically, these factors include sensitive demographic details such as race, gender, age⁵, disability, willingness to cooperate, cognitive ability and so on. One issue raised is fairness – whether a given implementation treats all users equitably and provides equivalent access to all those entitled. Another concerns the validity of test or pilot results for larger or different populations – indeed, testing standards in this area have only recently been developed. This arises as an issue in comparing biometrics used for passports or social benefit claims (which apply to different subsets of the population) to mandatory identity cards.

A second point is that the error probabilities are not fixed simply by the technical characteristics of the system, but also reflect the way in which it is used. This is obviously true in terms of the costs of errors – the loss function for a biometric system used to allow nightclub access obviously differs from that for the same system used for online banking. The error rates themselves may vary with application – this is particularly true of those systems that depend on “something you do” such as voice identification or signature dynamics, which may be affected by e.g. stress or ambient noise, or systems where measurements must be taken against a wide range of ambient conditions. But perhaps most crucially, the graph above lumps together ‘objective’ (technical or computational) error with ‘subjective’ error resulting from deliberate attempts to assert or avoid identification. For instance, the use of digitised face biometrics in the USVISIT programme is likely to attract different types and levels of risk than the same system used as part of the planned UK identity card scheme.

As mentioned, the actual market fate of biometric solutions depends on choices under perceived uncertainty. It is well-known⁶ that people tend to misperceive very high or very low probabilities, to behave differently depending on their reference position⁷ and to conflate the likelihood of a breakdown and the severity of the consequences. Thus, although the simple models stemming from decision analysis and management science are suggestive, they may systematically overlook likely features of the real world. The balance of this section briefly describes some of these departures and how they might be handled.

The first is that biometric decisions and impacts do not hang simply on probabilities (or risks) but also on consequences (or exposure). A simple model of the biometric adoption rate (described in the Annex) shows that, while adoption in low-exposure settings is likely to follow the conventional S-shaped path typical of many other technologies, increasing exposure (potential losses in the event of an identification failure) can lead to discontinuous jumps in adoption rates and path dependence (persistent inappropriate adoption decisions). This is illustrated in Figure 4. The left-hand pane graphs the adoption rate against risk and exposure, and the two right-hand panes show the response to changes in risk for high and low levels of exposure.

⁵ This is not uniform – it has been found that some biometric systems show a peak of performance for young adults, dropping off for children and the middle-aged. This has been seen as an unconscious reflection of the average age of those who develop such systems.

⁶ Ref Kahnemann and Tversky.

⁷ i.e. whether they believe that biometrics represents a way of increasing security or a way to prevent deteriorating security.

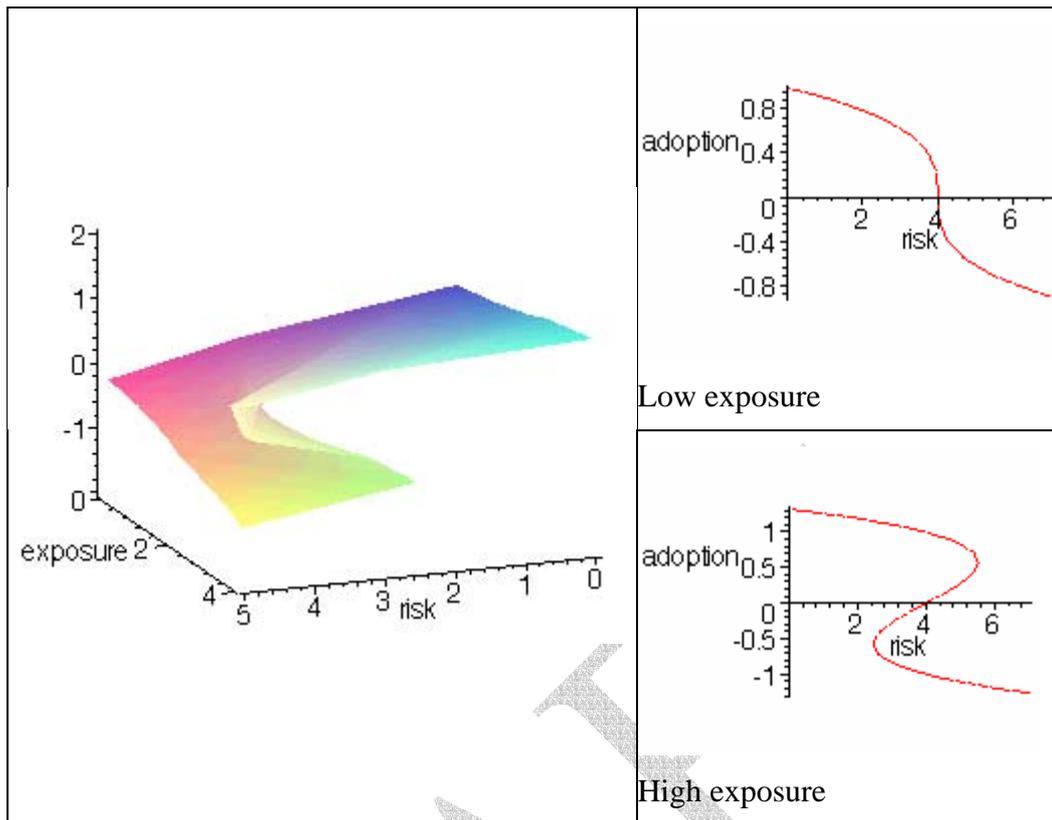


Figure 4: Risk, exposure and adoption

In this diagram, use of the biometric implementation is measured on the vertical axis. For low values of exposure, adoption follows a roughly sigmoid curve (top right), but as exposure increases, the change becomes discontinuous. Three practical observations may be made about this model.

- If deployment is widespread and/or stakes are high, progressive improvements in security may not lead to immediate or extensive increases in uptake;
- If increases in adoption lead to increases in (perceived) risk – perhaps through increased levels of spoofing or increased publicity for identification failures - adoption will not necessarily reverse or come to a halt immediately;
- As a result, rates of adoption may behave cyclically unless steps are taken to limit exposure.

A second point reflects the empirical finding that decisions about biometrics are typically undertaken to control risk –an enhancement to system function intended to insure against adverse consequences. As experience and acceptance accumulate, it will become increasingly likely that people will begin to recognise the positive benefits of biometrics in appropriate areas. When the conceptual ‘reference point’ shifts from averting losses to seeking gains, it is likely (according to both experimental and empirical evidence) that people will behave as though they are more risk averse. A further relevant finding from psychology and economics is the “Allais paradox” that reduction of the probability of an outcome by a constant factor has more impact when the outcome was initially certain – this suggests that if biometrics are regarded as nearly perfect (like DNA matching), failures will have a greater impact than if they are regarded as merely very good.

A final point relates directly to spoofing. A claim of identity commits the individual to a decision on the part of the identifying system and subsequent costs or benefits. If, through costly effort, an individual can bring their measured signal closer to that of another, the probability of a false positive rises. Both sides of the identity ‘transaction’ have strategic decisions to make, so the error rates in Figure 3 are not only technological but also strategically determined. Institutional changes (e.g. in penalties for false claims or false rejection of claims) as well as technological changes (in matching algorithms, decision modules or ‘spoofing’ methods) will thus produce their effects by ‘shifting’ the equilibrium outcome. In particular, it is possible to have either separating equilibria (in which individuals do not imitate one another and the level of ‘personalisation’ is high and ‘pooling’ equilibria – in which individuals can impersonate each other and/or where the level of personalisation is low.

In other words, the mere deployment of biometrics by itself does not necessarily mean that the system uses a high level of personal identity.

3.3 Costs and benefits

The operation of the system produces, in addition to more or less accurate matches, a series of monetised costs and benefits. These include a wide range of separate costs detailed in the business literature. The current discussion merely highlights a few of particular relevance to the economics of the sector as a whole.

Many costs turn out to be higher than expected from initial design or trials, including those associated with testing, enrolment⁸, operation (personnel requirements, speed and reliability), template maintenance (esp. template aging)), system integration and human resources, etc. Depending on the deployment, there may be additional liability costs as well. By the same token, however, the opportunity costs of adoption should be adjusted to take account of ‘hidden’ costs of any previous system (e.g. password reminders, document replacement, error costs, etc.).

Costs are also frequently cited as a barrier to adoption, but recent data show that the costs of sensors, in particular, has been falling rapidly. However, they remain substantial – especially for flagship large-scale public sector deployments. US estimates of the set-up cost of border-crossing biometrics fall in the range €1.05-2.18 billion, with annual recurring costs of between €26 million and €1.13 billion. UK estimates of costs of the identity card system place the development cost at between £415 million (Home Office) and £3 Billion (Home Affairs Select Committee) and the recurring cost at £85 million per annum⁹. To this system cost estimates must be added additional personnel, etc. costs (for example, the cost of face-to-face visa issuance in the USVISIT case) and the inevitable cost increases¹⁰ and delays associated with large public procurement and deployment of IT systems¹¹.

⁸ Early deployments (e.g. at Expo92) required higher than anticipated levels of user support. See M. Rejman-Greene, “Biometrics – real identities for a virtual world,” at <http://www soi.city.ac.uk/~kam/rejman-greene.pdf>

⁹ <http://politics.guardian.co.uk/homeaffairs/story/0,11026,1342061,00.html>

¹⁰ Within one week, the estimated cost of the UK card to users has risen from £77 to £85.

¹¹ This is particularly the case for the public sector. Not only is poor cost performance generally coupled with delays and performance problems, but the widespread introduction (full coverage rather than phased introduction) of systems that represent replacements rather than incremental upgrades has the effect of magnifying both actual costs and the ‘expectation risk’ associated with the leading customer role.

On the benefit side, the contributions of assured identity go beyond potential efficiency improvements to those associated with higher trust in the system's operation¹². On the other hand, trust is a form of *reliance* and 'too much' trust can lead some stakeholders to reduced levels of precaution – even when they are best-placed to manage risks accruing to the system as a whole. An example is provided by security: those in possession of classified information tend to be much more careful about revealing it in insecure environments. This increases the returns to corrupting an individual with clearance above the returns to penetrating the system. In other words, 'hardening' the outer boundaries or entry points of a system may reduce overall security if internal precautions or the interstices between parts of the system are relaxed. Of course, keeping internal barriers high is also costly. The true level of cost and the appropriate way to manage cost depends not just on the comparison of risky outcomes but on the comparison of different types of risk.

Because biometrics will be used primarily in the context of other systems, its true impact is impossible to estimate and may be difficult to measure *ex post*. Available data tend to fall into three categories:

The first source measures costs associated with problems biometrics are intended to solve. One particularly category is Identity theft¹³; in the UK it is estimated to impose annual costs of €1.95 Billion (10% of all fraud, and growing). By type of frauds, this is shown in Table 1.

Table 1: Scope of identity theft in UK

Category	Amount (€M)
VAT	€ 306.85
Money laundering	€ 563.75
Health Authorities	€ 1.07
Welfare fraud	€ 49.95
Immigration	€ 51.38
Credit cards	€ 528.07
Insurance	€ 356.81
False ID/impersonation fraud	€ 89.20
Total	€ 1,947.08

In the US, where it is growing at 300%/year growth, identity theft cost an estimated €5.5 Billion in 2003 and over the period 1998-2003, affected 28 Million US citizens. US estimates put the worldwide cost of identity theft in 2003 at €166.3 Billion (2003), a figure that is projected to rise to €903 million by 2005.

These estimates have three principle drawbacks as measures of benefit. First, they are inherently inaccurate. The estimates are in many cases based on existing processes for document issuance and checking that are themselves insecure. Reported financial losses – and even incidence counts – are often systematically underreported. Many instances of identity theft do not involve direct financial consequences. In any case, many cases of identity fraud are not promptly discovered, as shown in Figure 5 (based on data from the US Federal Trade Commission).

¹² Cave, J. (2005) "The economics of cyber trust between cyber partners" in R. Mansell and B. Collins (ed) *Trust and Crime in Information Societies*, Cheltenham: Edward Elgar, 380-428.

¹³ For US data, see e.g. <http://www.consumer.gov/idtheft/stats.html>. For the Cabinet Office report (2002), see http://www.homeoffice.gov.uk/docs/id_fraud-report.pdf.

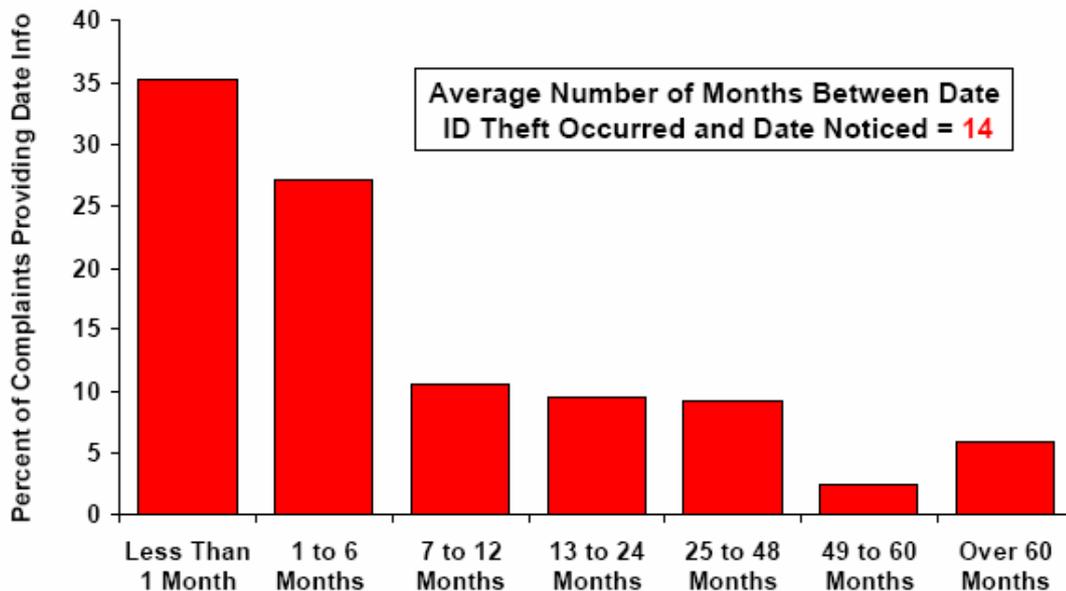


Figure 5: Delay in detecting identity theft

In some cases – especially those imposing no direct financial cost - they may never be known. It should also be noted that the effects of identity fraud are not limited to the immediate site of the identity theft, as the following ‘case study’ from the UK illustrates¹⁴:

The private sector fraud register CIFAS prevented one potential fraudster, who was eventually prosecuted and imprisoned, from succeeding with an intricately-planned £1m scam. Having sat 28 driving tests across the country, obtaining different identity documents from each, and registering on the electoral roll for a variety of rented properties, he was able to open multiple bank accounts. He cycled money between these accounts for several years, building up a healthy transaction history. This then enabled him to obtain multiple loans and credit. One CIFAS member became suspicious by the unusual nature of payments between bank accounts and the subsequent data search revealed the extent of the scam, which affected many more members.

Second, the actual costs of identity fraud consist not merely in the amounts taken, but should also include costly precautions and even the lost gains from trade associated with transactions foregone as a result of fears or costs of precaution. In this sense, such data understate the true costs of identity fraud.

Third, such a measure assumes that biometrics will be completely efficacious in eliminating identity fraud. This is highly unlikely; biometrics are not completely effective and there are costs associated with both false positives and false negatives. In addition, the strengthened security associated with biometrics may displace fraud to other locations. The Association for Payment Clearing Services (APACS), which deals bank and credit card fraud in the UK, estimates that credit card crime grew from €136 million to €587 million in the three years to 2001 and reach €28 million by the end of next year, primarily due to organised crime. The industry feels that strengthened credit card authentication procedures will shift fraud further upstream

¹⁴ UK Cabinet Office: “Identity Fraud: a study”, July 2002.

towards “account takeover” (whereby genuine accounts are hijacked for fraudulent purposes) and other identity fraud.

The second source measures specific cost efficiencies associated with immediate effects of biometrics deployment. Such data tend to come from industry sources and are often proprietary or associated with marketing literature – so they are not reported here – but they do represent an element additional to the savings in identity-related loss and crime. These process-orientated data should also include the costs of mislaid, multiple or corrupted identities to the extent that biometrics can influence these. Like all such estimates, they should be measured in terms of lifetime cost of ownership, make due allowance for capital changes (e.g. in financial, physical, IT and human capital) and recognise the many external effects of new identity management procedures on a whole range of internal processes.

Estimates of willingness-to-pay – or future revenues. These provide a lower bound estimate of the consumer surplus from biometrics. As usual, it is necessary to recognise that many of the gains will be inframarginal. As with many IT improvements, much of the gain in functionality is accompanied by falling costs: the two effects may offset each other in terms of measured market revenues, but mask simultaneous increases in consumers’ and producers’ surplus. A related factor stems from the role of biometrics in reducing and transferring risk – the consequence may be that risk averse consumers can substitute away from costly hedging or insurance towards more secure forms of transaction – again, there will be consumer surplus gains not measured in the biometrics market or those markets where the measures are used. In this case, however, there may also be external losses in sectors providing other means of identity management or insurance against this class of risk. The market of ‘information assurance (see www.iaac.org) may be particularly affected.

Specific costs and benefits are reflected in the criteria used to compare and choose among biometric solutions:

- Universality - Each person should have the characteristic.
- Uniqueness - No two persons should have the same characteristic (relative to the population involved) – this depends on the size of the codified template as well as the accuracy of the sensor and influences the proportion of false positives..
- Permanence - The characteristic should neither change nor be altered – this is more problematic for, say, hand or face geometry than fingerprints or iris scans, and increases false negatives and/or costs of template maintenance..
- Collectability - The characteristic can be measured quantitatively. Faithful and inexpensive sensing and codification of the characteristic can reduce the costs of decentralised solutions and the participation costs of small or medium-sized enterprises.
- Performance - The characteristic can be efficiently measured in terms of accuracy, speed, robustness and resource requirements. For instance, iris scans are easier to acquire than retinal scans, and facial information (e.g. photographs) are more likely to be available for e.g. antiterrorism or anticrime purposes than iris scans or even fingerprints. This reduces the burdens of identification and opportunities for spoofing. Fast scanning and processing can increase acceptance – and again reduce the ‘overhead’ burden on small firms.
- Acceptability - The characteristic should be acceptable to the public.

- Circumvention - There should be no easy way to fool the system.
- Robustness and auditability – Measurement of the characteristic should be immune to environmental variation and/or deliberate interference, and there should be an appropriate audit trail in the event of challenge.
- Appropriate functionality and level of identification for the associated transaction or service – one oft-cited barrier to acceptance of biometrics in private-sector applications is user perception that a given biometric system may be too intrusive or simply too strong for the service being sought¹⁵.
- Information security, including immunity from unauthorised access, alteration or destruction – this applies to templates and transmitted measurements, and typically involves use of encryption
- Acceptance to users on both sides and to legal, financial and other concerned authorities.

Because no single method is likely to offer optimal performance in all categories, it is necessary to make explicit tradeoffs and, where appropriate, to incorporate additional measures to compensate for deficiencies. For instance, uniqueness on a global level is almost certainly infeasible given reasonable template sizes and processing methods, but this might not matter for verifying identity against a short list of authorised persons. Where greater accuracy is appropriate, a combination of biometrics can be used.

4 Market situation

This section provides some background on the evolution of the biometrics market and the associated economic issues and prospects.

4.1 Market shares by region

The biometrics industry as a whole had its start in the US. Recent regional data are not available from public sources, but trade press reports indicate that the European share of the market is growing rapidly. This is particularly true in application areas such as banking, where consolidation and widespread closure of bank branches is creating strong demand for biometrics associated with ATMs. In addition, recent European government initiatives (e.g. the UK identity card scheme) are likely to lead to strong demand, though at the moment the USVISIT contract is the largest public procurement contracts in the field. Table 2 gives a rough idea of the pre-9/11 situation.

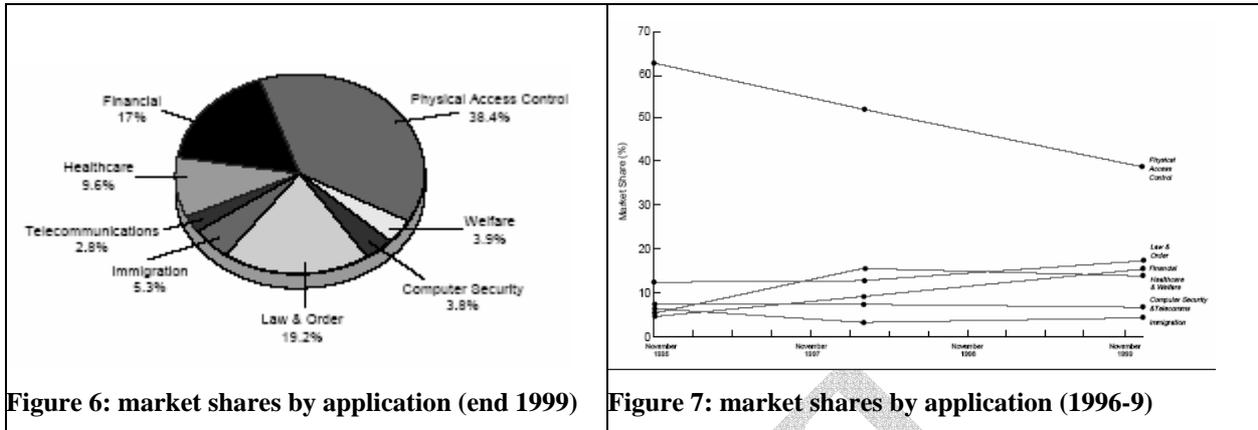
Table 2: regional market shares pre-9/11

<i>Region</i>	<i>11/1996</i>	<i>3/1998</i>	<i>12/1999</i>
W. Europe	12.2%	11.8%	18%
E. Europe	1.7%	1.3%	1.8%
N. America	70.2%	66%	57%
S. America	3.4%	6.9%	9.3%
Asia Pacific	9.9%	10.6%	9.5%
Africa & middle east	2.6%	3.4%	4.4%

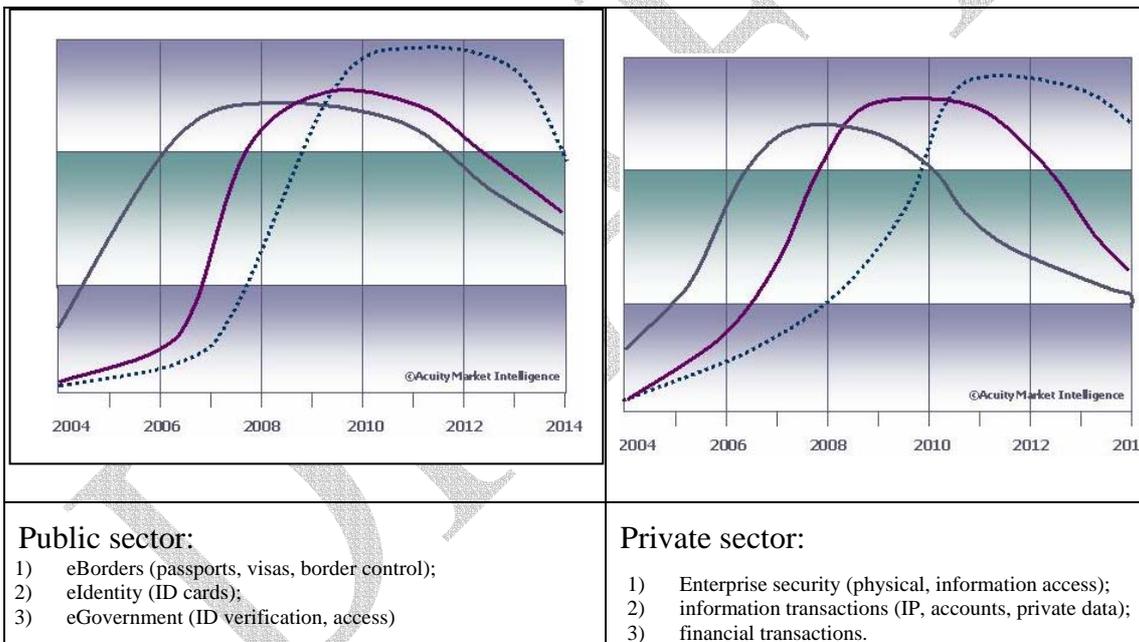
¹⁵ "Prepare to be scanned" Economist, December 4 2003,
http://www.economist.co.uk/displaystory.cfm?story_id=2246191

4.2 Market shares by application area

Figure 6 and Figure 7¹⁶ show the state and evolution of different application areas in the pre-9/11 world. Subsequent data from non-proprietary sources were not available.



Over the next decade, the public sector and private sector markets are expected to show phased growth. The following Figures show projected market growth for key public and private sector applications.



4.3 Market shares by technology

Figure 8, based on data from Biometrics Technology Today, shows the growth of market shares by biometric type.

¹⁶ Source: Biometric Technology Today

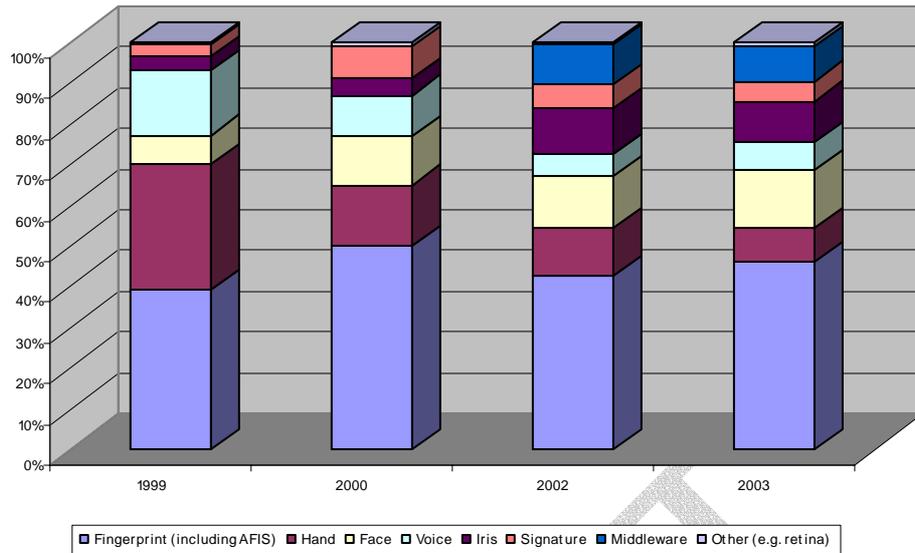


Figure 8: market shares by technology

This suggests a fairly consistent dominance of the market by fingerprint technology, with a dwindling emphasis on hand geometry and voice recognition, and the growth of iris recognition. Behind these data are other important changes – one being the growth of the non-US market, where hand recognition is only occasionally used and another being the growth of the overall size of the market, which can be seen in Figure 9.

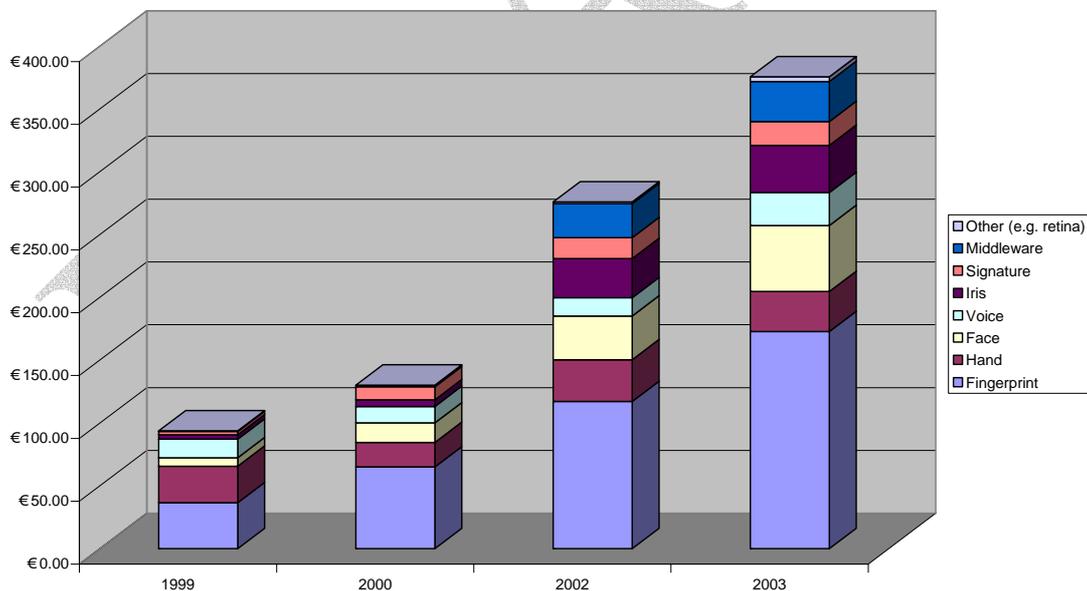


Figure 9: Revenues by technology

By contrast with the above, this shows strong revenue growth in fingerprint, which is certainly likely to grow substantially as fingerprint scanner costs fall (currently expected to be around €30) and these devices are bundled with computer hardware. Facial recognition is also showing growth and the growth in iris is much stronger.

4.4 Barriers to growth

- The industry follows in some respects a typical life-cycle for a high technology sector:
- an initial phase of exploration and widespread uncertainty, with low levels of highly-inelastic, high-income demand (typically public-sector), scarce capital available under tight controls (e.g. via exploratory procurement arrangements or ‘business angels’) and a small number of pioneering firms;
- a growth phase during which there are a large number of entrants, a maturation of products and uses, an increase in available capital (e.g. from mainstream venture capitalists, mergers and acquisitions and strategic alliances), and a much larger number of more diverse customers;
- a consolidation, maturation or ‘shakeout’ phase during which industry numbers shrink while volumes rise – the impact on profitability depends on the extent of effective competition, and the structure may change in other ways (see below). Capital is more likely to be raised through equity markets, with successful firms launching IPOs.

The cyclical pattern is not exogenous and is not the same for all technologies. Different technologies offer different mixtures of cost and performance.

The barriers were surveyed in 2000 by Clara Ferrando at Georgia Institute of Technology¹⁷ and shown in Table 3.

Table 3: Barriers to growth

<i>Barrier</i>	<i>Prevalence</i>
High cost	18%
Limited capital	10%
Lack of standards	26%
Regulations	2%
User acceptance	25%
Inertia	8%
Other ¹⁸	11%

It is clear that, in the perception of the surveyed members of a BC Listserv group, the dominant barriers were standards, cost and acceptance.

4.5 Structure

Most biometrics firms come from the private sector. The supply sector overall appears to follow the ‘experience curve’ pattern of an initially small number of firms, a lot of new entrants and a consolidation to a much smaller number of firms during the mature phase. This consolidation is well underway; despite the relatively strong demand growth in response to security concerns and spiralling levels of identity fraud; mergers and bankruptcies have prevailed in recent market reports. But the cycle is not the same in all technologies: it is much more advanced in fingerprint, while newer technologies like iris still have a large number of small firms pursuing diverse approaches. Data on market concentration were not available, but basic economics

¹⁷

¹⁸ a combination of factors; inertia of major players to adopt; immaturity of technology; and scarce information and education of users

and various commentaries on the industry suggest that concentration remains relatively high even during the expansion phase. The consequence may be that dominant firms survive disproportionately. This pattern is similar to that in the pharmaceutical industry, where small firms trial a range of approaches, with the successful ones being licensed or otherwise acquired by large incumbents.

This pattern of entry and shakeout is common to many industries with positive 'learning by doing' or 'reputation effects.' The underlying cost dynamic reflects the importance of cumulative experience: firms with longer experience can deliver functionality at lower average cost than novices or new entrants. Part of this reflects falling unit costs (including marketing, testing, etc.), but another part reflects the value of experience in better matching biometric functionality to user needs. The other side of this picture is a cyclical shift in the elasticity of demand (and thus in return on investment). Early adopters tend to be experimental (and thus as interested in learning for the future as in current performance), relatively risk-neutral and relatively well-heeled. These factors combine to produce inelastic demand and high return on investment, especially when the supply side is relatively small and can effectively coordinate its activities. This eases the availability of investment capital and encourages new entrants on both sides of the market. During the resulting shakeout phase, demand becomes much more elastic and returns drop – especially if oligopolistic control weakens. This leads ultimately to business failures and a concentration of the market to the mature phase, where returns rise again.

The tendency to concentration is reinforced in this industry by a number of specific factors. First, uncertainties surrounding the functionality of the technology result in relatively high fixed testing costs, which raises entry barriers to small firms – though not necessarily to the technologies and implementations they develop, which can be tested and deployed by large firms once their potential has been demonstrated.

A second factor is that the provision of biometrics is a way of helping a customer to cope with derived uncertainty relating to physical security, privacy, economic loss, etc. Especially during the early phase of a technology, the customer is likely to have inferior information about technological and functional characteristics and will therefore wish to bundle a degree of 'assurance' with their choice. This affects market shares in two ways – it favours incumbents with demonstrated track records and a willingness to work in partnership with customers (cheaper for large firms with ongoing relations) and it favours firms with a large installed base among a particular application area or customer group through the reputation effect. These same dynamics can be seen in e.g. large-scale government IT purchases.

A third factor is tipping equilibrium. Rosen (1981) pointed out that any industry with low reproduction costs and popularity effects may give rise to a highly concentrated "superstar equilibrium." Moreover, as Katz and Shapiro (1994)¹⁹ point out, complementarities in adoption of the same or interoperable technologies can accelerate the increase of concentration and market power: as each new user adopts a given technology, other users – particularly those who transact with those early adopters – find that particular solution increasingly attractive. The possibility that competition might decay is enhanced when account is taken of the 'layered' nature of the biometrics industry [see Figure 2]. Market power developing in one of these

¹⁹ Katz, Michael L., and Carl Shapiro (1994). "Systems Competition and Network Effects," *Journal of Economic Perspectives*, 8:93-115.

layers can often extend to the others, because each is complimentary to the others in delivering value²⁰. Relations among the layers (and the lead in developing the market) change as the market matures, which lends a certain complexity to the possible evolution scenarios.

A fourth factor is the importance of intellectual property rights (IPR). Increasingly, competition authorities have come to recognise the importance of IPR (especially patent) holdings in merger analysis and, by extension, in assessing the nature and efficiency consequences of market power. Put simply, if a firm holds key patents it need fear no competition; if it chooses to allow competitors to licence its technology to avoid the scrutiny of regulators, it can do even better²¹ by setting licensing fees appropriately. It can encourage entry of efficient rivals and extract not only the economic rent arising from its own technology but even the additional gains from new entrant's innovations and improvements (whether in product or process). Ultimately, such strategies are self-defeating; they encourage bypass competition and antitrust policy responses, they keep prices high and limit the growth of the market and they prevent the 'medicine of competition' from driving their own costs down further. But, as the recent situation in iris scan algorithm patenting shows, such self-defeating tactics still prevail²². Further ramifications of the IPR limitation come from the possibility of creating 'patent thickets' and 'patent clusters' to prevent innovative rivals from reaching the potential market.

While IPR represent a 'privatisation' of knowledge, standardisation represents – in some ways – a public good approach. Strictly, a standard or norm is a set of specifications adopted by a group of interacting parties. Standards are particularly important as means of protecting interoperability and as a way of decreasing information asymmetries between buyers and sellers. An open standard – whether given to the market under some form of general public licence or cooperatively developed – can enhance the effectiveness of competition (by lowering entry barriers to the 'core' of the market) and stimulate innovation (by providing guidelines to developers of complementary products). But standards may compete for market presence. One source of competition is where demand is differentiated – standards appropriate to one group of consumers may not fully meet the needs of another. If the economies of standardisation on the supply side are strong enough, there may be a period of standards competition culminating in either the triumph of one standard or

²⁰ Complementarity – and the accompanying incentive of a dominant 'bottleneck' player in a value chain, can directly affect the internal security of a set of applications. Typically, this effect weakens security within the set of applications (though it may strengthen security at the 'borders' of the system. In this situation, open standards pertaining to part of the system may have a perverse rebound effect on the functionality of the system as a whole. This is not unlike the well-known 'second best' principle in industrial economics. See e.g. R. Anderson "The economics of trusted computing" at http://www.netproject.com/presentations/TCPA/ross_anderson.pdf

²¹ More concretely, suppose that a market requires a combination of 2 technologies (say, recognition algorithms and scanners), and that a firm owns exclusive patent rights on one of them, with associated constant marginal cost c_1 . For the second technology, it has a constant marginal cost of c_2 for a combined cost of $c_1 + c_2$. If the firm excludes other firms, it can charge the monopoly price $P(c_1+c_2)$, which is larger than the competitive price (c_1+c_2) . If potential competitors have lower cost for the second technology ($c' < c_2$), the incumbent could refuse to license the algorithm (and go on earning $P(c_1+c_2)$) or it could charge a licensing fee to its more-efficient rivals equal to $P(c_1+c') - c'$. Even if the new firm(s) behaved competitively, they would charge the same price as a monopolist who had the best technology – and then surrender this extra money to the incumbent. Innovation would still proceed, but almost all the returns would go to the incumbent, and only trivial amounts to entrants and customers.

²² The main patent holder is Iridian Technologies and the patent is due to expire next year. The incumbent has guarded its rights jealously, launching attacks against actual or potential rivals – notably a licensing dispute with LG technologies and a patent dispute with IriTech.

the emergence of a ‘second best’ compromise. In the former case, investment in the losing standard may be lost. Another source of competition is rent-seeking. Some standards are owned by their developers. This may be an overt proprietary standard, or it may be a hidden consequence of a ‘voluntary’ standard bundled with a dominant position in a complementary market. This is apparently the case between the ‘public’ (or collective) BioAPI standard and the Microsoft (BAPI) standard. It is not impossible that they might converge in subsequent generations, but it is by no means guaranteed. A further consideration involves the cumulative nature of standards.

Leading customers can also promote lock-in to specific providers as well as specific solution. This is not inevitable; public sector clients in particular can adopt such ‘innovation friendly’ strategies as design competitions and multiple sourcing to balance creativity with value for money. But in cases where the lead customer is under pressure rapidly to choose a solution, where the proportion of near- to medium-term demand represented by the lead customer is large, where the lead customer has or seeks long-term relationships with a small number of suppliers for the solution or for integrated services that incorporate biometrics, and where many other customers need to interoperate with the lead customer, it is highly likely that the early winners will sustain their advantage as the market matures.

4.6 Sectors

The following sections discuss some of the specifics of biometric applications in different sectors. This does not attempt to provide a complete market analysis of these sectors or a projection of future revenues and market performance. Rather, it attempts to identify implications arising from the application area and the stakeholders involved for future competitiveness, efficiency, interoperability, etc.

Government and other public sector

This continues to be the leading sector in terms of volume, new technology adoption, scale of individual projects and prominence. The balance of biometric modes and application shifted after 9/11 towards transport security and immigration, with widespread plans for adoption of biometric passports. The specific aspects of this sector include the emphasis on international interoperability, different procurement and contracting arrangements and high levels of political risk. The public sector is also one of the leading clients in the health, security and transportation sectors.

In response to security concerns and public sensitivity to the economic and other impact of immigration biometrics are being used in a range of border control (visas, expelled persons, asylum seekers) and passport initiatives. [see ‘Transportation’ below: possible case study on US-VISIT procurement: scale, roll out from air to land, international matching and additional biometrics]

National identity cards incorporating biometrics are being developed in the UK, Canada, Serbia and a number of Gulf States (Bahrain, Oman and UAE). [Case study material on UK identity card: costing, impact assessment, procurement, pilots, industry consultation]

In addition, concerns relating to benefits entitlement and the increased use of smart cards to control provision of public services have created a further area for application – either in a stand-alone mode or in combination with a national identity card.

Physical access control

This has been the dominant application area since the advent of biometrics, but is rapidly being supplanted by computing-based uses. In 2000 it still accounted for 42% of the biometrics market (including time and attendance). This share was dwindling but has revived strongly since 9/11. The dominant trend in this area has been to expand the system to incorporate other functions such as time, attendance or physical location.

Retail and other payments

Biometrics are already being trialled in a wide range of retail payment applications. Although quantitative data are not yet available, it is clear that some combination of biometric identification (for non-cash payments) and smart cards (for electronic cash payments) will come to dominate retail payments within the next decade. There are several potential models.

One model is the 'trusted identity card' model in which one or more trusted pieces of biometric identification are used to provide 'public good' authentication services. Typically, one of these is state-issued (to provide security of personal identification) and the other is proof of financial standing (a cheque guarantee card or a major credit card). In this model, no liability attaches to the identity provider. Typically, this is also a decentralised model, with local template storage (since the 'public good' provider does not participate directly in or benefit from the transaction).

A second model is the 'closed-system' or proprietary model, in which a small cluster of retail outlets provide their own identification arrangement. This model is compatible with a centralised implementation, since the identification managers benefits directly from the trade and thus have an incentive to provide computational, storage, security and communication services. It may be that a specialist third party provides these services – either a systems integrator or a commercial credit institution, operating in a manner analogous to that of the issuing bank in a credit card transaction.

A third model is a converged smart card/biometric model in which the individual's token is both a means of identification and a store of value. In this case, the function of the biometric is to provide 'theft insurance' to the user of the smart card, and the template can be stored locally. It should further be mentioned that 'open' retail systems are now being deployed (e.g. Electracash) that will accept a wide range of biometric identity certifications, with transaction limits appropriate to the security provided.

Of course, in any transaction the buyer, as well as the seller needs a degree of assurance or protection. Biometrics could be used to enhance consumer protection as well. One way is simply to have the seller verify (and record) his identity as well, providing assurance of the validity of the transaction. Seller identification could also be used to establish an ensured 'liability tail' in the event of faulty merchandise or other claims arising from the transaction. This may also involve buyer identification: for instance a record of a transaction that incorporates both parties' identities (and the date and other salient details) could be recorded by means of hash functions in the same way as contracts are electronically time-stamped at the moment. This provides both parties with a maximal level of protection and removes many of the agency problems involved in commercial trades.

One further comment is that retail transactions are increasingly likely to be electronic in nature. Given soaring levels of fraud, it may happen that the availability of biometrics in some (but not all) commercial channels will ‘tilt the competitive playing field’ against those that cannot offer equivalent protection. This need not be an insurmountable obstacle in view of the proliferation of biometrics in combination with IT and telecommunications devices. However, it seems likely that some ‘hierarchical trust’ arrangement would be needed in order for these devices to compete successfully with other channels. At the moment, the third-party protections against fraud, abuse, repudiation, etc. available in off-line commerce – such as limits on customer liability – are being withdrawn from remote eCommerce. This creates an opening for biometric identity intermediaries.

Telecommunications

Telecommunications services are increasingly capable, increasingly integrated with other services and increasingly linked to individual data and resources. To date, the growth of mobile telecommunications – let alone the kind of ‘ubiquitous’ or ‘converged’ telecommunications in which physical presence is neither technologically not economically relevant²³, has been restrained both by cost and the sheer difficulty of spanning the gap between the ‘personal net’ of communications and the ‘device net’ through which it is implemented. Concretely, while it is possible to route a call to any office in a major building, and even to use a sensor to locate the individual to whom the call is directed, it is not easily possible to ensure that only that individual receives the call or data. With biometrics, this becomes (in principle) straightforward.

Theft of communications devices and services is another widespread problem where biometrics is being tested (e.g. Identix).

Financial services

The financial services sector has been the third largest area for biometrics (15% in 2000). This is likely to accelerate (in size if not proportion) as a result of new types of fraud (e.g. ATM fraud), changing needs for financial identity management and the changing structure of banking itself.

The use of sophisticated devices to intercept PIN numbers has made the development of stronger identity certification essential.

Increasing participation in an ever-broader range of financial transactions (from asset trades to mortgages to balance transfers, etc. and the possibilities for fraud and money-laundering also increase the need for identity verification. This can provide a direct means of controlling the explosive growth of borrowing (and the consequent societal problems), closing tax loopholes and limiting possibilities for money laundering. In addition, other forms of transaction (cheque-cashing, foreign exchange, credit and debit cards) are already increasing the need for secure electronic identification – this need is set to increase as converged communications devices are used for payments.

On the structural side, as retail banks cut back on physical offices in favour of ATM, telephone and internet banking, and as they lose market share to ‘virtual’ banks and banking services offered through non-bank institutions (e.g. supermarkets), the

²³ Such ‘location independent’ communications are already being tested by a number of telecom providers (e.g. France Telecom, BT).

need for a secure and uniform means of managing identity is essential if governance of the financial system is to be maintained.

Health

In the health area, biometrics is widely used to prove entitlement (especially in public health systems) and to link individuals to electronic health records. The economic effects are likely to be increases in the efficiency of health care provision, a reduction in the transactions costs associated with lost records or fraudulent use of services, and possible knock-on effects on health care economics. One of the more significant developments is the rebalancing of provision from physicians towards physician extenders and (for some services) chemists and diagnosticians. It is essential for this expanding range of providers, who come from a range of backgrounds and have widely varying ability to collect, share and manage records, to have a common data store in order to serve their patients. This need will certainly increase as a result of advances in e.g. genetic medicine with its greater degree of personalisation and integration of treatment at the patient level. At the same time, this information is of direct and financial interest to insurance companies, so the needs for integrity, privacy and access and use control are stronger than in many other applications. One significant development in this area is the US Health Insurance Portability and Accountability Act, which directly encourages the use of biometrics to protect confidentiality and security of healthcare information. [possible case?]

IT, etc.

The computing industry has commanded the second-largest share of the market (25% in 2000). This is growing with the inclusion of e.g. fingerprint sensors in laptops, the development of the Windows-specific BAPI interface standard and the growth of biometric implementations in converged computing communications equipment. Because the primary vulnerabilities of computing systems, from laptops to intranets, are human, biometrics are a natural part of the quest for information assurance and critical systems protection. While much of this protection can also be seen as part of retail transactions, it is fair to say that IT-enabled transactions bring their own issues²⁴. There is no doubt that the pace of electronic commerce is accelerating rapidly and that ecommerce fraud and identity theft are increasing even faster. Biometrics do not hold the answer on their own – but in combination with e.g. digital certificates and digital signatures²⁵ they can form part of a solution.

Transportation

In the transportation sector, biometrics are increasingly used to provide physical access control – both to transportation system and to the destinations it connects (including immigration). Indeed, airports and government ministries concerned with migration have led the way in deployment of new biometric entry controls. In the post-9/11 environment, this has obviously accelerated enormously. Major initiatives include the endorsement of digitised facial images as the biometric passport standard by the International Civil Aviation organisation, visa schemes such as US-VISIT: the

²⁴ One such issue not explored here is the question of non-human or virtual ‘avatars’ that are increasingly used to make transactions in electronic environments. These do not have biometrics, and the questions raised by the possibility of delegating biometrics go beyond the scope of this background piece.

²⁵ Ref. to EU Digital Signature and eCommerce directives and US ESIGN Act.

European VIS scheme²⁶, and the border-crossing iris check used to identify those who have been expelled from the UAE. As a further by-product of security, biometrics are increasingly used to control physical access in air- and seaports.

Case study: Biometric travel documentation

The United States Visitor and Immigrant Status Indicator Technology (US-VISIT) initiative is intended, when fully in place, to cover all (land, sea and air) points of entry to the US with a biometric identity system that provide comprehensive information on all would-be entrants, who will be linked biometrically to their travel documents. All visas and all passports from visa-waiver countries will incorporate biometrics. The project was rapidly pushed out, with a stop-gap system deployed at the main 130 ports of entry (115 airports, 15 seaports). In addition, the US has a pilot 'registered passenger' programme at Minneapolis-St. Paul and pilot biometric physical access controls at 20 major airports.

In the meanwhile, a large and somewhat vague tender was let to a consortium led by Accenture to develop and deploy the 'real' system over the next decade. The tender documentation was not very detailed considering the unproven technology, likely size (estimated as high as €1.3 billion), and the winning consortium apparently proposed a methodology instead of a detailed technical solution²⁷ of the project. In consequence, many of the important technical decisions rest with the private sector and the functional specification of the system will be fixed only after the technology and associated processes are in place. In another departure from procurement 'good practice,' the requirements of the programme fall in the first instance on non-US citizens. This means that the programme will necessarily stimulate the development of solutions outside the US before such solutions are developed within the US. In consequence, the US solution remains uncertain as to scope and functionality, the balancing of economic and societal issues and the development of legal and commercial frameworks will be delayed and the US is unlikely to retain its initial lead in this policy area.

By contrast, the European Community is taking an more measured and balanced approach, debating issues of privacy, civil liberties and societal acceptance and establishing a harmonised legal framework for developing technical and societal standards before beginning deployment. In the meantime, Australia has taken a technological lead in facial recognition for passport issue and border control through extensive *in situ* testing, closely followed by Chile, which became the first country to 'go live' with a fully-deployed facial recognition kiosk system that checks digital photographs against the passport template, a local watch list and the Interpol database. Some key initiatives²⁸: are listed in Table 4.

Table 4: Transportation biometrics

<i>Country</i>	<i>Status</i>
Electronic passports	

²⁶ States comprising more than half the European population have publicly committed to biometric travel documents: Bulgaria, Denmark, France, Germany, Ireland, Italy, Netherlands, Poland, Slovenia, Sweden, Switzerland and UK. On June 8, 2004, the Interior Ministers of all EU member states agreed to incorporate biometrics in travel documents.

²⁷ C. Most (2004) "Biometrics and border control: beyond US-VISIT", *Digital ID World*, September/October

²⁸ *ibid.*

<i>Country</i>	<i>Status</i>
Australia	RFP for chip integration, issuing checks
Belgium	Prototype
Canada	Announcement
Denmark	Contract awarded
Germany	Initial study
Italy	Pilot completed
Netherlands	Pilot
Sweden	Request for interest
Switzerland	Request for interest
UK	Pilot
US	RFP issued
Border control	
Australia	Smartgate pilot
Chile	Deployed in Santiago
Germany	Pilot at Frankfurt
Netherlands	Deployed at Schiphol
UK	Pre-procurement, Heathrow pilot

5 Economic outcomes

The economic consequences of different biometric policies and scenarios include profitability, competition, and efficiency. The latter comprises three distinct types of efficiency: allocational efficiency measures the extent to which all current participants could be made better off by an alternative arrangement. Technical efficiency measures whether current levels of service could be provided at lower (opportunity) cost. Dynamic efficiency refers to the incentives for future RTD.

These are aggregate measures; a further relevant outcome is distributional fairness reflecting changes in patterns of access. This includes inclusiveness: whether biometrics makes it more difficult to opt-out, the spillover impact on other forms of identification, the difficulty of repudiation of a system that is known to be accurate, etc. As indicated in the previous section, it will also consider likely cost and benefit incidence of across sectors and countries to the extent indicated by available or analogous data.

These elements will be fleshed out by a summary of major related initiatives and experience in Member States and selected reference countries (e.g. the US and Australia), and two exemplary ‘good practice’ studies examining the supply- and demand- sides of the biometrics industry.

Broader issue analysis will draw implications of enhanced biometric identification for economic transactions reflecting ‘economics of identity’ and trends towards ‘mass personalisation.’ It will analyse *trust* and participation (see next section). Another issue is the trade-off between privacy and security in terms of actual and perceived levels of implied consent, ownership of associated data streams and the likelihood of their reuse or disclosure. One key aspect will be the applicability or effectiveness of existing data protection. A further important issue cluster concerns crime. It seems obvious that biometrics reduces fraud; the impact analysis will assess reduction in direct and indirect fraud costs – including those of current, less-effective fraud

reduction, mitigation or compensation measures. However, as with genetic information, the presumed higher accuracy of biometrics creates other risks - especially in online transactions where biometric information may be easier to forge or mistake than to repudiate. Identity theft may simultaneously become less likely and more serious, and mutual assurance in on-line criminal enterprise may become easier to provide. A further risk is abuse of genetic biometric information. The upshot is that these technologies change levels of security and efficiency – the effects can be quantified using conventional industrial organisation tools providing care is taken to account for opportunity costs – including the impact of increased *reliance* and changes in the channels and locations of economic activity. Accurate prediction is difficult, but existing studies on other aspects of, e.g. e-commerce can be used to calibrate the likely outcomes.

5.1 Evolution Scenario Issues

The evolution of biometrics in the economy has been largely constrained by costs and acceptability.

Many researchers have commented on the degree to which the sectors that seemingly have the most to gain from enhanced identity (notably retail commerce) have been the slowest to adopt it. Initially, it was believed that this reflected the *level* of costs – available technologies offered (or were believed to offer) poor cost-effectiveness and increases in accuracy, security and reliability were associated with purchases of expensive embedded technology. In recent years, the costs of necessary equipment (and the associated computing power) have declined sharply, but take-up rates are still modest and growth remains slow. Part of this, no doubt, can be explained by the *distribution* of costs and risks and uncertainties about costs, performance and (commercial) benefits.

This in turn reflects the underlying technological model. Biometric technologies differ in terms of cost of encoding, location of processing, storage requirements, ability to reissue or cancel compromised identities, cost of scanning, etc. Consider a technology used in a retail outlet to verify identity – in other words, to facilitate, say, a credit card transaction. An individual's physical characteristics must be read, and the codified version matched with an enrolled template. If the template is stored locally (e.g. on a smart card), there is no need for the retail outlet to communicate with a remote database. On the other hand, the matching algorithm must be executed locally, and there must be some way to ensure the integrity of the local template. An alternative is to have a local biometric scanner that sends data to a remote database for matching and verification. This may be subject to interception and manipulation *en route*. For repeat transactions, it may be efficient to store the individual's template locally.

To analyse the appropriate degree of identification and the appropriateness of different identity management technologies, it is useful to separate a transaction into a number of potentially distinct *steps* – not all of these need be part of every transaction:

- Entry to a facility
 - Positive (membership on a specific list)
 - Negative (absence from a specific list)
 - General (identification from a large population)

- Access to product information/product specification
- Make purchase
 - with independently validated means of payment (e.g. cash)
 - using third-party payment (e.g. credit card)
- Make contract
- Accept delivery
- Pursue claims

Associated to these economic functions are various *technical functions*:

- Scanning
- Template retrieval and/or transmission of scanned data
- Template matching
- Transaction data recording

These in turn create specific *risks* that may be incident on the parties:

- Interception
- Template corruption
- Computational error
- Falsification
- Data compromise
- Etc.

The result of these factors is to define a range of possible evolution scenarios. A scenario is a partial description of a possible future.

This approach employs a theoretical foundation that we term the "XLRM" for understanding multiple scenarios developed recently at RAND and shown as Figure 10.²⁹

²⁹ Robert Lempert, Steven Popper and Steve Banks, *Shaping the Next One Hundred Years: New Methods for Quantitative, Long Term Policy Analysis* MR-1626-CR, 2003 available at <http://www.rand.org/publications/MR/MR1626/> (visited on 25 September 2004)

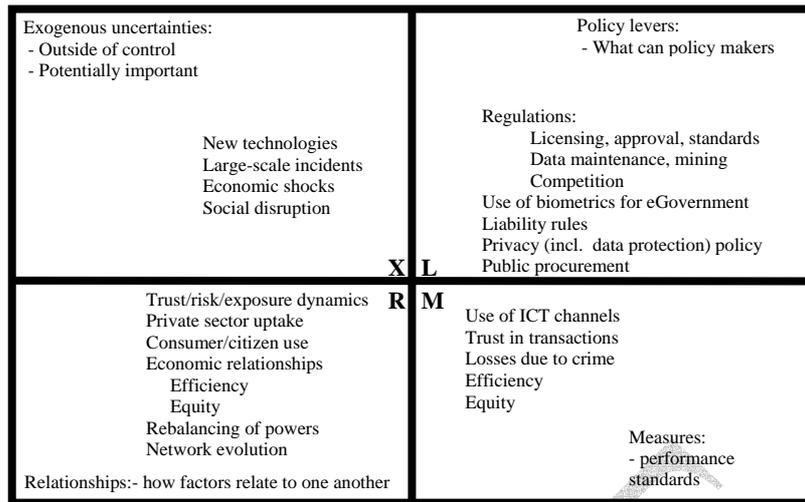


Figure 10: XLRM scenario framework

The **eXogenous** factors are ways in which the future is uncertain that are not under the control of the actors. **Policy Levers** are the options that are open to the various actors. There may be one or more actors who have the ability to posit policy; moreover, their separate policies may be jointly or separately implementable. **Relationships** are the ways in which exogenous factors are connected with each other and how policy will affect the world. **Measures** are ways of assessing the world, in order to ascertain in as quantitative a way as possible how desirable or undesirable any scenario is from the point of view of any of the actors.

The XLRM structure was originally developed for use in exploring very large numbers of scenarios. Many dimensions of exogenous variables are considered, which in combination can lead to literally millions of alternative futures. A computerized model is constructed that contains the relationships calculates the measures. Policy levers are then applied to the model to obtain values for the measures. Because of the multiplicity of dimensions and therefore scenarios, the results are displayed collapsed across dimensions to show how major measures behave according to variations in two or three dimensions at a time.

For present purposes, scenarios combine the exogenous environment and the relationships. Those using the scenarios generate their desired policy levers, as well as ways in which they measure the desirability of the scenarios presented to them.

The first step in constructing the modelling base for scenario construction and modification starts is the assembly of a listing of the salient components, combined in a generic top-level logic model of the main transactions (the main actors and the main types of transaction affected). The second step is to identify the main (XLRM) feature(s) of the scenario(s) actually under construction. Where relevant, these aspects may be calibrated or fleshed out by empirical data projected into the future using estimated or hypothesized relationships. The third step is to choose the main 'corners' of scenario space – identifying the 'big idea' for each scenario, and the main implications for the elements. The final step is to construct the scenarios themselves. This involves identifying the distinguishing characteristics of the scenario and tracing through the main elements of the model to identify important aspects, the order in which they are altered, and the spillover changes elsewhere. This process creates a 'shadow model' of impacts. The use of the models can enhance the 'concept' for the

narrative, the story-board logic of tracing effects and cross-linking and the use of numerical projections for colour and instantiation.

Network effects

The adoption of biometric technology offers several very strong externalities. One is *interoperability*. Many industrial organisation economists (Katz and Shapiro, Economides, etc.) refer to this as a network externality, but the concept is at once more complicated and simpler than networking in the ‘graph’ sense considered below. Interoperability is concerned with the formation of groups whose internal structure is almost irrelevant. What has a profound influence is the *complementarity* among users of compatible technologies. Two specific implications are a tendency towards standardisation and a ‘tipping equilibrium’ tendency for markets with interoperability externalities (e.g. biometric identity management systems) and markets complementary to them to be ‘captured by a small number of firms capable of exercising enormous market power. In the area of biometrics, tipping may be in the services, hardware or software markets, or via dominance of standards or a very concentrated holding of key patents). It may have its origins in organisational interoperability – for instance, the procurement contracts associated with large public-sector procurements can result in an enormous installed base of users of a particular solution, which will encourage following clients – particularly those whose operations require interoperability with the public sector lead customer) to adopt the same or similar biometrics, and will encourage suppliers of complementary biometrics-related goods and services (see Figure 2) to also standardise on the emerging dominant incumbent. Tipping, like standardisation, can affect compartmentalisation. This could increase if different segments are dominated by different players, but is most likely to decrease, which in turn can magnify the impact of system failures or attacks. In terms of the underlying progress of technology and implementation, tipping can produce an adverse trade-off between interoperability and cost-reducing process innovation on one side and diversity, robustness and functionality-enhancing product innovation on the other. More broadly, markets with strong complementarities may have inefficient equilibria or no equilibrium at all.

A second externality is *interconnection* or *linkage*; individuals or firms adopting a particular kind of BM technology in effect join (or move closer to) others using the same technology. This does not mean that they cannot otherwise be connected, but only that, for some purposes, the biometric connection is ‘stronger.’ It may offer lower transactions costs, or a different incidence of costs. It may also change the type of network (the structure or pattern of linkage, the degree of directionality or durability). In addition, it may change the dynamics of network change³⁰. Models of networked behaviour show asymmetries in individuals’ ‘power³¹’, the possibility of forming ‘small worlds’ clusters in contrast to broadly connected networks, and a tendency for stable networks to be inefficient and *vice versa*. In addition, network effects depend as much on indirect as on direct connections. Thus, in the example discussed on page 19, one form of false identity was used to trigger the acquisition of others and to commit frauds in remote contexts.

³⁰ Many models of network formation assume that links are formed when both parties wish to form them, but can be broken unilaterally.

³¹ This can be seen as a form of tipping where a few well-connected players emerge as ‘hubs;’ even if they do not control large market shares, they may have disproportionate influence on efficiency.

Public goods

Of course, neither of these externalities are intractable – nor are their adverse effects inevitable. Network access is something of a *public good* – even if it is subject to e.g. congestion effects (whether these involve delayed transactions or escalating fraud or accident costs) effective competition can nonetheless produce substantial benefits. The trick is to determine efficient provision (how much identity, security, privacy and interconnection is efficient), feasibility (how to induce society to pay the costs of this level of provision) and equity (how to align costs and benefits among the participants). In the absence of externalities, these tasks are all performed simultaneously by competitive markets. Identity in general and biometrics in particular face a number of specific challenges. They are information goods; they cost little or nothing to reproduce, but are difficult to verify without being disclosed. They are produced, mediated and used along extensive supply chains, which makes it difficult to fix liability for error, negligence or fraud on the technology provider, the system operator, the user, etc. Moreover, it may be difficult to monitor or enforce compliance with standards, privacy regulations, etc..

Evolution and intellectual property

The development of biometric (and other) identity management solutions is the result of an evolutionary process. All such processes have three fundamental components.

- Variation (here, innovation)
- Selection (here, regulatory, commercial and public acceptance and the challenges of accident and opportunism)
- Heredity (here, codification, embedding in durable systems and cumulative innovation with backwards compatibility).

Together, these forces determine the nature and effects of societal development. Variation may be viewed as the deployed results of R&D and a range of formal and informal innovation practices. According to the conventional argument, society must reward innovation in order to get the right kind and amount. Because markets simultaneously reallocate outputs and reward production, it is usual to create markets in the fruits of innovation. These take the form of intellectual property rights. [standard discussion]. Patents vs. secrecy or NDAs or bundling. Spillovers between the market for goods and services and the market for innovation.

Open source software replaces the private good paradigm with a limited form of public good paradigm. Open source vs. open standards;

Security

By increasing the accuracy of identification, biometric identity management solutions can increase security or assurance. But they may decrease security or produce uneven benefits. In this context, a secure system is one whose continued effective function can be relied on in the face of a range of uncertainties. Effective functioning includes:

- The identities of qualified personnel will be recognised
- The identities of unqualified personnel will not be recognised

- Information held by the system will not be compromised, corrupted, disclosed inappropriately, etc.

As a result, security is likely to be multidimensional. A decrease in the rate of false positives does not necessarily entail a reduction in the rate of false negatives and the *information security* of a system may not coincide with the security of the *system's function*. The weighting of these dimensions is also likely to differ from party to party. A system operator may care more about false positives (letting in unauthorised personnel) than about false negatives – thus an increase in security from his point of view may decrease it from a user's point of view. Definitions of information security may effectively translate to immunity of system information from unauthorised viewing or to immunity from unauthorised modification. Different users may care more about one than another. Some may care exclusively about functional security – will I make my plane if I rely on the airport's iris-scan to allow me to skip the immigration queue? Others may be more concerned with information security – will the recorded iris template information be used to create a false identity tied to me?

The aspects that are optimised are the ones most important to those whose actions control biometric application deployment, taking costs and liabilities into account. A system operator who faces relatively little competition (e.g. a public sector agency) and substantial liability for false positives might be expected to deploy a system with greater redundancy or intrusiveness than another operator without those constraints. An operator facing substantial competition to serve individuals primarily concerned about security of access might deploy a system with less security. An operator whose system had to interact with complementary systems might produce a system with greater protection towards the outside and lower internal security protections than one who could not profit from locking in producers of complementary systems³².

Security is additionally a *subjective* feeling in respect risk, so it is path-dependent and sensitive to external shocks – in other words, it may depart systematically from objective risk levels. Like other judgements under uncertainty, it may: not reflect a 'rational' separation of likelihood and consequence, weight very large or very small probabilities inaccurately, place undue weight on the experience of others; or conflate objective risks of accident or systemic failure with 'subjective' (e.g. intentional and thus endogenous) risks due to attack.

One consequence is that levels of security – and thus use of secure channels in preference to an insecure channel – may not behave smoothly in response to changes in e.g. perceived risks. This possibility and the consequences for adoption and levels of security are illustrated in the model in the annex and shown in Figure 4 on page 16.

These considerations complicate the objective measurement or valorisation of security, which are sensitive to fine details, starting from technological characteristics but stretching well beyond them. It further suggests caution about some 'obvious' principles – for instance, that more security or trust is better. However, individual levels of security could be measured and allocated efficiently if, e.g., there were a full – or sufficiently rich – set of markets for risk or liability

³² R. Anderson, "Cryptography and Competition Policy: Issues with 'Trusted Computing'" at <http://www.ftp.cl.cam.ac.uk/ftp/users/rja14/tcpa.pdf>.

The uncertainties include: accidental breakdown; systemic failure; and intentional attack. These are assessed, transferred and managed differently from each other.

As a systemic property, security depends on the alignment of risk tolerance, ability to manage or affect risk and ability to take further action in response to insecurity. Security is thus connected to *resilience*: for example, if strong security is available via a biometric solution, it is tempting for the parties to economise on other precautions such as the use of skilled or experienced personnel. Failures may be less frequent, but their consequences may be more severe. In addition, if the skilled personnel are more efficient at other aspects of their job, system performance may suffer in other ways.

Trust

Biometrics can build trust by increasing parties' mutual knowledge of each others' identity. Trust is a key element of social capital but carries the potential for abuse. The acceptance of biometric identification requires trust on the part of the user and may reduce the scope for trust on the part of the system operator. The user is typically interacting with an IT system rather than a human being, and must trust it in several ways:

- The user must trust the system to confirm his claim of identity or identify him correctly;
- The system operator must trust the system's report enough to act on it;
- In the case of a purely automated system, the user must trust the integrated system to act correctly on his behalf (this may include trusting the system to identify others with whom it may transact)
- The user must trust information provided by the system (following identification) enough to act on it
- The user must trust the system (or its operator) not to misuse the template information provided or other information relating to use of the system (e.g. transactions records)

From the economic point of view, trust is sometimes represented as an incomplete contract – in this perspective, trusting and trustworthiness work in the opposite direction to assurance and certification³³

Privacy

Concerns relating to biometrics have economic consequences insofar as individuals may attempt to secure their privacy by 'opting out' of economic activities secured by identity management or by resisting attempts to integrate identity through compartmentalisation. As a scenario variable, it would be interesting to explore different models for legal protection – different levels and types of privacy and the existence or otherwise of fundamental privacy rights. This is particularly critical as regards globalisation, since the European Union has already seen the consequences for financial transactions of differences in personal data protection between the EU and the US and, more recently, as regards disclosure of passenger data relating to transatlantic air flights. The economic consequences are fairly straightforward and

³³ Just as security and privacy may be seen as opposing.

take the potential form of changes in levels of transatlantic (or other) travel and in the allocation of (especially electronic) commerce revenues.

Commerce

One potential implication of biometric identity is enhanced and 'delocalised' access to a range of services, especially utilities. For example, the concept of utility computing is now widespread, but concerns remain over the security of this service and the degree to which individuals would have access to their data and programmes while outside their normal secured environments. As mentioned earlier, enhanced identity will also affect retail commerce in a number of ways, including: enhanced forms of electronic contracting and payment; enhanced consumer protection; simplified search and price comparison and expanded possibilities for electronic marketplaces. The impact on the effectiveness of competition depends on a range of scenario-specific choices such as: whether 'personalising' data relating to individuals belongs to them or to their trading partners (as at present) and whether biometric implementation favours or penalises competition from SMEs and firms in remote jurisdictions. From the public policy point of view, it is possible that enhanced levels of identification can improve the efficiency (both higher compliance and reduced cost) of VAT collection.

Work

As the work environment becomes more security conscious, control of access receives greater emphasis. At the same time, increasing information-richness of productive activity and new methods (locations) of work mean that individuals need greater remote access to information and communication resources than ever. It is almost inevitable that biometrics will play a central role in this – indeed, without productive development it is unlikely that the full potential of teleworking could be realised. At the same time, the inhibiting factors noted for commerce and citizen-government interaction are at play in this environment, and the regulatory and contractual environments will have a strong influence.

Societal discourse

The reciprocal impacts of biometrics and societal discourse are not primarily economic, but rather reflect concepts of privacy, identity and anonymity. However, it is worth recording that public debate on a range of policy issues will doubtless be influenced by economic considerations.

Under this general heading, it should also be noted that the diffusion of biometrics will naturally spread to many aspects of societal interaction. In large part, this process will be driven by user economies of scale and scope: not only is it easier to carry a single strong form of identification, but it also becomes simpler for the individual to keep track of their own traces. Some privacy concerns could be allayed, for instance, if individuals had a record of where and in what circumstances their biometric identification was invoked. Particularly in a centralised deployment, it would be possible to apply the same access, accuracy and approval protections to data combination and transactional data as are now applied to data collection and re-use. This could also greatly enhance protections against the theft of legitimate identities (as opposed to creation of false or multiple identities).

It is also important to stress that the technological aspects of biometrics will not wholly determine either the future trajectory or its impacts. The degree to which

consumers (or system operators) have a choice of providers and technologies will not only influence the combination(s) of technologies in actual use, but also the gap (if any) between cost and price and the extent to which biometric and other applications will converge. The evolution of novel technologies and novel uses will be mediated by market competition, standards and policy. These could produce either virtuous or vicious cycles; any biometrics implementation involves both changes and shift in cost, and it is not always obvious whether the resulting identity management arrangements will become simpler, more specific, more reliable, etc.

5.2 Competitiveness and efficiency

Competitiveness

5.3 Equity and distribution

As mentioned, the costs of different biometric implementations do not fall evenly across sectors or across different sized or organised firms within sectors. In rough terms, the fixed costs associated with database and secure communications infrastructures mean that centralised systems (e.g. those used for user identification and multiple-identity checking, or used where local template copies or match processing are not practicable) are likely systematically to favour large firms over small and networked or affiliated firms over independent firms. This is likely to affect economic competitiveness, the density and availability of the services offered by such firms, the employment and other local economy benefits of their operation and the general efficiency gains associated with effective competition. To the extent that such firms survive by using less secure means of identity verification, their cost and market presence disadvantages are multiplied. The disadvantages facing SMEs may be further magnified by proprietary software coupled with strong ‘network externalities.’ If an ‘identity divide’ opens up in, say, a retail market, it may spread to further divisions along sectoral, geographic or other lines, and may slow the acceptance of biometrics among the populations who trade with these firms. Such an outcome is highly likely without active policy intervention: not only does the complementarity engendered by interoperability favour concentration (see Section 4.5) but the connections of such firms with each other are likely to follow a ‘power law’ in which a few central firms dominate a periphery of less well-connected firms³⁴.

On the consumer side, the competing public and private uses of biometrics can also have distributional consequences. For instance, in countries where biometrics are lead by passport applications rather than identity cards or private sector applications, or where private sector applications are primarily associated with high-service credit card applications, they will first be used by those in higher socioeconomic brackets.

6 The role of policy

Acting as ‘launching customers,’ governments have been among the prime movers in the development and deployment of biometrics. This role seems set to continue for some time to come. However, other aspects of policy can also contribute to the development of the field and the management of potential problems and concerns. This section considers some of the problems and possible avenues for policy to

³⁴ See e.g. D. Watts *Small Worlds*. The result follows from a model of identity network formation in which firms weakly prefer to link to firms whose information-sharing networks are larger.

address them. It does not try to deal with non-economic issues of public trust and acceptance.

6.1 Potential areas for policy intervention

One set of potential problems comes from the ‘tipping’ tendencies of economic competition (see section 5.1): too few firms, excessive market dominance; slow or distorted technological development; high prices for hardware and software³⁵; possibilities for overt or tacit collusion among suppliers and integrators. On a more positive note, pro-competitive policy could help promote diversity (in approach, firm size and application area); balance scale and scope economies with economic efficiency; restrain vertical foreclosure (footnote 21, p. 26) whilst encouraging appropriate integration; and encourage product and process innovation.

A more general or classic basis for policy intervention comes from the ‘public good’ aspects of identity. Society at large benefits from stable and reliable individual identity, even when the ‘benefit’ to the individual is small or negative (as with criminal activity). It is thus likely that identity will be under-supplied and there is scope for public policy to encourage ‘opting in.’ However, identity is not necessarily a ‘pure’ public good; as discussed above, there are dangers to both the individual and the system if the identity management structure becomes too comprehensive³⁶ or is not sufficiently secure, transparent and accountable. Identity is also subject to congestion; the efficiency and effectiveness of identification (especially for large populations) is markedly less than that of identity verification. Thus for, security, equity and efficiency, a degree of compartmentalisation is necessary, which may require public policy support. More generally, the expanding use of biometrics may impose a range of externalities which policy can help to internalise (e.g. by liability rules).

Beyond the biometrics or identity management sector itself, we can already see ‘ripple effects’ with which policy should be concerned. For instance, the same foreclosure risk that connects e.g. a firm with a monopoly on iris scanning algorithms to providers of iris-based biometric solutions may extend to public or private sector users of these products. One particular issue concerns the effective competition among different channels for commerce. It is conceivable that competition and consumer protection policy should address the differential impact of biometrics across competing (e.g. physical and electronic) channels. A related issue is the allocation of costs associated with identity verification along the market chain. We know that such costs tend to accumulate where parties have limited substitution possibilities, regardless of benefit. Thus it may be that consumers have to bear the costs of identity management systems intended primarily to protect sellers³⁷.

Finally, the public sector has its own economic objectives in regard to biometrics. Generically, governments seek to maintain competitiveness by supporting domestic research and production capacity and further by encouraging practices that improve the efficiency of economic transactions. They also seek to increase the value for

³⁵ And, as a result, slower-than-optimal decline in costs.

³⁶ This can raise privacy concerns – especially around data mining), issues of compromise of personal data and difficulties of repudiation: if identification becomes more accurate, the probability of error drops but the severity of error increases.

³⁷ Typically, a seller is more concerned with false positives and a consumer with false negatives.

money represented by public eGovernment procurement programmes. In pursuit of both goals, procurement policy is an important tool. The ‘launching customer’ role can stimulate supply and influence R&D, but in the case of public-facing procurements (esp. national identity or entitlement cards) it can stimulate acceptance of biometrics and stimulate demand as well. Moreover, eGovernment biometrics implementation is not limited to public-facing changes, but can improve the efficiency and accountability of back-office transactions in government as in any other large organisation.

6.2 Policy levers

Regulation

There are a range of regulatory functions and competences that can directly address these issues. The most obvious is competition policy, since tipping can lead to structural market failure and because foreclosure, predation, tie-in arrangements and collusion are failures of market conduct. The effectiveness (and political willingness to use) such tools vary from country to country, and there are complexities in defining market boundaries and measuring costs and even prices, but the procedures for developing policy in this area are straightforward. One particular aspect calling for EC leadership is the international dimension (whether coming from overseas ownership of key property rights or the insistence by overseas governments on use of proprietary or non-neutral standards for biometric identification). Competition policy may be a valuable tool in addressing this issue, either internally or via organisations such as WTO and WIPO. A second issue is the need to balance market control via market share with market control via dominance in related markets (esp. the ‘market for innovation (IPR) and the systems integration market).

A related set of regulatory fora is provided by sector-specific regulations, especially in the areas of telecommunications (pricing, approval, licensing, mergers, etc.), health care (esp. medical record collection, exchange and maintenance) financial services and transportation.

Finally, a range of consumer regulations address specific issues associated with private sector use of biometrics; principally consumer protection (for retail and credit transactions) and personal data protection (for privacy-related issues).

Laws

On the legal front, it is worth stressing the importance of policy relating to intellectual property rights (especially mutual recognition of overseas IPR and the legal status of copyleft and general public license), law relating to the position and powers of standardisation bodies, and contract or commercial law. In particular, there have been discussions of biometrics in relation to digital signatures and certification and in terms of the enforceability of contracts that mandate biometric identification or impose ‘contracts of adhesion’ relating to biometrically-secured information.

A further ‘legal’ policy area with strong economic implications is liability: wherever the activities of individuals or market participants affect others not party to the transaction, it is likely that the private outcome will diverge from the social optimum. Liability rules are the device by which stakeholders are encouraged to take account of these impacts. In many cases, they are designed to correct informational asymmetries between buyer and seller (much of consumer protection is concerned with this). But in other cases they are aimed at inefficient separation of incentives and control. For

instance, the technology of a biometrics solution influences the extent and type of errors, but the specific implementation also makes a contribution and determines the severity of consequences. If the technology provider bore full liability the result would be greatly retarded and inefficient development, but if the user bore full liability only the simplest and most transparent methods would be used. Optimal one-sided approaches put the liability on the 'least-cost avoider' of the problem³⁸, but intermediates such as joint and several liability or options built into the contract can provide greater efficiency and better incentives to reduce (rather than compensate) problems. It is also worth remarking that the same comments hold for unplanned benefits as for unplanned losses – where the deployment of biometrics can lead to additional benefits through rearrangement of working arrangements or rebundling of goods and services, there is scope for public policy to consider supporting innovative contract forms.

International policy

On the international front, the relevant policy tools include trade policy, policies relating to exchange of data and security policy. These have a direct impact on the ability of firms to gain fair access to international markets, and also on the international mobility of both capital and labour.

Support for R&D

Public support for R&D, as always, has a role to play in the development and deployment of this technology. There is little that is specific to biometrics: it is necessary to tune such support carefully to avoid distorting competition and to select projects with a reasonable mix of public and private benefits. Many of the existing vehicles for technology development can be applied in this area. In addition, lessons are being learnt from the modalities of research and development support found in the military context, where biometrics has already been substantially elaborated.

Procurement policy

These lessons reflect the general importance of procurement policy. In this connection, the new EC public Procurement Directives provide a more 'innovation-friendly' structure than heretofore. In particular, there is scope for allowing procurement authorities to choose between standards or equivalent performance in drawing up specifications, in engaging in certain forms of 'pre-competitive engagement' with research and industry stakeholders, in using multiple-sourcing and design competition to balance innovation with cost-effectiveness, etc. Other aspects of procurement policy include the use of intellectual property options in procurement contracts (to encourage maximal effectiveness of wider deployment and subsequent development), open standards requirements and insistence on open and transparent supply chain management. This recognises the enormous effect of large government contracts, which are often the first major demand component, underwrite private financing and create industry leaders in a short space of time. While interoperability generally makes it impossible to divide up procurement among many different firms until open standards have been developed, the procurement can be structured in such a way as to leave even 'losers' of tender competition with valuable intellectual property

³⁸ Brown (1973).

and to provide opportunities for integrators, licensees, etc to participate in future development.

In considering the use of procurement, it is essential to develop sound estimates of costs, effectiveness, uptake and knock-on effects – on the private market and on other parts of the public sector.

Costs fall in different phases – at a minimum; development, initial deployment and mature operation. Some of the costs are unrecoverable, while others can be partially recovered – through operational savings, user fees or the benefits of wider deployment. Some costs are ‘fixed’ –independent of operational scale, while others vary with deployment and/or utilisation. For these costs, good predictions of coverage and utilisation are essential.

Note that deployment and utilisation are definitely *not* the same for public services. Early assessments of eGovernment services provided through web sites show availability well in excess of demand, and thus that utilisation rates are well below capacity. This gap may close over time, and will necessarily do so as alternative channels are closed. However, both the public and private sectors should be concerned about any underlying mismatch between demand and supply. In the case of the UK biometric identity card, for instance, there will be no way to avoid the costs of the card – or to avoid using it in order to get services. But for marginal populations the gap may be enough to cause them to opt out. Even for inframarginal groups, the costs may exceed the benefits – but without an alternative there is no *a priori* way to ensure efficient provision.

On the other hand, the cost basis for biometric programmes is improving continuously, as governments move towards resource accounting (which allows accurate tracking and analysis of fixed and variable costs and of spillovers to other parts of the system).

7 The economic future of biometrics

This section provides some final considerations on the unfolding of biometric futures from the economic point of view.

7.1 The market for identity

In the realm of identity verification, it is likely that a distinct market niche will emerge based on the connection of identity to increasingly personalised economic ‘access’ – the conversion of personal life to a paid-for experience foreseen by Rifkin (2000). The collection of cumulative experience and preference information and its protection (and indexing) by precise identity information can facilitate matching – but also make consumer comparison or search harder. By analogy with the information assurance market, we can expect this identity market to be dominated by intermediaries and integrators unless this is prevented by e.g. privacy regulation. What is fairly certain is that, as the economic value of identity increases, there will be a struggle for ownership of the essential technologies and facilities and for the data themselves (by analogy with e.g. credit ratings).

In addition to the ‘white’ market, there will also undoubtedly be ‘grey’ or ‘black’ markets for false, multiple and limited or partial identities. The more society relies on strong identification as a substitute for other parts of enforcement and monitoring, the greater will be the rents earned in these markets. There may even be a degree of

'regulatory competition' between nations seeking to attract profitable enterprises by offering 'light-touch' regulation of biometrics.

In an ideal world, the collection and encoding of biometric information would be kept open and competition would drive improvements in matching and information management.

One main uncertainty is the extent of convergence – will we have a single 'public' biometric identity or many 'proprietary' ones? Another is whether identity management (particularly for identification and multiple identity screening) are sufficiently scalable to permit widespread convergence.

As the market expands, issues will arise regarding the scalability of its desirable and undesirable characteristics; changes in incentives; selection effects; and the strategic evolution of strategies.

As pointed out by Grijpink (2004) and others, the balance of costs and benefits associated with a biometrics implementation may not be scalable. As the use of the system spreads, its ability accurately to balance stakeholder interests and incentives changes, the attractiveness of the system to attack, its vulnerability and the likelihood that problems will not promptly be recognised also increase. So, too, do the potential deadweight losses associated with technology or IPR dominance. On the other hand, the proliferation of small, bespoke systems may carry a large overhead (especially for implementations that rely on user-provided 'local' templates), inhibit interoperability and limit the potential returns to new technologies. One could therefore imagine two 'futures' differing in the degree to which 'communities of interest' defined by biometrically-secured mutual identity are, in fact, cross-linked.

A further aspect of biometric evolution that will influence the 'joined-up' character of the system is the course of learning and the formation of conventions. As for learning, biometric identity, like all other forms of technologically-based security, is essentially a dynamic game in which the important players are those who supply identity management services, those who demand them, those whose identities are verified or ascertained, and those who would subvert the process for purposes of identity fraud. Among the first three players, the critical aspects are the alignment of responsibilities and liabilities with incentives, information and capability. The arrangements should not place decisions on those to whom they do not matter, those who do not have the necessary powers of action or knowledge, or those whose actions are prone to errors that others cannot efficiently guard against. Between these three parties and the 'defectors' the game is more dynamic – it is not just that technologies can be circumvented, but more importantly that the way in which they are subverted and the returns to doing so depend as much on how the technologies are 'socialised' – and how much they are relied upon. In addition to the type I and type II errors (false positives and negatives) referred to above, there is a danger of type III error – solving the wrong problem precisely. At the moment, we tend to think of identity fraud in terms of individuals assuming false identities, usurping other people's identities in order to gain access to goods and services without payment or authorisation, and the creation of multiple partial identities to hide the full scope of illegal or unethical activity. Given the costs of false positives and false negatives and the possibilities of the technology, it is possible to define an identity hierarchy, with higher and more burdensome levels of assurance offered in areas where the risks or exposure are higher. But new technologies will change the nature of criminal attacks. For instance, it has been observed that the proliferation of surveillance technologies (including

potentially, a biometric component for identification) runs the risk of displacing crime from public areas to more private spaces, or to on-line environments. In addition, a strong biometric system could give rise to a form of ‘denial of identity’ blackmail, similar to the ‘denial of service’ blackmail currently in vogue. This makes the game of ‘catch-up’ harder to play; it can also lead to identity crime played for higher stakes than those of the arena where identity was compromised. The ultimate result may be an undermining of trust. Given that the ‘defectors’ may have a range of motives (from the venal to the political), that they adjust with a lag and that societal recognition and mitigation of harms take time, the best policy may be *festina lente* – to make haste slowly.

Among those for whom biometrics is intended to advance a common interest, it should also be recognised that the acceptability of identity is a matter of societal convention as much as of law or contract provision. For instance, in California, the default form of documentary identity consists of a California driver’s licence and a major credit card – even for transactions that involve neither driving nor credit (like check cashing). As a result, both the state and the credit card providers have developed ‘attenuated’ forms of these identity documents: driver’s licences that do not give the right to drive and credit cards with minimal credit limits. This tendency for identity to diffuse follows linked societal groups and the transactions where identity must be verified. We might therefore expect that dynamic analysis of the formation of social conventions³⁹ would be useful for describing both the diffusion of biometrics and the purposes for which they wind up being used.

7.2 The impact of identity on market outcomes

Currently, markets have a broad range of degrees of ‘identification.’ In markets for durable goods, especially where there is potential need for buyer-seller contact over an extended period of time (e.g. real estate, financial services) levels of identification are high. In other cases (e.g. public markets, trade in used commoditised goods, illicit trade) formal identification is quite low. Part of this is no doubt cultural, but part reflects the balance of costs and benefits associated with identification on one side and the uses to which it is put (non-repudiation, warranty, etc.) on the other. Expanded identity can bring added security to one or both sides, but can also restrict choice and impose costs.

Consider the case of cash transactions. Their drawbacks are that they offer the buyer little or no recourse, that they are difficult to monitor and tax and that they do not promote efficient competition when the characteristics of goods are not common knowledge. On the positive side, they are fast and low-cost, expose the parties to little risk of identity theft and carry no danger of repudiation. As biometric costs fall and reliability, accuracy and acceptance improve, this balance can be expected to change and with it, the prevalence of high-identity transactions. If it were simply a matter of economic efficiency, the combination of biometrics with advances in ICT would probably spell the end of the ‘cash economy.’ But it is by no means certain that individuals who choose to use non-biometric forms of identification will not opt out of a biometric world, leading to an expansion of the ‘grey economy.’

³⁹ Young (1993, 1998)

8 References

- Acuity Market Intelligence at: http://acuity-mi.com/Industry_Evolution.html
- Anderson, R. (2001) "Cryptography and Competition Policy: Issues with `Trusted Computing`" at: <http://www.ftp.cl.cam.ac.uk/ftp/users/rja14/tcpa.pdf>
- Anderson, R. (2002) "The economics of trusted computing" at: http://www.netproject.com/presentations/TCPA/ross_anderson.pdf
- Barabasi, A. (2002), Linked: The New Science of Networks, Perseus, Cambridge, Mass.
- Biometric Technology Today, Market Reports 1999-2003.
- Brown, J. (1973), "Toward an Economic Theory of Liability", *Journal of Legal Studies*, 2 (June), 323-349.
- Cave, J. (2003) "Trust: how much is enough?" Under submission, *Information, Computing and Society*.
- Cave, J. (2005) "The economics of cyber trust between cyber partners" in R. Mansell and B. Collins (ed) *Trust and Crime in Information Societies*, Cheltenham: Edward Elgar, 380-428.
- Clarke, Roger (1994) "Human Identification in Information Systems: Management Challenges and Public Policy Issues" at: <http://www.anu.edu.au/people/Roger.Clarke/DV/HumanID.html>
- Economist, "Prepare to be scanned" *Economist*, December 4 2003, at: http://www.economist.co.uk/displaystory.cfm?story_id=2246191
- Ferrando, C. (2001) "Survey of Biotechnology Practitioners" Georgia Institute of Technology, unpublished report.
- Gordon, L. Loeb, M. Lucyshyn, W. (1999), "An Economics Perspective on the Sharing of Information Related to Security Breaches: Concepts and Empirical Evidence", presented at Workshop on Economics and Information Security University of California, Berkeley May 16-17, 2002, at: <http://www.sims.berkeley.edu/resources/affiliates/workshops/econsecurity>
- Grijpink, J. (2001), "Biometrics and Privacy", *Computer Law and Security Report*, **17**(3), 154-60.
- Grijpink (2004a), "Identity fraud as a challenge to the constitutional state", *Computer Law and Security Report*, **20**(1), 29-36.
- Grijpink, J. (2004b), "Two barriers to realising the benefits of biometrics," mimeo.
- Guardian newspaper report at: <http://politics.guardian.co.uk/homeaffairs/story/0,11026,1342061,00.html>
- International Biometric Group, <http://www.biometricgroup.com/>
- Jackson, M. and Watts, A. (2002), "On the formation of interaction networks in social coordination games", *Games and Economic Behaviour*, **41**, 265-291.
- Kahneman, D., Slovic, P. and Tversky, A. (eds) (1982), Judgment Under Uncertainty: Heuristics and Biases. Cambridge University Press.

- Katz, Michael L., and Carl Shapiro (1994). "Systems Competition and Network Effects," *Journal of Economic Perspectives*, 8:93-115.
- Lempert, R. S. Popper and S. Bankes, (2003) *Shaping the Next One Hundred Years: New Methods for Quantitative, Long Term Policy Analysis* MR-1626-CR, 2003 at: <http://www.rand.org/publications/MR/MR1626/>
- Machina, M. (1982), "'Expected Utility' Analysis without the Independence Axiom", *Econometrica*, **50** (2) 277-323.
- Most, C. (2003) "Biometrics and trusted Identity" in *Biometrics Market Intelligence*, **2**, 3.
- Most, C. (2004) "Biometrics and border control: beyond US-VISIT", *Digital ID World*, September/October: 18-21.
- Piper, F. M. Robshaw and S. Schwiderski-Grosche, "Identities and authentication" in R. Mansell and B. Collins (ed) *Trust and Crime in Information Societies*, Cheltenham: Edward Elgar
- Rejman-Greene, M. "Biometrics – real identities for a virtual world," at: <http://www.soi.city.ac.uk/~kam/rejman-greene.pdf>
- Rifkin, J. (2000), *The Age of Access*, New York: JP Tarcher/Putnam.
- Synovate (2003) "Federal Trade Commission – Identity Theft Survey Report", mimeo.
- The Association for Biometrics (see <http://www.afb.org.uk/>)
- The Biometric Consortium, <http://www.biometrics.org/>
- U. K. Cabinet Office (2002), "Identity Fraud: a study" at: http://www.homeoffice.gov.uk/docs/id_fraud-report.pdf.
- U.S. Department of Consumer Affairs at: <http://www.consumer.gov/idtheft/stats.html>
- US Federal Trade Commission (2003) "Information on identity theft for consumers and victims from January 2002 through December 2002," mimeo.
- US Federal Trade Commission (2004) "National and State Trends in Fraud & Identity Theft January -December 2003," mimeo.
- US National Bureau of Standards (1979) "Federal Information Processing Standard 65, Guideline for Automatic Data Processing Risk Analysis"
- Watts, D. *Small worlds : the dynamics of networks between order and randomness*, Princeton University Press.
- Woodward, J. (2003), *Biometrics and Strong Authentication*, Emeryville CA: Osborne/McGraw-Hill.
- Young, P. (1993), 'The Evolution of Conventions', *Econometrica*, 61(1): 57-84.
- Young, P. (1998), *Individual Strategy and Social Structure* Princeton NJ: Princeton University Press.

9 Annex: A simple model of biometric adoption

9.1 Description of the one-sided model

Consider a population of linked, but heterogeneous individuals, each of whom can choose between biometric and a non-biometric channel. We assume that players have the following evaluation of the biometric channel:

$$\Delta_i = \theta_i [1 + \nu h_i] - \rho_i$$

where

Table 5: Model variables

Variable	Definition
Δ_i	The payoff advantage to player i of using the biometric channel
ρ_i	Player i 's perceived risk -cost from using the biometric channel
h_i	The proportion of i 's neighbours using the biometric channel
θ_i	Player i 's idiosyncratic taste parameter, independently and identically distributed according to density $f(\theta)$ and cdf $F(\theta)$
ν	Exposure (or the 'network externality' among users of biometrics)

The 'taste' parameter θ_i measures the relative value of these benefits compared with expected costs ρ_i , which include the cost of switching to the biometric channel.¹ θ is independently and identically distributed on a compact interval $[\theta^-, \theta^+]$.

The 'benefits' term includes both a fixed component (for example, reduced transaction costs associated with trusting a biometrically secured channel) and a term that varies with the number of other biometric channel users with whom player i potentially interacts (the *network externality*) – This term can alternatively be interpreted as the degree of *exposure* created by use of the biometric channel.

We assume a fully connected network involving a large population, so $h_i = h$, where h is **adoption** - the proportion of individuals employing the biometric channel. ν measures the strength of the network interaction: if $\nu = 0$ there is no network externality, while $\nu = 1$ corresponds to the 'peer-to-peer' case where only network interactions matter. There is no asymmetry in risk cost assessments, so $\rho_i = \rho$ for all i .

9.2 Equilibrium behaviour

Nash equilibrium implies the set of biometric users is $[\theta^*, \theta^+]$, where the cut-off value, θ^* , satisfies:

$$\theta^* = \frac{\rho_i}{1 - \nu + \nu h}, \text{ and for consistency}$$

$$h = 1 - F(\theta^*)$$

Substituting for the proportion of biometric users h gives:

$$\frac{\rho_i}{F^{-1}(1 - h)} = 1 - \nu + \nu h$$

Figure 4 (page 16) plots the equilibrium values of adoption (h) against risk (ρ) and exposure (ν) where the distribution of taste is normal.

9.3 A two-sided version

Assuming fixed network/exposure effects, risk responds instantaneously to the adoption of biometrics: $d\rho/dt = \phi(h)$.

For *assurance*, risk responds to biometric adoption in the ‘expected direction’ - $d\rho/dt$ is positive (negative) if adoption is above (below) a critical value h^* where expected returns to attempted fraud just balance expected costs and the system tends to a corner solution: high adoption ($h=1$ and $\rho=0$) or low adoption ($h=0$ and $\rho=1$).

For *predation* or *opportunism* we consider the possibility that the widespread reliance on biometrics works to increase risk by displacing other costly precautions (vigilance, use of skilled or experienced personnel, etc.). Thus, $d\rho/dt$ is positive or negative if adoption is larger or smaller than h^* . If network/exposure effects are small enough for adoption (or use of the biometric channel) to respond continuously to risk ($v < v^*$), the model converges monotonically to $h=h^*$ and $\rho = [1-v+vh^*]F^{-1}(1-h^*)$. If network or exposure effects are ‘too large’ the system cycles (clockwise) in a hysteresis loop.

The model assumes users know the ‘true’ of risk, adoption and network/exposure levels but the same result can be obtained from a Bayesian learning model where subjective risk estimates are adjusted according to a dynamic equation:

$$\rho_t = (1-\lambda)\rho_{t-1} + \lambda h, \text{ or } \dot{\rho} = \lambda(h - \rho)$$

9.4 An incomplete-information model

Qualitatively similar results (roughly S-shaped time-paths for adoption and decay of the biometric channel) can be obtained from a simple incomplete information model where individuals become aware of the current differential pay-off to the biometric channel by randomly sampling the population – say by media or survey reports. Dynamics depend critically on the credibility of this information and any inherent bias. One interesting observation is that S-shaped adoption paths can arise even from a *uniform* distribution of prior beliefs.ⁱⁱ The population of biometric users at any give time is a representative sample of the distribution of tastes θ rather than an ‘upper interval’ of the form $[\theta^*, \theta^+]$. In the equilibrium model sketched above, average trust per capita is given by

$$\tau = \frac{\int_{\theta^*}^{\infty} [1-v+v(1-F(\theta^*))] f(\theta) d\theta}{\int_{\theta^*}^{\infty} f(\theta) d\theta}, \text{ which for } \theta \sim N(\mu, \sigma^2) \text{ gives}$$

$$\tau = [1-v+v(1-F(\theta^*))] \frac{\theta^* \sigma^2 f(\theta^*)}{1-F(\theta^*)}$$

By contrast, in the ‘survey’ model with the same Normal distribution, each type θ ’s propensity to use the biometric channel is $\tau = \theta[1-v-vh^*]$, so per-capita average adoption is: $\tau = \mu[1-v-vh^*]$.

We can compare these two expressions to track the response of adoption to a secular decrease in perceived risk –equivalent to a fall in the critical taste parameter θ^* at which the individual is indifferent between the biometric and non-biometric channels.

Change in τ as θ^* falls	Equilibrium	Survey
No network externalities ($v = 0$)	$\frac{\sigma^2 f(\theta^*)}{1 - F(\theta^*)}$	μ
Peer-to-peer ($v = 1$)	$\sigma^2 f(\theta^*)$	μh^*

ⁱ An equivalent formulation would assume quadratic utilities; treat ρ_i as the (subjective) variance of returns, and θ_i as an inverse measure of risk aversion.

ⁱⁱ Jensen (1982) obtains a similar result in a technology diffusion model.

DRAFT