Attachment 2:

Why does Australia have negative interchange for EFTPOS?

9 July 2004

Why does Australia have negative interchange for EFTPOS?

Contents

1 Introduction

Frontier Economics has been asked by the Australian Merchant Payments Forum to write a Report explaining why Australia has a 'negative' interchange arrangement for EFTPOS whereas much of the rest of the world has a zero or positive interchange.

During the last two decades, writers in many different fields of economics have tackled the issue of why particular standards or institutional arrangements come to be regarded as settled and why these standards or institutional arrangements vary from one place to another. These literatures have many elements in common – such as increasing returns, externalities, positive feedback loops and path dependence. The recent decision by the Australian Competition Tribunal gives a history of the development of EFTPOS interchange in Australia that is remarkably consistent with this literature. This Report consists of an argument that the literature on self-reinforcing systems provides valuable insights into why the Australian system has continued as it has.

This Report is arranged in three substantive sections. In the first (section 2), we provide a brief guide to this, rather diverse, set of theories and we give some examples of how the elements that are common to these theories have been applied. Section 3 shows how many of the elements of the theory are commonly said to apply to interchange arrangements for EFTPOS. The issues raised in other applications of the theory are the same issues that have been raised about EFTPOS interchange. The final substantive section of the Report (section 4) offers a few tentative suggestions as to why other countries may have adopted different systems.

2 Multiple equilibria and path dependence

2.1   UNIQUE EQUILIBRIA AND CONVEXITY

Until the last twenty years, much of economic theory was founded on the idea that social interactions could be explained by models, each of which had a unique equilibrium. An obvious example was the unique equilibrium of a competitive market. Supply and demand curves were always drawn with appropriate slopes so that they only crossed once; and the point of intersection was the unique price and quantity that would be traded in the market. Each firm in the market would be assumed to have an upward-sloping marginal cost curve so that the output of each firm in the market would also by 'solved' by the model.

The need to produce models with unique equilibria generally required the mathematical assumption of convexity. Translated into social outcomes, convexity had the nice property that the solutions to the resource-allocation problems in the models were intermediate, moderate solutions. The assumption of convexity ensured that households spent their budgets on mixed baskets of commodities; that countries produced mixed bundles of goods and that markets were not monopolised by a single firm. The assumption of convexity was not based on any generalisations of human or social behaviour: it was based on the perceived need to structure models so that each produced a unique equilibrium.

The way economics 'explained' a phenomenon was to construct a model with a unique equilibrium. Indeed, the uniqueness of the equilibrium was widely considered to be a characteristic of an acceptable theory. Joseph Schumpeter, the outstanding twentieth-century historian of the discipline of economics, spoke for the overwhelming majority of the profession when he explained the relationship between 'explanation', 'meaning' and 'equilibrium' as follows:

> If the relations which are derived from our survey of the 'meaning' of a phenomenon are such as to determine a set of values of the variables that will display no tendency to vary *under the sole influence of the facts included in those relations per se,* we speak of equilibrium: we say that those relations define equilibrium conditions or an equilibrium position of the system and that *there exists* a set of values of the variables that *satisfies* equilibrium conditions. This need not be the case, of course – there need not be a set of values of variables that will satisfy a given set of relations, and there may exist several such sets or an infinity of them. Multiple equilibria are not necessarily useless but, from the standpoint of *any* exact science, the existence of a 'uniquely determined equilibrium (set of values)' is, of course, of the utmost importance, even if proof has to be purchased at the price of very restrictive assumptions; without any possibility of proving the existence of uniquely determined equilibrium – or at all events, of a small number of possible equilibria – at however high a level of abstraction, a field of phenomena is really a chaos that is not under analytic control.[1]

---

[1] Joseph A Schumpeter, *History of Economic Analysis,* George Allen & Unwin, 1954, p 969.

At the time of his writing (half a century ago), Schumpeter did not need to debate his propositions. It was sufficient to state them, because 'of course' they were obvious to all well-trained economists. During the 1950s and 1960s, a few (left-wing) economists at Cambridge or Oxford did not accept these propositions as obvious. These economists persisted with the (Marshallian) idea that important areas of economic activity were characterised by increasing returns rather than by convexity; but these researchers were regarded by the rest of the profession as heretical or even as mad.[2]

2.2   THE ELEMENTS OF SELF-REINFORCING MECHANISMS

In the half-century since Schumpeter wrote his *History of Economic Analysis,* mainstream economists have identified and modelled self-reinforcing mechanisms in a large range of economic activity. These mechanisms generally incorporate some or all of the following four elements[3]:

- large set-up costs, which lead to falling unit costs as output expands and the fixed costs are spread over a larger number of units;

- learning effects, so that costs decrease or products improve as they become more prevalent;

- co-ordination effects, so that co-operation is facilitated by the adoption of common systems; and/or

- self-reinforcing expectations, where people act in certain ways in anticipation that certain systems are likely to be widely adopted.

Examples of these self-reinforcing mechanisms have been provided by technical standards such as the QWERTY keyboard[4] and Sony's Betamax video technology in the 1980s.[5] In these cases, alternative systems were available; but one system became dominant; and its dominance (and expectations that it would continue to be dominant) created incentives for people to continue investing in the system in such a way as to perpetuate the dominance of the already-dominant system. Such systems are sometimes said to exhibit 'tipping'.

Arthur[6] observes that such systems exhibit four properties:

---

[2] Prominent among them were Joan Robinson, Nicolas Kaldor and PWS Andrews.

[3] The list is a version of that provided by W Brian Arthur, "Self-Reinforcing Mechanisms in Economics", in *The Economy as an Evolving, Complex System* (ed by Philip Anderson, Kenneth Arrow and David Pines) 1988; reprinted as chapter 8 in W Brian Arthur, *Increasing Returns and Path Dependence in the Economy,* University of Michigan Press, 1994 at p 112.

[4] See Paul David, "Clio and the Economics of QWERTY", *American Economic Review,* vol 75, 1985, pp 332-7; and S J Liebowitz and Stephen E Margolis, "Fable of the Keys", *Journal of Law and Economics,* 1994.

[5] See Michael, Cusumano, Yiorgos Mylondadis and Richard Rosenbloom, "Strategic Maneuvering and Mass Market Dynamics: The Triumph of VHS over BETA", CCC Working Paper, No 90-5, Center for Research in Management, Haas School of Business, UC Berkely, 1990.

[6] See Arthur, *op cit.*

- *Multiple equilibria.* In these cases, two or more quite different 'solutions' could have been produced. 'The outcome of the process is indeterminate; it is not unique and predictable.'

- *Possible inefficiency.* It is possible that one of these technologies is inherently 'better' than the other; but, because it did not gain early market acceptance, it was beaten by a poorer solution. In the cases of QWERTY and video technology, the technical experts have debated as to whether the victorious system was or was not the best.[7]

- *Lock*-in. Once a solution is reached, it becomes difficult to change. Because people have invested on the assumption that the system will continue, any change of system will cause a re-incurring of these (sunk) costs.

- *Path dependence.* The early competition among systems (which may – at least in part – be determined by chance events) can determine which system ultimately prevails.

In section 2.3 of this Report, we give further examples of self-reinforcing mechanisms. In section 3, we suggest that the adoption of a particular system of interchange for debit transactions has many of the qualities of a self-reinforcing system. Many of the issues that have been raised in the debate within Australia over EFTPOS interchange are very similar to those that have been raised with respect to many other examples of self-reinforcing systems.

2.3   SOME EXAMPLES OF SELF-REINFORCING MECHANISMS IN THE ECONOMICS LITERATURE

In section 2.2 above, we gave some examples of debates in the economics literature over the adoption of technical standards such as the QWERTY keyboard and the Sony Beta video system. However, in the last few decades, economists have identified (and modelled) self-reinforcing systems in many circumstances. We list a few examples.

2.3.1 International trade theory

Many models of international trade allow for economies of scale in the production of a particular commodity. If there is a model of two countries and two goods and production of each good exhibits increasing returns to scale, the obvious prediction is that each country will specialise in the production of one of the goods. However, it is not clear which country will specialise in which good. That is, there are two possible equilibria.[8]

---

[7] See S J Liebowitz and Stephen E Margolis, "Network Externality: An Uncommon Tragedy", *Journal of Economic Perspectives,* Spring 1994, pp 133-50.

[8] See B Ohlin, *Interregional and International Trade,* Harvard University Press, 1933; and E Helpman and P Krugman, *Market Structure and Foreign Trade,* MIT Press, 1985.

2.3.2 The growth of cities

Many writers have developed Marshall's idea of agglomeration economies associated with the growth of cities. These suggest that a city may start to grow by chance; but once it has started on a path of growth, self-reinforcing mechanisms will propel it along a path of growth. For example, it has been argued that the growth of Silicon Valley in California was largely a matter of chance. A few key people happened to establish businesses near Stanford University in the 1940s and 1950s; and this created a pool of labour and inter-firm markets that led to further agglomeration economies which led (over time) to hundreds of other businesses being located in the area.[9]

2.3.3 Industrial organisation

Since the 1980s, various authors have developed models exploring the implications of network effects for patterns of competition. Katz and Shapiro used the term 'network externalities' to refer to advantages that consumers gain by using systems that are compatible with the systems of other consumers. For example, if a large number of consumers believe that Microsoft's operating system will be used by most of their work colleagues, each will have a strong incentive to purchase software that is based on that operating system; and various software producers (including Microsoft) will have incentives to produce a large number of products that are based on the Microsoft operating system. The result will be a self-reinforcing mechanism.[10] This analysis by Katz and Shapiro was the catalyst for much further work, for example that of Economides, which explored the implications for pricing of goods that were produced as part of a compatible package of products.[11]

---

[9] On Silicon Valley, see DL Cohen, "Locational Patterns in the Electronic Industry: A Survey", mimeo Stanford. On the theory of agglomeration economies, see Walter Isard, *Location and Space Economy,* Wiley, 1956; and W Brian Arthur, "'Silicon Valley' Locational Clusters: When do increasing returns imply monopoly?", *Mathematical Social Sciences,* vol 19, 1990, pp 235-51.

[10] See Michael Katz and Carl Shapiro, "Network Externalities, Competition, and Compatibility", *American Economic Review,* vol 75, June 1985, pp 424-40; "Technological Adoption in the Presence of Network Externalities", *Journal of Political Economy,* vol 94, August 1986, pp 822-41; and "Systems competition and Network Effects", *Journal of Economic Perspectives,* vol 8, Spring 1994, pp 93-115.

[11] See Nicholas Economides, "Desirability of Compatibility in the Absence of Network Externalities", *American Economic Review,* vol 79, December 1988, pp 1165-81; and Nicholas Economides and Steven Salop, "Competition and Integration among complements, and network market structure, *Journal of Industrial Economics,* vol 40, March 1992, pp 105-23.

3 The EFTPOS interchange system as a self-reinforcing mechanism

3.1   EFTPOS INTERCHANGE AS AN EXAMPLE OF A SELF-REINFORCING SYSTEM

The present Australian system of EFTPOS interchange has a number of distinctive characteristics. Two of these have attracted some controversy during the last few years. The first of these controversial characteristics is that it involves a 'negative' flow of funds from issuers to acquirers. The second is that the rates at which these funds flow is a matter of bilateral negotiation among the various pairs of financial intermediaries.

In the matter *Re EFTPOS Interchange Fees Agreement* [2004] A CompT 7, the Australian Competition Tribunal found that EFTPOS was introduced into Australia in 1984 by the major banks.[12] EFTPOS was not widely accepted until substantial investment had been undertaken by acquirers and retailers. Negative interchange fees were negotiated to finance the investments that were undertaken by acquirers.[13] There still remained the problem of the customisation of the card payment systems to meet the needs of merchants. This was resolved by merchants taking the initiative to introduce their own equipment in the early 1990s.

> The commercial consequence of this was that once major retailers had committed to install their own networks they were in a position to negotiate what for ease of reference may be called a fee (sometimes it took the form of rebates or credits in relation to other transactions) from acquirers to compensate them for the investment in EFTPOS infrastructure and processing costs and the corresponding reduction in the acquirers' costs. Acquirers also benefited from lower transaction costs for the existing transactions, due to the high volume of transactions being delivered to them.[14]

The Tribunal adopted an explanation of the current system of negative interchange (in which funds provided by the issuer are shared, to some extent, between the acquirers and the large retailers) in terms of path dependence. Certain historical imperatives launched Australian banking along the path of negative interchange.

The Decision of the Tribunal also suggests why, once Australia headed along the path of negative interchange, we became locked into the system. Once particular institutional arrangements for flows of funds were established (and were expected to continue) incentives were established for certain patterns of investment. And the system becomes self-reinforcing.

---

[12] EFTPOS Tribunal Decision, para 30.

[13] EFTPOS Decision, para 38, quoting the evidence of Mr Gove.

[14] EFTPOS Decision, para 41.

The flow of funds from issuers to acquirers and to merchants created strong incentives for acquirers and merchants to invest in equipment and in promotional activities to attract further EFTPOS transactions. The number of EFTPOS terminals increased dramatically over time and, today, more than half of all EFTPOS transactions are processed through terminals and PIN pads owned by merchants.[15] The flow of funds from issuers also created an incentive for banks to reduce their investment in branches. The net result of these (and other) pressures was that in 1990 there were slightly over two EFTPOS terminals for every bank branch; and today there are 89 terminals for every bank branch.[16]

Australia's system of negative interchange for EFTPOS is not the only controversial system for the flow of funds that seems to be self-reinforcing. Another example seems to be the system in Australia (and Europe) for calling-party-pays (CPP) for mobile telephone calls. Under CPP the person who initiates a call to a mobile handset will pay for both:

- the cost to its own telecommunications operator for taking the call to the point of interconnection with the mobile network of the receiving party; and

- the charge that the mobile operator of the receiving party charges for completing the call.

In the early days of mobile calls, the United States settled on a different system, whereby the calling party pays only for the cost of the call up to the point at which it leaves the network of the caller's network service provider; and the receiving party pays for the cost of the call from the point of interconnection up to the receiver. This is called receiving party pays (RPP).

It has been argued that the CPP system in Australia and Europe leads to high charges for the service of terminating mobile calls and this causes a flow of funds from providers of fixed networks to mobile network operators. The argument has been accepted by the Competition Commission of the United Kingdom and, more recently, by the Australian Competition and Consumer Commission.

In a recent paper, Stephen Littlechild analyses the reasoning of the UK Competition Commission. The Commission acknowledged that there was vigorous competition among the mobile network operators (MNOs), so the flow of funds caused by the CPP system did not lead to enhanced profits to the MNOs. Rather, the flow of funds created investment incentives that dissipated the funds that flowed from fixed to mobile networks as a result of the CPP system:

> But are the mobile operators able to keep these increased revenues from the higher termination charges? Not if the mobile sector is competitive. Each additional subscriber is likely to be called by subscribers on other mobile and fixed networks, so will bring a flow of termination charge revenues to the mobile operator. Each mobile operator will therefore try by a variety of means to attract

---

[15] EFTPOS Decision para 45.

[16] EFTPOS Decision, para 44.

subscribers to its network. This may involve incurring marketing costs, cutting other prices, and offering various subsidies. As the Commission [UK Competition Commission] put it, "There is vigorous competition among the MNOs to attract and sign up subscribers to their networks, for example through the payment of incentives and discounts to retailers, and handset subsidies to customers, but this is funded by excess returns from termination charges." The commission calculated that "the average net cost of acquiring a new customer was around 100 pounds."

In the limit, mobile operators would find it worthwhile to spend up to the expected value of the net revenue from termination charges in order to attract new subscribers. The more competitive the market, the more likely this is to happen, because the mobile operators have no alternative but to do so. Conversely, the less competitive the retail market, the more that the surpluses would remain with the operators instead of being passed through to the subscribers.[17]

Once a country has adopted a CPP system or an RPP system for the flow of funds within telecommunications networks, incentives are created to invest in certain activities. These incentives will affect bargaining positions into the future, so a country will be 'locked in' to a system that was adopted many years in the past. So the adoption of a particular system is said to be 'path dependent'.

The Tribunal decision in *Re EFTPOS Interchange Fees Agreement* is consistent with a similar explanation of Australia's system of EFTPOS interchange. In the late 1980s the issuing banks had to create incentives for merchants to place machines in their stores. Merchants had to ensure that the new EFTPOS systems were consistent with other elements in their payment systems. The result was a system of bilateral contracts with negative interchange, some of which flowed on to those merchants who had invested in the new machinery. The expectation that this flow of funds would continue has continued to create incentives for patterns of investment by issuers, acquirers and merchants. That is, the system has become self-reinforcing. The initial investments created a flow of funds which created incentives for further investments and so on …

3.2 CAN VARIOUS SYSTEMS OF INTERCHANGE BE RANKED ACCORDING TO THEIR EFFICIENCY?

As we noted in Section 2.2 above, a common theme in the literature of multiple equilibria and path dependence is that an economy may settle on an inefficient system and the inefficient system will be locked-in by self-reinforcement. That is, because of the problem of lock-in, the least fit may survive.

In some cases of multiple equilibria, there is no obvious efficiency ranking of the equilibrium paths. A well-known example is driving on the left or the right side of the road. There are two equilibria: everyone drives on the left or everyone drives on the right. Any other arrangement is not an equilibrium. But the two equilibria are equally good.

---

[17] Stephen C Littlechild, "Mobile Termination Charges: Calling Party Pays vs Receiving Party Pays", April 2004, CWPE 0426.

In the cases of mobile telecommunications, there is a large literature on problems associated with CPP and RPP systems; but there has been relatively little analysis of the respective merits of the two systems. Littlechild points to two lessons that might also be applied to a comparison of the relative merits of negative and positive interchange in EFTPOS.[18]

The first is that any change in the direction of funds will reduce some prices and increase others. To the extent that there is competition in acquiring among merchants, the funds that flow in the form of the interchange will be dissipated in investments in assets of various kinds. If Australia were to move from a system of negative interchange to a system of zero or positive interchange, the prices of the services yielded by these assets would increase.

The second lesson follows from the first. If a change of systems increases some prices and decreases others, then any comparison of the relative efficiency of the two systems must involve detailed modelling of welfare gains and losses. The relative efficiency of the two systems is essentially an argument about allocative efficiency. The accepted way to resolve such debates about relative orders of magnitudes is to undertake empirical estimation of these magnitudes. To quote Littlechild with respect to mobile termination charges:

> Oftel's approach, which the [UK] Commission adopted, thus focuses on what economists call 'allocative efficiency': an improved allocation of resources. It is worth emphasising that the logic of this approach demands that detailed welfare calculations be done, no matter how reliable or otherwise they can ever hope to be. Moreover, the use of allocative efficiency as the benchmark means that the net benefits are always likely to be small. In this respect the price controls on mobile termination charges stand in contrast to price controls on other utility networks.[19]

There is a third lesson that must be added to the two suggested by Littlechild. Lock-in is caused largely by sunk costs of one kind or another that various people have incurred. If a system jumps from one equilibrium path to another, incentives to invest will change and new sets of assets will need to be acquired by the parties involved. These costs of moving from one equilibrium path may be very substantial. In particular, they would need to be included in any weighing up of the relative merits of moving from one system to another.

---

[18] See Littlechild, *op cit,* pp 14-15.

[19] Littlechild, *op cit,* p 14.

4 Why do other countries have different systems?

In *Re EFTPOS Interchange Fees Agreement,* the Tribunal was confronted with confusing evidence as to the nature of EFTPOS interchange arrangements in countries other than Australia. After reviewing this evidence, the Tribunal stated that they were not convinced of its relevance:

> Frankly, we do not see a great deal of value in overseas comparisons. The way a banking system operates in a given country is a result of a complex mix of historical, geographical, political, cultural and socio-economic factors. It is not likely to be a profitable exercise to engage in a detailed examination of these with a view to seeing what features are or are not replicated in the Australian experience.[20]

This paragraph goes to what is sometimes referred to in the literature of self-reinforcing systems as the 'selection' problem.[21] The problem arises in the case of a system that has multiple equilibrium paths: what determines which particular path is taken?

There are two approaches to this problem The first is to rely on probability theory and to say that the final equilibrium path that is chosen depends on a series of accidents. The second approach is to attempt to analyse the 'historical, geographical, political, cultural and socio-economic factors' that might be subsumed under the umbrella of the stochastic shocks to which the process was subject in its early years.

One factor that seems to have been important in causing Australia's system (negative interchange, bilaterally negotiated) to be different from many others is that it was introduced before many others; and, in particular, it was introduced before electronic credit transactions were at all common. This can be contrasted with other countries. EFTPOS was not introduced into Canada until the early 1990s;[22] it was not introduced to the United Kingdom until the late 1980s;[23] and it was not introduced to Germany until 1990.[24]

As the Tribunal noted, EFTPOS was introduced into Australia in 1984, when banks had issued electronic cards to use at ATMs, but before merchants had installed equipment to process electronic transactions. This meant that the issuing banks had to persuade merchants to install machines for electronic processing of EFTPOS when the merchants had no idea of their acceptance in the market.

---

[20] EFTPOS Decision, para 81.

[21] W Brian Arthur, *Increasing Returns and Path Dependence in the Economy,* University of Michigan Press, 1994 p xv.

[22] Interac Association, *Interac Association: A Backgrounder,* January 2003, p 3.

[23] Bank for International Settlements (Committee on Payment and Settlement Systems), *Payment and Settlement Systems in Selected Countries (Red Book),* April 2003, p 405.

[24] European Card Review, *European Payment Cards Yearbook, Volume 2 – Country Reports,* 2001, p GE-5.

As the Tribunal stated, there were (doubtless) many factors that led Australia to adopt the system that it did. Nevertheless, once the system was established it was hardly surprising that it became self-reinforcing. The flows of funds that were established, created incentives for investments to be funded in such a way as to lead to the perpetuation of the system.

5 Conclusion

Until the last twenty years, economists had a strong bias in favour of explanations of economic phenomena by means of models with a single equilibrium. However, economists have come to realise that many processes have multiple equilibria. Such processes have been observed in a large range of situations. The system of EFTPOS interchange seems to be yet another example of a system that can yield multiple equilibria.

Although there is unlikely to be a simple explanation of why Australia came to adopt the system it now has (negative interchange, bilaterally negotiated). Any system will lead to particular investment incentives which are likely to see the system being self-reinforcing.

It may be that one equilibrium is more efficient than another. But the ranking of systems is a difficult task, because investment incentives are dependent on the financial flows. This means that any system will result in some prices that are higher and some prices that are lower than in any other system. This means that the ranking of systems would need to be the outcome of a process of quantifying the costs and benefits of the various possible systems. If the quantification is to be the prelude for change of systems initiated by regulation, the regulator should also consider the extent to which sunk costs (that have been incurred in response to the established system) would need to be reincurred as a result of the change.