

The Operation of Hawala-style Transjurisdictional Value Transfer Systems in the Contemporary Global Order: a historical and comparative perspective

Roger Ballard

Access to an effective system of payments system has always been a necessary prerequisite for the conduct of long distance trade. The reasons for this are plain to see. If merchants were to avoid the limitations of conducting their transactions on a basis of straight forward barter – or in other words the exchange of one pile of goods for another, once both parties had agreed that in their view the two piles were of equivalent value – the use of stored units of value of which was as easily to divide as it was to consolidate, or in other words a locally agreed-upon currency, was an immensely convenient innovation. Nevertheless the invention of coinage, often based on carefully ordered units of gold or silver bullion, also had its downside, especially when merchants sought to move their capital and/or their accumulated profits from one market-place to another. Merchants were always worried about the prospects of finding themselves exposed ship-wrecks and piracy; however their alarm increased exponentially when they found themselves faced with the prospect seeking to transport large sums of concentrated, and easily portable value in the form coins or bullion in value from one market place to another, since this raised the prospect that they could lose huge sums overnight to predators of one kind or another. In these circumstances construction of a transjurisdictional value transfer system of kind that would enable them to transfer their accumulated capital assets from one market-place to another without exposing themselves to the risk of pillage proved to be just as significant an innovation as was the invention emergence of coinage in the history of the growth and facilitation of long-distance transjurisdictional trading networks. However the task of maintaining the integrity, and hence the value, of the currency used to facilitate financial activity in any given local jurisdiction is an entirely different kind of activity from that of facilitating the transfer of value *as between* such jurisdictions. The regulation of, and hence the maintenance of the integrity of the currency circulating within each such local jurisdiction has always been necessity a role fulfilled by the state, whether it takes the form of a monarchy, a confederation of locally based traders, as in Italian city-states such as Venice, or – as in contemporary contexts – by the Central Bank whose core role is to sustain the financial integrity of every sovereign state; in so doing, such institutions provide the operational bed-rock on which the superstructure on which all forms of economic activity within that jurisdiction revolves: its currency.

In so far as currency in this sense is a compact, stable and readily portable store of value, it gains its utility in the market place, above all because it eliminates the complex, unwieldy and time-consuming process of barter: in other words as soon as currency became available many centuries ago, it promptly became a no brainer. Since coinage unleashed opportunities to engage in all sorts trade with unprecedented ease – at least within the jurisdiction within which it was current – coins (and in due course printed notes) were readily accepted as standardised tokens of value which could be used to facilitate the transfer of goods and services of every conceivable kind: a role which it continues to fulfil to this day. However there was one major snag in the financial systems to which the use of such tokens of value gave rise, which not only remains with us to this day, but has been hugely magnified by contemporary processes of globalisation: the inherently parochial foundations of such tokens, since it is only within the bounds of the jurisdiction by which they were issued, and by whose financial authorities their value is guaranteed that the units of any currency enjoys the status of being an unchallengeable form of legal tender.

This has several unavoidable consequences. Since the ultimate guarantor of any given currency is of necessity the state by which it is issued, it follows that so long as on the global socio-political order is composed of a multiplicity of sovereign states, every jurisdiction strives to protect its condition of financial (and hence its political and economic) autonomy by continuing to issue its own currency. The potential negative consequences of sacrificing that condition of autonomy by inextricably linking the value of the local medium exchange with that deployed by a larger and more prosperous neighbour is plain to see in the light of the intense difficulties which Greece, together with a number of other peripheral jurisdictions within the Eurozone, are currently facing. In each of these jurisdictions the expected boost to the manufacturing capacity as a result of monetary union failed to materialise, leading to ever higher levels of both personal and fiscal indebtedness. But having locked themselves into the Euro, none of these jurisdictions are now in a position to devalue their local currency as a ready means of extricating themselves from their position indebtedness.

The same argument applies in spades with respect to the prospective emergence of single global transferable unit of exchange. Whilst a local currency is a convenient, and in the contemporary world a necessary, means of reifying value, the better to facilitate economic activity within the parochial limits of the jurisdiction within which it circulates, from a more global perspective relative value of those units of value is far from being absolute: rather its relative worth – or more specifically its relative buying-power constantly goes up and down,

depending on the economic and commercial success (or otherwise) of its local economy viz-a-viz those of its neighbours. As the current phase of globalisation has ‘shrunk’ the world order to a wholly unprecedented degree, the consequences of this state of affairs have become steadily more pressing.

Ever since Adam Smith wrote the *Wealth of Nations*, it has been well understood that the relative value of the currencies circulating in any two jurisdictions is a function of the relative success of their respective economies, and most especially of the pattern of trade between them: if any given jurisdiction imports more goods from a counterparty (or more realistically a range of counterparties) than it exports, a balance of payments deficit will of necessity occur. There are only two ways in which such a jurisdiction can escape from the resultant deficit: either by allowing its own currency to decline in relative value, such that the imported goods become more expensive, thereby reducing the demand for them, or by depleting its own capital reserves by exporting bullion to its overseas debtors. In historical terms this has been a commonplace phenomenon. Two millennia ago, the Roman Empire exported bullion to finance its trade deficit with India; the Spanish Empire deployed the same tactic half a millennium ago, when their access to the products of silver mines in Peru and Mexico of which they had recently taken possession gave them (and their western European trading partners) access to enormous volume of ‘free money’, such that they were able to settle their trading debts with Indian and Chinese merchants (Chaudhury 1985, Frank 1998).

However it is only in exceptional circumstances (such as the discovery of giant oilfields beneath their deserts, for example) that contemporary jurisdictions find themselves in such a happy position – but only for so long as the ‘free money’ does not run out. Hence if the jurisdiction takes advantage of its position by engaging in conspicuous consumption, rather than investing in a manufacturing base in order to provide an alternative stream of income, it will in due course run into a balance-of-payments crisis – an event which has become steadily more commonplace in our current phase of ever more in globalisation. In these circumstances only two options are available: either to devalue, thereby significantly reducing relative value of their local currency, or to borrow value from elsewhere. But in the absence of an adequate value-producing base, there will be little prospect of them meeting the interest payments, let alone the capital repayments which their creditors will inevitably demand. In such circumstances the ‘beneficiaries’ of debt will simply find themselves sinking ever more deeply into the mire – leading in the fullness of time to the an inevitable outcome: a reduction in the relative value of the local currency.

It follows that despite the extent to which contemporary processes of globalisation may be shrinking our socio-economic universe, such that everywhere is becoming ever more intimately connected with everywhere else, any suggestion that this precipitating a steady shift towards global homogeneity – whether in political, social, financial, and conceptual terms – is seriously mistaken. Diversity remains just as salient a feature of human affairs as it has ever been. Moreover whilst the current phase of globalisation has radically expanded the scale of the global economy, it has also had a radical impact on its structure: one of the significant consequences of the resurgence of manufacturing activity in the Asia-Pacific region has been that the smokestack industries on which Euro-American jurisdictions have relied as their prosperity for the past two centuries have collapsed, precipitating an unprecedented degree of turbulence in the global economic order.

But whilst established patterns of relativity have consequently been swept away, inter-jurisdictional competition remains the order of the day, if only because the determination of all the contemporary world's multiplicity of jurisdictions (for splits and secessions still remain the order of the day) to take advantage of whatever assets they can lay their hands on remains as vigorous as ever. Hence globalisation has increased, rather than decreased, competition between local jurisdictions. In these circumstances it follows that the current phase of globalization has no way eliminated the significance of jurisdictional boundaries, despite exponential increases in the passage of people, of goods, of services, and hence of *value* across them, as each jurisdiction struggles to advance its own interests, and above all to retain its sovereign freedom of action in the face of the boundary-shredding impact of globalisation.

A journal article is no place in which to address all the many consequences of the current phase of globalisation. Nevertheless, these developments are irreducible aspect of the broader context within which it essential to set the much more specific set of issues on which this coalitions of inter-personal reciprocity as vehicle by means of which to facilitate the transfer of value across jurisdictional borders on a cheap, speedy and reliable basis remains just as viable, and indeed just as a complete as it was ten centuries ago, despite – and indeed in many senses because of – the challenges and opportunities which have been thrown up in the contemporary phase of globalisation.

The emergence of transjurisdictional value transfers: the logistics of Hawala

Although the logistics which underpin the operation of contemporary value transfer system are as complex as they are varied, at root all such systems face similar challenges, and have devised similarly structured solutions by means of which to address them, as those which were devised to meet the needs of traders in the ancient world. Hence ever since long-distance trade began to emerge regular feature of the economic order, value-transfer networks began to emerge to meet merchants' demands for what would be described in contemporary parlance as a reliable transjurisdictional payments system. In ancient times the central concern of long-distance traders with a means of avoiding the danger associated with constantly transporting highly concentrated caches of value in the form of coinage or bullion as they moved from one market place to another, thereby constantly exposing their working capital to theft as much as to disaster. Hence when specialist value transfer agents offering what must initially have seemed to be a miraculous service – namely the capacity to engineer the delivery of substantial packets of value in distant market places in return for a small (and often invisible) consideration – long-distance traders, Imperial tax collectors and many others were only too pleased to outsource this otherwise risky aspect of their activities to specialist financial services providers.

Since then the value transfer business has never looked back: over many centuries the scale of long-distance trading has grown exponentially, and with it the demand for long-distance payments services which were simultaneously speedy, inexpensive and above all thoroughly reliable. However in no way did those providing this increasingly essential service start shipping vast quantities of cash or bullion as between the many market places, and hence the many different financial jurisdictions into which their payments systems reached: on the contrary this was a practice onto which they only fell back as a last resort. The same point holds in spades in the contemporary world. No less at present than in the ancient past, of long-distance payments systems – from Western Union onwards – are not significantly engaged in the transfer of cash, let alone of bullion. To be sure the take in their customers' cash at one end, and pay it out again at specified destination elsewhere. But what such agents actually transfer within their payments networks is something far less substantial in character: caches of value, even these are due course delivered as cash, often in an entirely different currency from that in which the sender initially dispatched it.

How, then is this outcome actually achieved in logistical terms? And just what does the intrinsically insubstantial phenomenon of 'value' amount to in contexts of this kind? And just

how does it get to move from A to B? A little straightforward algebra provides a convenient way of illuminating just what occurs.

- i. Let us assume that a Trader in jurisdiction A, in which the local currency is denominated in units of D^a , wishes to settle a debt with a fellow trader resident in jurisdiction B, where the local currency is denominated in units of D^b .
- ii. Let us also assume that the debt accrued in return for the despatch of the consignment of goods which the recipient P has just received from a distant supplier Q amounts to $x D^b$
- iii. Let us also assume that the exchange rate between the currencies circulating in each of these jurisdiction is currently such that $n D^b = 1 D^a$

If so it follows that to settle in order to settle his debt to merchant Q in B in terms of the local currency D^a , he will need to approach his local transfer agent, whom we can identify as H^a , for a quotation for the settlement of his debt to merchant Q in B amounting to $x D^b$; H^a will then make a quick calculation, and quote $(x/n) D^a$ as his fee for the settlement. If the quotation is accepted, and the requisite sum in currency D^a is handed over, H^a will send a message to his partner H^b in B, instructing him to hand over the sum $x D^b$ to merchant Q, thereby settling the bill for the consignment of goods which he had recently despatched to P.

But whilst H^a and H^b thereby cooperated to fulfil the service required, namely the settlement of a trading debt to the two merchants' satisfaction, at another level the debt itself was manifestly still in existence, from the perspectives of the two transfer agents had simply placed themselves in the same positions which A and B had formerly occupied, since H^a had become better off to the tune of $(x/n) D^a$, whilst H^b was short of the sum of $x D^b$ which he had paid out to Q on H^a 's behalf. Not that that was any concern to either A or B; as far as they were concerned the matter was settled. In other words all that has happened is that the debt which the two traders had originally owed to one another had been lifted off their respective backs, and assumed instead by the two value transfer service providers – whom we can now conveniently identify as Hawaladars – who had stepped in to take over responsibility for settling the matter as between themselves. How H^a eventually settled the debt which he now owed to H^b was of no concern to Q, or indeed to P. All that mattered to them was that the agreed upon value transfer had taken place. From a financial perspective there is clearly nothing odd about that: in the contemporary world many millions of customers regularly insert their plastic cards into the nearest convenient ATM, no matter how far distant they may be from their home base. It is no concern of their as to how the transfer of value from their account in some distant jurisdiction, and its prompt emergence in some obscure local

currency is engineered. All anyone cares about is that the system works reliably, and does so at minimum a cost.

Back-office procedures: the conceptual roots of Hawala in Islamic Law

Most ancient civilizations developed sophisticated value transfer systems, all of which operated on broadly similar underlying premises. Most remain in operation to this day. But whilst the Chinese typically rely on *Feiqian* (otherwise known as flying-money') and Hindu *Shroffs* rely on *Hundis*, this article focuses on the operation of the contemporary Hawala-style 'informal' Value Transfer Systems whose global hub is currently located in Dubai, whose Exchange Houses are currently in a position to facilitate value transfers to almost every jurisdiction around the globe (Ballard 2005, 2010).

Such Hawala networks are of extremely ancient origin, and serve much the same purpose as they did when they came into existence before the birth of Christ, let alone of the Prophet Mohammed: namely to facilitate value transfers generated by the seaborne traders circulating through the vast arena ranging from East Africa, the Red Sea and the Persian Gulf, up and down the western and eastern coasts of the Indian sub-continent, through straits of Malacca and onwards to Indonesia, China and Japan (Chaudhury 1985, Ballard 2009). Whilst records of financial transactions are inevitably fragile, documents which came into being as a result of this trading network – of whose existence there is plentiful archaeological evidence – is to be found in Arabic sources generated from the ninth century CE onwards, when the conceptual underpinnings of the practice Hawala began to be incorporated into system of Islamic law – in other words the Shari'ah – which Muslim scholars were in the midst of articulating (Schacht 1944, Udovich (1970).

The term Hawala is of Arabic origin, and is derived from the root *h-w-l*, which points to all conceivable forms of change, exchange and transformation, including, in this case, value transfer. As Islam spread outwards from Mecca and Medina, Muslim merchants were keen to ensure that their commercial transactions were *halal* in character, such that they were congruent with the prescriptions of both the *Qur'an* and the *hadith*. In that respect two requirements derived from the Prophet's teaching were of pressing concern: firstly that gaining unearned profits from an enterprise without having shouldered any portion the potential risk was intrinsically exploitative, and hence should be avoided at all costs; and secondly that whilst debts were 'real', in the sense of being grounded in patterns of interpersonal obligation arising from real trading transactions which could either succeed or

fail, could consequently be identified *facto* as *halal*; by contrast the loaning money on interest-bearing credit, without the lender in participating in any way in the underlying risk, was inherently exploitative, and hence *haraam*.

Whilst Hawala procedures are consequently grounded in the premises Islamic Law, close inspection of the content of those premises reveals that they are notably sensitive to the financial challenges faces by merchants conducting transactions in transjurisdictional contexts, as is immediately apparent from masterly account of the conceptual underpinnings of Hawala which Schacht (1964) set out many years ago in his much admired *Introduction to Islamic Law*.

[Hawala] is in the first place, a mandate to pay, i.e. I owe something to *A* but charge *B* to pay my debt. It can also be an assumption of my debt by *B*. The practical prerequisite in both cases is that I have a claim against *B* which is equal to or higher than the claim of *A* against me. This is not necessarily a debt. It can also be for the return of an object, e.g. a deposit or something taken by usurpation. Normally, therefore, the *hawala* amounts to an assignment: I assign to *A* a claim of mine against *B*, in order to satisfy a claim of *A* against me.

But the existence of a claim against *A* against me is not a necessary prerequisite, and the *hawala* then amounts to a mandate to collect, i.e. I charge *A* to collect my claim against *B*. The element common to all cases is merely that an obligation of *B* towards *A* is created. The acceptance of the *hawala* by *A* extinguishes my obligation; it revives only if *B* dies bankrupt or denies the existence of the *hawala*. Performance of *B* towards *A* extinguishes my claim against him only if the *hawala* was conducted with specific reference to this obligation, not if it was unconditional. (Schacht 1964:148-149)

Moreover as Schacht goes on to show, once the legitimacy of switching and/or of swapping personal obligations of debt on this basis had been established on this basis, the way was open to offer a formal foundation in Islamic law to the arrangements which merchants who had made entered into a multiplicity of transactions back and forth in any given local marketplace to facilitate a comprehensive settlement of all the debts which they had accrued as between themselves, as well as between themselves and their customers, during a given period in one fell swoop. In other words it both facilitated and legitimated the arrangement of a series of reciprocal swaps between participants in a collectivity, such that they could settle their mutual debts with resort to the services of a clearing bank – or in other words without recourse to the services of the of a current account in a clearing bank, the resource on which virtually all contemporary Euro-American rely to collect our salaries and to pay our debts. It follows that Hawala can be accurately, although anachronistically, described as mode of facilitating ‘banking without banks’.

However the strategic potential unleashed by the physical and legal underpinnings around which the *modus operandi* of Hawala was constructed amounted to much more than a mere ‘fill-in’ for the absence of contemporary banking services. Where Hawala systems really came into their own was with respect to their capacity to provide logistically reliable and legally grounded foundations to underpin the transjurisdictional value transfer system which by then spread out across right out across the Indian Ocean and into the South China Sea.

The essence of Hawala (at least in the contexts which have so far considered) was grounded in the arrangement of mutual agreements between merchants exchanging goods and services which enable them to settle their consequential debts in due course either by exchanging them directly with one another, or by nominating some other person to whom the debt might be discharged on the creditor’s behalf. However it also followed that once such a coalition of reciprocity (a concept which I will go on to examine in much closer detail in due course) had been constructed between the participants in any given Hawala network, there was no need for those conducting business on this basis to settle their debts on a daily basis, or even to have daily face-to-face contact with their trading counterparties. Provided that they had confidence that all such debts would in due course be discharged, they could confidently do business with counterparties sitting in market places several thousand miles away across the ocean. As Schacht goes on to explain:

One of the practical advantages of this institution is that it enables me to make payments in another place through *B*. Its effect is the same as that of the *suftaja* or bill of exchange. This defined as ‘a loan of money in order to avoid the risk of transport’; I lend an amount to *B*, in order that he may pay it to *A* in another place.

In other words *hawala* procedures with respect to the transfer of debt provided ideal means by which Hawaladars could to facilitate long-distance value transfers on a risk-avoidant basis on behalf of their customers. Yet just how were the logistics of this outcome actually achieved? When I left off the argument, all that had been achieved was that a pair of cooperating Hawaladars had taken the debt which their customers had initially owed to one another onto their own shoulders, such that H^a was better off to the tune of $(x/n)D^a$, whilst H^b was short of the sum of xD^b which he had paid out at some distant destination to settle the merchants’ commercial transaction on behalf of H^a . To make sense of the operation as a whole we need to take a wider view of its conceptual and logistical foundations.

Hawala and the settlement of debts

As Adam Smith long ago insisted, transjurisdictional trade is of necessity a two way process: if a merchant wishes to obtain goods which are scarce, sought after and expensive in his own jurisdiction, but cheap and plentiful in a distant jurisdiction, the best way to get to implement a trade is to export goods which are readily available at his home base, but much sought after, and hence more valuable elsewhere. In circumstances where opportunities for were unavailable – as for example when western European traders gained direct access to the fabled wealth of the East when they found their way round the Cape of Good Hope half a millennium ago – they were disturbed to find that there was no significant demand for the trade-goods which they brought with them in that then immensely prosperous region in the globe. Hence they found themselves in the same position of the ancient Romans: the only way in which they could buy their way into Indian markets was by paying up in bullion. But whilst unbalanced trade of that kind rapidly precipitates severe economic grief, the trade routes around the Indian Ocean which had emerged around the Indian Ocean of the course of centuries, which gave rise to complex exchanges as between traders in Chinese ceramics, spices from Indonesia and Kerala, Bengali rice, cotton textiles from Western India, Pearls from the Gulf, Carpets from Iran, and Ivory from Africa (to name but a few) circulated these goods around the entire region on an ever expanding scale; and since most of the trade was despatched by sea, and most of the seamen, as well many of the merchants were by now Muslim, a Hawala style payments system emerged as the principal means of settling trading debts all the way from Zanzibar to Shanghai (Ballard 2010). To be sure much has changed since Europeans broke into this Asiatic lake having found their way round the Cape of Good Hope and Cape Horn, Hawala networks, no less than those of *Feiqian*, continue to thrive to this day. How, then, does it operate?

Hawala does not operate on a one-off basis. Rather it thrives in circumstances where there a multiplicity of traders (and a rather smaller number of Hawaladars) in business in multiplicity of jurisdictions, all connected by a complex network of trade routes. So long as business is booming, there will be a multiplicity of traders at each node with invoices from counterparties in need of settlement as and when their goods arrive. In these circumstances everyone involved in making such deals will have a strong interest in seeing their debts settled on a reliable basis, as swiftly and above all as cheaply as possible. It was this logistical challenge which Hawala networks were developed to meet, long before the arrival of Euro-

American traders, or Euro-American banking systems which the imperial powers brought with them in the latter part of the nineteenth century.

With this background in mind we can begin to unpick the essence of contemporary Hawala procedures, the back to back swap, by adding in the parallel transactions which accompany unidirectional algebraic formula which I set out earlier. Whilst that formula focussed on the logistical and conceptual of a single value transfer implemented by single pair of pair of cooperating Hawaladar, in the real world any given Hawaladar can expect to execute similar transfer on behalf of a multiplicity of customers on a daily basis. More is that all: many of those customers will have a multiplicity of debts to settle, as like as not requiring transfers to a variety of destinations; moreover since trade is a two-way business, he can also hope to receive a similar volume of disbursal instructions from his overseas partners, such that the amount cash deposited with him by his own customers is broadly congruent with that which he disburses to his distant partner's behalf to his partners local customers. Hence in the ideal world which the logic of system gives the Hawaladars in any given network a strong incentive to achieve, the sum total payments in and payments out at every node in the system should be as close to zero as possible.

So it is that all value transfer institutions, including banks, give the appearance – at least from the customer's perspective in the front office – of transferring cash on their behalf to a specified recipient elsewhere, since cash is what they pay in or take out as the case may be, all that happens in such institutions' back-offices is a transfer of an abstraction: value. Meanwhile currency notes stay put, even if they make regular excursions through a revolving door – as every banker is well aware. It follows that the key determinant of profitability for those engaged in value transfer is the efficiency with which they can implement the complex, and potentially exceedingly labour-intensive, logistics of their back office settlement processes. This exercise becomes steadily more challenging when settlements involve deal with other branches of the same bank, with other banks in the same jurisdiction, and worst of all with other banks in other jurisdictions. Hawala networks have always operated at the toughest end of this scale.

Necessary prerequisites for the implementation of value transfer

It follows that value transfer systems, no matter what form their organisational characteristics may take, face a similar set of logistical challenges. In meeting these challenges, consolidation is invariably the name of the game: in few other arenas do economies of scale

generate such a dramatic diminution of overhead costs. Hence the key to success in the value transfer game is to avoid the time consuming process of settling each transaction one by one, and the even more expensive process of transporting bullion, or even packages of currency notes from one place to another. But whilst the use of the insubstantial construct of value as a substitute for its physical and jurisdictionally specific counterpart, money, undoubtedly eases the implementation of that task, in no way does it eliminate all the underlying of logistical problems. Those which still need to be resolved can conveniently be considered under three main heads:

- *Communication*: how to move *information* on a secure basis from $a \rightarrow b \rightarrow c$ and back again
- *Logistics*: how to move *value* from $a \rightarrow b \rightarrow c$ and vice-versa
- *Security* : how to contain the *risk of malfeasance*, and above all *systemic integrity* when a, b and c are widely separated in both physical and jurisdictional terms

Communication was once a much more challenging task than it is today. In the past messages could easily take months to reach distant destinations. However, ever more dramatic technological developments during the course of the past century and a half have changed all that. Given broadband cables and the internet, further supplemented by mobile and satellite phones where fibre connections are unavailable, instant and secure data transfer facilities are now available 24/7 from anywhere to everywhere else. But if issues of communication ceased to be problematic in a wired world, the processes of globalisation of which these developments are an integral part have also thrown up new challenges. Whilst global trading activity, and the consequent need for invoice-settlement services, expanded exponentially during this period, the parallel growth of long-distance migration led to a hitherto largely unprecedented source of customers for value transfer services: migrant workers looking for a cheap and efficient way of remitting a significant portion of their wages to their families back home. The scale of long distance migration, and the scale of the resultant scale of remittance transfer has grown exponentially during the past two decades; it currently amounts to at least \$500 billion per annum on a global scale. However meeting the demands of this novel section of the value transfer market has proved to be exceedingly challenging in logistical terms: not only are the packets of value which migrants seek to remit little more than ‘penny packets’ from a banking perspective (often as little as \$100, and with a mean somewhere around \$500), but the destinations to which delivery is sought are frequently in remote rural areas in the global South, and as such well beyond the banking frontier. It follows that whilst the

current overall volume of migrant remittances is of immense financial significance, most especially for jurisdictions in the global South (Maimbo and Ratha 2005), the task of implementing the delivery of millions of ‘penny packets’ of value to remote rural destinations has presented value transfer agents with logistical challenges of unprecedented severity; and if Hawala-style networks have experienced a dramatic revival during the course of the past two decades, it is largely because their underlying logistical model, and most especially the communications network on which their operations relied, proved to be of a kind which could readily be adapted to cope with this novel challenge. As a result they were able rapidly to develop the kind of speedy, reliable and low-cost delivery service that their customers required.

Yet despite the real-world efficiency of the remittance-based value transfer service which Hawaladars have been able to deliver to millions of customers in recent decades, neither its expansion has not been plain sailing: indeed if one types Hawala into Google, it comes up with over half a million results, the vast majority of which suggest that Hawala networks are no more than a vehicle for criminal conspiracy, on the grounds that their principal purpose is to facilitate money laundering (Ballard 2006). The reasons why Hawala has acquired this reputation are diverse, but two are of particular significance: firstly the tendency of virtually all economists and bankers whose training and practice has been wholly restricted to Euro-American conceptual universe to find themselves totally bewildered when faced with a financial system which is grounded in relationships of trust rather than those of contract, with the result, amongst other things, that Hawaladar’s back offices, to the extent to which they exist, lack most of the documents on which auditors would expect to rely in the course of figuring out what was really going on. Secondly and just as significantly, in the immediate aftermath of 9/11, the USA took the opportunity to pass the Patriot Act, in which it took advantage of the US Dollar’s position as the *de facto* global unit of account to demand that all foreign financial institutions wishing to maintain trading relationships with the United States should take steps to ensure that all aspects their operations complied with AML/CFT (Anti-Money-Laundering and Countering the Financing of Terrorism) regulations.

As everyone concerned with the contemporary financial services sector is well aware, compliance with AML/CFT, and especially with its Know Your Customer requirements, adds considerably to their existing overheads. They are also deeply sceptical as to whether they are actually fit for purpose, given the objectives they are ostensibly designed to serve (Ballard 2005). Nevertheless to the extent that the new global regulations assume that all

reputable institutions will order their financial operations in terms of contemporary Euro-American methodologies – namely that all transactions are contractually grounded and formally documented, not just as between the service provider and the customer, but at every stage of the transfer process, such that auditors can readily check the whole process for accuracy, and in doing so constantly keep an eye out for malfeasance – the requirement to implement an additional dimension of transparency and traceability required to maintain compliance with AML/CFT may not have been particularly welcome, but was nevertheless no more than an extension of their existing practices. Moreover however costly compliance with the new regime might be, and it often turned out to be substantial, every institution could at least take comfort in the fact that that all their competitors have to bear the same burden. It also followed that these additional costs could readily be accounted for under the heading of ‘security’, and in any event would undercut profitability, since the burden could readily be transferred to customers as service charges – always provided that everyone obeys the same rules.

However the additional costs precipitated by the new compliance regime did not fall evenly across the whole spectrum financial activities, largely as a result of the impact of economies of scale. Hence the cost of compliance with requirements of AML/CFT, transparency and so forth with respect to the transfers of packets of value \$100, of \$1,000, of \$1 million or of \$100 million are in no way proportional to the quantum value being transferred: on the contrary they are much the same in each case. Indeed fiddling around with small packets of value on behalf of customers with potentially dodgy IDs and no credit record will like as not require a significantly greater degree administrative attention than the transfer of \$10 million to a similar destination on behalf of an established corporate customer. The same considerations also apply when currency conversion is an aspect of the transfer. When UK Bank is contacted by a local corporate customer to settle an invoice for Rs.10 million from its Indian service provider, there is no need to mess about: the bank can readily purchase Rs. 10 million on the spot market (most probably in New York), send a messages over SWIFT to implement the delivery of the tranche Rupees purchased in New York to the Indian service provider’s bank account in Bangalore (or more likely the Banks correspondent account in another Wall Street bank), and having recalculated the USD cost of purchasing the Rupees in Sterling as well as adding a small commission for its troubles, debit its customer’s account accordingly.

If, however, we consider the logistics of another equally commonplace transfer, but one which lies at the other end of the scale – as for example when a recently arrived Pakistani migrant worker wants to send £200 to his wife and children in Azad Kashmir – the whole business becomes a great deal more complicated. In the first place the customer may run into difficulties with respect AML/CFT compliance, especially if the bank is uncertain about the legitimacy of his presence in the UK. Secondly the packet of value which he wants to send to Pakistan is minute in banking terms: no-one goes to the spot market to buy £200 worth of Pakistani Rupees; and last but not least even if his wife had an account in a bank located in the remote part of Azad Kashmir in which the family resides (an unlikely prospect), the prospect that any such bank would have a direct connection to SWIFT is even more remote.

But even the family was based in Islamabad or Karachi, where such facilities are indeed available, a local branch of a UK bank would not dream of despatching such a small packet of value on a transjurisdictional journey on a freestanding basis. Rather the details would be sent to a central international office in which this transaction could be bundled up with several thousand, and perhaps tens of thousands, of similar value transfer instructions, which could in due course readily be processed en bloc on wholesale basis through the money markets, before being disaggregated in Pakistan, and eventually finished off with the delivery of a packet of rupees to his wife. But in the aftermath of all this time consuming rigmarole, and of the imposition of all sorts of handling charges accumulated along the way, his wife would be lucky to see £100 in value at the other end several months later – and even then the bank would probably have made a loss on the whole transaction.

It follows however good a transjurisdictional value transfer service Euro-America's mainline banks may be able to provide to their substantial commercial customers, and of course to high net wealth individuals, the service they are able to provide for less substantial customers standing as the other end of the spectrum are much less attractive, no less in terms of speed of delivery than in terms of their services charges. Whilst migrant workers by no means the only section of the population who suffer as a result of this deficiency, since travellers of all kinds regularly run into this problem, they are nevertheless deeply disturbed by the prospect of large chunks of their hard-earned savings disappearing in service charges in the course of despatching them back home. In these circumstances it should come as no surprise that they have developed their own alternative structures by means of which to facilitate remittance transfers than the expensive and sclerotic services currently on offer from institutions grounded in the mainline financial services sector.

But this also raises a further set of questions. Whilst the communications systems to which the mainline financial services sector has access as a means of facilitating value transfers, running from access to SWIFT onwards, enables messages to be despatched to almost any destination around the globe virtually instantaneously, and whilst globalisation has precipitated a dramatic speed-up in almost every other sphere of manufacturing and commercial activity, the process of transferring value transjurisdictionally on a retail basis has been largely untouched by these developments – except in the case of Hawala-style networks.

Why should this be so? Is it because the networks are inherently criminally in character? Or is it, to the contrary, that the Hawaladar's preferred logistical model precipitates outcomes which are substantially superior from their customers' perspective than those which institutions operating the mainline financial services sector routinely deploy – and indeed with which they are *required* to comply, thanks to the global impact of the Patriot Act? However if we cast the paranoid concern precipitated in the aftermath of 9/11 to one side for a moment, a yet deeper issue also emerges. Just how to Hawaladars manage to run such an logistically efficient and systemically stable transjurisdictional value transfers without deploying the multiplicity of administrative checks and balances which the Euro-American financial system deems to be a necessary prerequisite for the maintenance of systemic integrity? Why, to put the matter yet more pressingly still, did the Euro-American banking system collapse into a condition of systemic failure in the autumn of 2008, whilst the Hawala system has not required – or sought – any kind of state support in order to maintain its integrity? To answer that one we will have to go right back to first principles.

Value and the exchange of debt

Whilst there is clearly nothing more valuable than value itself, it is at the same time a wholly evanescent phenomenon, and all the more so when it is transferred from one spatial location to another. It is easy to see why: transferred value, like debt, has no physical substance. Rather it is a measure of the obligation which, as a result of a transaction of one sort or another between two persons, one party owes to another. In face to face contexts, the obligations arising from such transactions can readily be extinguished, as for example when a farmer barter a sack of potatoes for a newly fired pot, or in a monetised context, by anyone with sufficient cash in his pocket purchasing the pot from the Potter for an agreed sum of money. But once the parties are spatially separated, everything becomes a great deal more complex: even if I had the means to send the cash with which to buy a pot fired in Bombay,

my sterling transfer would have to be exchanged for rupees before it was of any value to Potter. Moreover the value of the cache of rupees which the Potter received would only come into its own if he was able to exchange the notes for something more useful: a sack of rice, perhaps. It follows that transjurisdictional value transfers invariably give rise to a double exchange. In this hypothetical example I would have deployed the (sterling) value generated by cutting several people's hair to swap for a similarly valued cache of rupees, thereby enabling me to purchase the pot, and the potter to purchase a bag of rice. Expand on that transaction a million times over and one arrives at a realistic a model of global trading networks. However it is not the physical outcome of such transactions – such that two haircuts in the UK serve to underpin the bag of rice in Bombay – with which we are concerned here; rather it is with the logistics of *value* transfer, given that nothing of any physical substance found its way to Bombay in the course of the implementing that transaction. Not that there was anything unique about that. Exactly the same was true in the case of my earlier example of the outsourcer's settlement of the of his Bangalore service provider's invoice. In that case the instructions sent by SWIFT also had no substance, and there is a strong sense in which the same was also true of the way in which the resultant settlement took place - namely by means of an electronic message altering the data on the hard disks located in the computers of the Wall Street banks in which the UK bank and the Indian bank respectively held their US Treasury approved Correspondent accounts.

Hawala and the Exchange of Debt

Whilst these examples suggest that value is an extremely slippery concept, a shift in terminology such that value is reinterpreted as debt, and more specifically as an obligation by one party to make good his debt to another, we find ourselves on much firmer ground. Moreover once we do so the foundational concept of Hawala, namely that what such a transaction involves is an exchange of debt, begins to throw much more light on what is actually going on when transjurisdictional value transfers of this kind are implemented, than does more straightforward forms monetary theory. This is particularly so when it comes to appreciating the logic of a back to back swaps.

Much earlier on I set out an algebraic representation of the logistics of a one-way value transfer between a pair of cooperating Hawaladars, in which they took the opportunity to settle the debt which customer P owed to customer Q, effectively by taking the debt onto their own shoulders; hence as a result of H^b 's pay-out to Q, H^a was better off to the tune of $(x/n)D^a$, whilst H^b was short of the sum xD^b . But let us imagine that at much the same time as

←

H^a made this arrangement, H^b was approached by customer R, who had purchased a consignment of goods in far-away A, as a result of which he owed his supplier P the sum of yD^a . In these circumstances H^a would be delighted: once P had agreed to pay him $(n \times y)D^b$ to arrange the settlement of his purchase in from the supplier in A, H^a could promptly sent H^a instructions to pay out on this basis, such that his indebtedness to his partner would be reduced from xD^b to $xD^b - (n \times y)D^b$.

With such considerations in mind, the essential foundation on which the implementation of all transjurisdictional value transfers rest comes into focus. As we have seen, the principal objective of Hawala networks, and indeed of transjurisdictional value transfer services of every conceivable kinds, *is to obviate the need to physically transfer currency (as opposed to value) from one location to another*. Moreover in the long run, the sum total of transjurisdictional value transfers within such networks must by definition be zero, for obvious reasons: if a financial crisis is to be avoided, the total value of goods and services flowing into any given jurisdiction must ultimately match the total flow of value in the reverse direction. In the absence of such a balance, the exchange value of any relatively overvalued local currency will eventually decline until a more appropriate value for the exchange rate, n , is arrived at.

Hawala-style settlement procedures, no less in the present than in the past, build on precisely that insight. Hence whilst Hawaladars' front offices present retail customers with what appears to be one-way value-delivery service, the secret of their success lies in their capacity to broker a parallel transaction (or more accurately a set of transactions) which have the effect of transferring a similar quantum of value in the reverse direction. The result is what can best be described as a back-to-back swap, such that $\sum D^a \leftrightarrow \sum D^b$. So long as that outcome is achieved, the two hawaladars will be able mutually to extinguish their debts towards one another. In other words back-to-back Hawala enables collaborating Hawaladars to settle the sum of the debts which each of them have assumed on behalf of their customers by means of an exchange of debt on a consolidated basis as between themselves.

One further issue is also worth addressing at this point: how on earth to Hawaladars turn a profit when they appear to charge nothing for their services, and appear to be playing a zero sum game? The answer is quite simple: by manipulating the value which they assigned to n , the exchange rate, in the course of calculating the sum required in the local currency to precipitate a pay-out of the requisite size at the specified destination in their favour; by doing so they could both contrive to make a small (but invisible) margin by way of profit on each of

the transfers they initiated. With this in mind a further point also springs to the fore: in a globalised world order all contemporary Forex markets, whether located in New York, London, Dubai or Hong Kong are closely intermeshed. It consequently follows that it is the traders in value located in these markets who constantly renegotiate, and hence set, the open market exchange rates as between the currencies issued by, and circulating within, each and every one of the globe's financial jurisdictions – no matter how much their respective central bankers may seek to manipulate those rates to the (short term) advantage of the inhabitants of their own nominally sovereign jurisdiction.

Trust and the construction of coalitions of reciprocity

But whilst those responsible for implementing transjurisdictional transactions have a great deal of power, transactions of this kind are inherently risky: how can someone who makes a deal in one jurisdiction be certain that delivery will actually implemented in some distant sovereign jurisdiction way over the horizon? If something goes wrong in such circumstances, what remedies would be available to set things aright? To be more specific, how can an obligation to perform a financial transaction as agreed be enforced at arm's length? One way of doing so is instruct expensive lawyers to prepare elaborate contracts by means of which obligations can be enforced and malfeasance contained, backed up by squads of accountants and auditors whose task is to ensure that all such transactions have been implemented as specified. But in that case how can networks operating on the basis of mutual trust, as opposed to legally enforceable contracts manage to contain the twin challenges of systemic integrity and the detection of malfeasance? The risks involved are obvious enough. Not only do customers routinely entrust their Hawaladars with large tranches of value for delivery far over the horizon, but the debts which Hawaladars acquire viz-a-viz each other not only just as far off, but likely to accumulate rapidly as time passes. How, then, can customers be sure that their local Hawaladar will not run off with all their money – or indeed that their local Hawaladar will not be let down in a similar fashion by his partner in some distant jurisdiction?

As I have argued at length elsewhere (Ballard 2005), the analysis of the underlying logic of the risk-containment strategies deployed by Jewish traders operating across the length and breadth of Mediterranean during the course of the 10th century under the ultimate aegis of the Baghdad caliphate set out by Greif (1989) is most illuminating with respect to both these issues. Whilst Greif makes no mention of Hawala (which was at that time a specifically Islamic context), the associational networks established by the long-distance traders on whose

activities he focuses appear to have been closely congruent in character with those deployed by their Muslim counterparts operating down the Red Sea and into the Indian Ocean. Taking his cue from games theory, he concludes that:

The Maghribi traders overcame the contractual problems associated with agency relationships ... through a non-anonymous organizational framework, the coalition. Within the coalition an internal information-transmission system served to balance asymmetric information, and a reputation mechanism was used to ensure proper conduct. This reputation mechanism explains the observed "trust" relations among the traders.

The "trust" did not reflect a social control system or the internalization of norms of behaviour (although these factors play a role in any economic system). Rather, the Maghribi traders established a relationship between past conduct and future economic reward. As a result, agents resisted the short-term gains attainable through deception, since the reduction in future utility resulting from dishonest behaviour outweighed the associated increase in present utility. Since this fact was known beforehand to all traders, agents could acquire a reputation as honest agents....

Each trader benefited from being a coalition member more than he could have by establishing agency relations based upon a reputation mechanism outside the coalition. Obtaining the benefits of coalition membership depended upon proper conduct in the past, while the short-run gain from cheating today was less than the long-run benefit an honest coalition member could obtain. Since this situation was common knowledge, the merchants perceived that the agents could not do better by cheating. (Greif, 1989: 858 - 9.

However Greif also takes care to spell out the context within which these coalitions were constructed, for membership was of such coalitions was highly selective: in addition to being strongly hereditary in character (always provided that sons were careful to maintain the reputation which their fathers had established), only Jews with strong prior associations with existing members of the coalition ever appear to have been recruited as members. Hence whilst the network which Greif studied had its base in Cairo, whilst in spatial terms coalition members were residentially distributed right across the southern section of the Mediterranean, no members of the long settled Jewish communities into which their diaspora extended appear to have been recruited into the coalition. Hence as Greif goes on to emphasise

Coalition members enjoyed internal information flows that facilitated the reputation mechanism. These information flows provided the information required to uncover cheating, and contributed to the "capital premium" available to honest coalition members. Information was crucial to business decision-making; however, coalition members blocked a cheater's access to the coalition's internal information flows....

Information flows within the coalition also enabled agents to signal that they were honest. Just as modern firms hire auditors to establish the legitimacy of their financial statements, eleventh-century Maghribi agents generally conducted important business in the presence of other coalition members, including in their reports the names of those witnesses whom the merchant knew, thus enabling him to verify the agent's report (*ibid*: 880 – 81).

All the points which Grief makes in his analysis chimes closely with developments in the Indian Ocean, no less in the present than the past. Like the Maghribi traders, the resources on which Hawaladars relied to sustain their strategies of risk and reputation management, as well as their complementary processes of information transfer has typically been rooted in the prior presence of tight-knit but spatially extended networks of kinship and quasi-kinship, often yet further reinforced by membership of a specific Sufi *tariqa*. (See, for example, Ho 2006). The construction of value transfer networks within such contexts not only brought substantial material benefits to those who had gained a sufficient reputation to be invited aboard, just as Greif suggests, but also meant that malfeasance precipitated a further sanction of a kind which does not make an overt appearance in the documents on which he based his analysis: namely that any such betrayal of trust served not just to destroy a malfeasant's reputation within the coalition itself, but also that of his entire family within the wider whole diasporic community from which coalition members were drawn. In the circumstance the consequence of deliberate malfeasance (as opposed to an impersonal disaster such as a shipwreck) was not just expulsion from the coalitions, but also from access to the resources of the entire diasporic community within which it was set – an exceptionally powerful sanction in the midst of the Indian Ocean region's multiplicity of caste-, sect- and *biraderi*-based communities (Chaudhury 1985).

Last but not least, one further feature of such coalitions of reciprocity is that they are intrinsically self-regulating, with the result that had – and have – no need to look to outside agencies (such as state-sanctioned courts of law) to guarantee their financial integrity. Since they are, and always have been, well prepared to contain the intrinsic risks associated with transjurisdictional value transfers, they have continued to flourish in the midst of the contemporary world's multiplicity of sovereign financial and legal jurisdictions, and continue to do so even in those regions in which all traces of a coherent jurisdiction have collapsed – as in Iraq, Afghanistan, and Somalia for example.

The impact of European Empires on the operation of Hawala networks

During the first three centuries of their presence in the Indian Ocean region, European traders – as well the semi-Imperial institutions such as the East India Company – continued to rely heavily on indigenous Hawala-style networks to conduct value transfers on their behalf, both within and as between the new-found jurisdictions they were in the midst of founding. It was not until 1806 that the European-style Bank of Calcutta, soon to be renamed the Bank of Bengal, came into existence, but it took the two other Indian Presidencies, Bombay and

Madras, the best part of four decades to follow suit. Thereafter change was rapid. As in all other spheres, the scale of Imperial hegemony expanded rapidly in the face of the 1857 uprising, in the aftermath of which all government business, as well as the bulk of financial activities implemented by European businesses began to be directed through ‘formal channels’: European owned banks whose operations were grounded in terms of relationships contract rather than in coalitions of reciprocity. As a result indigenous service-providers found themselves steadily pushed to out to towards the edge of almost every market-place, initially India, and eventually throughout the length of the Indian Ocean and the China Seas. But despite their demotion from the position which they formerly occupied, the ‘informal’ service providers – a label which simply indicates that their trust-based *modus operandi* stood, and still stands, in sharp contrast to the ‘formal’ contractual procedures deployed by European-style banking houses still found there was plenty left to play for, not least because the services they were capable of offering to anyone prepared to seek out and to trust their service regularly found that their services were superior in quality, at least in regional contexts, to those offered by their formal counterparts. It is also worth noting that besides being able to implement value transfers much more speedily than their newfound rivals, the passage of funds through these networks was largely invisible to the Imperial authorities.

This opened up a further set of contradictions whose implications proved to have exceptionally far-reaching consequences in the longer run. Hitherto such networks operated as autonomous, self-regulating coalitions of reciprocity, governed by their own internal customary conventions; as such they were subject to little or no administrative interference. But once subject to the hegemonic structure of Britain’s Imperial Raj, they found themselves in the midst of a jurisdiction that was unprecedentedly intrusive in character: one which sought to ‘update’, and in due course to regulate, virtually all aspects of its newfound subjects’ behaviour according to its own preferred premises. And given that one of Raj’s prime concerns was to defray the cost of its administration by taxing its new-found subjects, it made strenuous attempt to ensure that all financial transactions within the jurisdiction were implemented on a contractual basis, thereby giving rise to written records, rather than on the ephemeral basis of back-to-back swaps.

Hence the 1879 Indian Stamp Act – which required details of all such transactions to be recorded on expensively stamped officially printed paper – was introduced not so much to render them legally enforceable (the basis on which the Act was formally legitimated), but rather as a means of raising revenue, as well as to undermine the competitive advantage

which informal value-transfer networks enjoyed over their heavily bureaucratized rivals, the ‘properly constituted’ European Banks. But whilst such ‘proper’ Banks soon began to enjoy a position of financial hegemony in major financial centres such as Bombay and Calcutta, as well as in their counterparts elsewhere in the Indian Ocean region, those who expected more informally constituted value transfer systems to wither on the vine were in for a disappointment. Stamp duty could readily be evaded by continuing to make as many agreements as possible on the basis of mutual trust, rather than contract; moreover since their own preferred *modus operandi* had its own well-established means of guaranteeing system security, recourse to the newly established British legal system – the ‘advantage’ which the use of stamped paper was held to precipitate – was simply an irrelevance. Nor did matters change much when India – along with all the other colonial jurisdictions which had been similarly established across the region – eventually gained their independence. In almost every case the indigenous administrators took over these newly sovereign jurisdictions retained the ‘modern’ administrative structure which they had inherited from their Imperial predecessors; amongst other things, although few admitted it out loud, the ‘steel frame’ which underpinned every such colonial administration placed those who stepped into their predecessor’s shoes in a position of untrammelled privilege – and power. To be sure many sought to use that power to steer the economy they had inherited in a more ‘socialistically oriented’ direction, in the expectation that a planned economy would precipitate rapid progress towards much greater levels of prosperity.

Unfortunately in most jurisdictions in which chose that route that route found the outcome deeply disappointing. Given the overweening power of ‘modernising’ administrations state-run initiatives grew increasingly sclerotic, indigenously inspired initiatives were suppressed, taxation regimes grew steadily more punitive, and foreign exchange crises began to loom, access to convertible foreign currencies became ever more tightly rationed. In the face of what came to be described as the licence/permit Raj (a phenomenon with parallels throughout the Indian Ocean region), India’s so-called ‘black economy’ burgeoned, as did the ‘invisible’ value transfer networks which rendered that often more productive and profitable dimension of economic activity viable, thrived as never before. To be sure the processes of liberalisation which have stimulated unprecedented levels of economic growth throughout the region wherever the most egregious features of the licence/permit Raj have been swept away. Nevertheless, the heritage of the colonial past has still not been fully swept away. Restrictions on access to foreign exchange are still in place in many jurisdictions – sometimes, although

by no means always, for good reason. On the negative side of the equation, an aversion to paying taxes, and hence an interest in squirrelling away excess capital in offshore jurisdictions, is at least as strong, and perhaps even stronger, amongst members of newly affluent elites in Asian contexts as amongst their counterparts in Euro-America; meanwhile from a more positive perspective, so long as restrictions on access to foreign exchange tie the arms of the owners of SMEs behind their backs, it is only to be expected that they will turn to alternative sources of value-transfer to meet the invoices of their overseas suppliers. So it is that for at least as many good reasons as bad ones, Hawala-style value-transfer networks continue to thrive to throughout the region.

Hawala in the Contemporary World

With this in mind five major factors, all of which are closely interconnected with one another and as well as integral by-products of the current phase of globalisation have facilitated the resurgence of hawala-style value transfers during the course of the past two decades (Ballard 2010). They include:

- The failure of the formal banking system to develop an effective, efficient and accessible global cross-jurisdictional payments system, such that both retail customers and SMEs face substantial charges when they seek to settle invoices through their banks;
- The recent explosive growth in the volume of international travel and trade, which has precipitated a further explosive global demand for a service of this kind;
- The emergence of the internet as a vehicle for instant (and secure) communication on a global scale
- The huge outflow of migrant workers, some highly skilled but mostly unskilled, especially (although by no means exclusively) from South and East Asia to Arabia and Northern America, who have in turn generated a reverse flow of remittances currently well in excess of \$US 100 billion per annum for the Indian sub-continent as a whole;
- The resultant capacity of Hawaladars to utilise a significant slice of this massive inflow of liquidity on the basis of which fulfil counterparty settlements in hard currency in every jurisdiction in the Indian Ocean region (many of which still remain heavily regulated), as well as in Dubai, New York, London and so forth.

With such considerations in mind it is clear that whilst one of the central roles which by contemporary Hawala networks still fulfil is a continuation of their ancient function, namely to provide value-transfer services for long distance traders, particularly (although by no means exclusively) in South-South contexts, much of the remainder of the operation has

changed dramatically in character. In the first place a vast new market for their services has opened up: millions of migrant workers who have established a foothold in jurisdictions other than their own, no less in the global South than the global North, with an urgent need for a means of remitting a substantial portion of their earnings swiftly, safely, and above all with the minimum degree of overhead loss to their wives and families back home. Secondly, and just as significantly, the shape and size of these remittances differs significantly from those implemented by merchants in settling their commercial invoices. On the one hand migrant remittances are much more numerous, and hence dramatically smaller – averaging out at no more than a few hundred dollars a time – than commercially grounded settlements; and on the other few if any the army of recipients to whom these remittances are directed are based either in the central business districts or industrial centres of major conurbations. To be sure some families may have established a foothold in the chaotic periphery of one or other of the developing world's mega-cities; however the great majority of such recipients are still based way out in the sticks, since it is restricted opportunities in more rural contexts which constitute the principle driving force behind contemporary flows of transjurisdictional migration.

As ever in financial contexts, the rewards associated with economies of scale are enormous. But although the total flow of migrant remittances on a global scale may well now amount to as much as a trillion dollars per annum, especially when all the transfers which fail to register in statistical bean-counter's reports to the World Bank are included, it is a market which mainline Euro-American banks have found it almost impossible to service, despite the initially tempting prospect of tapping into the profits potentially available in a trillion dollar value transfer market. The central reason for their disappointment is not hard to determine: despite the huge size of the total sum of value which migrants remit on a global scale, not only is the size of each of millions of packets of value dispatched by migrant workers worth no more than peanuts from a banking perspective, but the logistics of implementing deliveries to a host of remote rural destinations also amounts to a nightmare, at least in terms of their own established organisational procedures. Most Banks have consequently dropped out of the remittance business – if, indeed, they ever made a significant effort to provide such a service for migrant workers.

Meeting the logistical challenge

As we have seen, the foundation on which the implementation of all Hawala-style value transfers lies in the capacity of Hawaladars operating in the context of an established

coalition of reciprocity to swap consolidated tranches of debt as between themselves. In the light of that exchange – and very often simply the prospect that that such an exchange will in due course be brokered as between themselves – each Hawaladar will utilise the sums of local currency deposited by customers with them with instructions that they should be transferred as value overseas to implement a parallel set of pay-out instructions in the local currency to further set of customers as specified by his overseas partner, in what is best identified as a pair of back-to-back swaps; and whilst contemporary Hawala networks utilise just the same principles as those deployed in the ancient world, those operating in the contemporary world not only operate on a much vaster scale than any of their predecessors, but also face a much more complex set of logistical challenges as they set about coping with the discrepant needs of a widely varying and globally distributed clientele. Moreover the demands of their customers are equally challenging: what they are looking for is a service able to implement value transfers, no matter how small, to remote destinations in distant jurisdictions on a swift, cheap and absolutely reliable basis.

Contemporary hawaladars have relied on two complementary sets of resources as means of meeting a set of logistical challenges which have largely defeated mainline banks. On the one hand on their capacity to construct globally extended self-regulating coalitions of reciprocity based on precisely the same principles as those outlined by Greif; and on the other hand by making the most of current electronic technologies in the processes acquiring and distributing data as between themselves, such that they are able to negotiate and execute a web of individually negotiated transfers and swaps on a real-time basis. By bringing these two arms together they have managed to construct highly effective – but from a Euro-American perspective an intrinsically informal – global payments system, implemented within the context of a series of acephalous but globally distributed networks of reciprocal trust.

With such considerations in mind the tasks which these networks are required to fulfil are multiform, but can conveniently be broken down into main categories. On the one hand those which are broadly physical in character, beginning with the conversion of each customer's funds from one currency into another, followed by securing the delivery each of those re-denominated packets of value into the hands of distant recipients; and on the other hand the issue and distribution constantly re-tailored instructions sets to network members responsible for implementing all the many dimensions of these value transfers. In the midst of all this we also need to remember that central function of all such networks is to transfer *value*, not money. Moreover as we have seen, the object of the whole exercise is to ensure that in

general *currency* stays put in the jurisdiction in which it is current, and is simply swapped between senders and recipients who will in general have no knowledge whatsoever of each other's existence – any more than has one any knowledge of through whose hands the cash dispensed by an ATM had previously passed.

The implementation of back-to-back swaps

When we consider each of these processes step by step, each turns out to present a distinct set logistical challenge. With respect to the issue of currency conversion, we have already seen that economies of scale are such that all significant financial operations are best done in a consolidated basis, given that the marginal cost of implementing the exchange of a mass of penny-packets of cash or value is immensely time consuming, and hence generates a prohibitively expensive level of overheads. As a result contemporary hawala networks are invariably ordered in terms of a hierarchy of layers, at the bottom of which a multiplicity of retail hawaladars take small-scale deposits from and/or implement pay-outs to retail customers, typically migrant workers sending remittances back home. It follows that retail hawaladars not only interact in face to face terms with to the customers, but since each such normally provides a services to members of a local ethnic colony all of whom are drawn from the same equally specific locality overseas, he can readily complete the initial stage of a value swap by sending a fax setting out delivery instructions to a parallel bunch of retail hawaladars at that destination with whom he has established cooperative relationship. But by the very nature such transfers are invariably strongly asymmetric: at this stage in the game number of migrants sending money home to support their kinsfolk, to pay for marriage celebrations and build houses back home far outnumber those wishing to repatriate the profits accrued from selling land or other familial assets in their home village to the alien destination in which they have established only recently established themselves. In other words there is very little prospect of retail hawaladars dealing with customers of this particular kind being able to set up comprehensive back-to-back swaps as between themselves. Indeed since the value transfer runs overwhelmingly in one direction, and since the delivery Hawaladar pays out as instructed to each of the recipients named in the fax which his sending partner routinely fires off at the close of business every day, if nothing else was done the distributing Hawaladar H^b would find himself ever more seriously indebted to H^a (i.e. $\sum D^a \rightarrow \sum D^b$), with no sign of H^b being able to access reciprocal remuneration anywhere in sight.

A further salient feature of contemporary Hawala networks is that the transmission of delivery instructions (or in other words those of the kind which H^a faxed horizontally to his

distributing partner H^b) often follows a completely different route from that followed by the value transfer instructions, which – in keeping with the premise that financial matters economies of scale invariably precipitate massive commercial benefits – proceed upwards through a hierarchy of consolidating Hawaladars until they reach a sufficient size for fully fledged Hawala swaps to be brokered, typically in one or more consolidated tranches of value, each of which is worth 100,000 US dollars, or failing that of a similar sums of either Euros or pounds sterling.

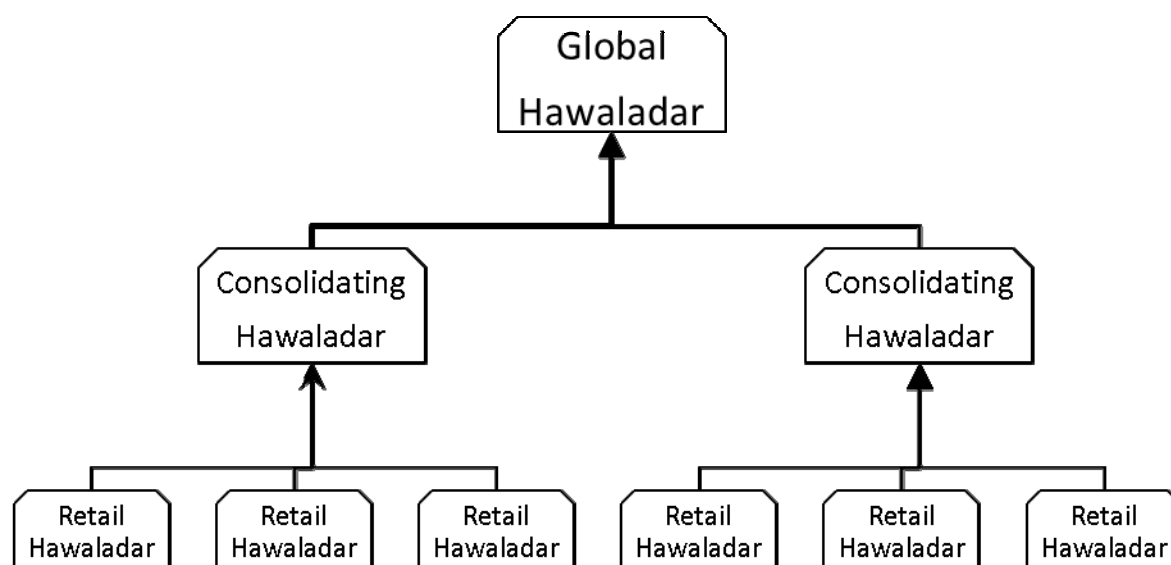


Figure 1 Outline structure of a value consolidation hierarchy in contemporary Hawala networks

Hence whilst each of the retail Hawaladars in the diagram will fax off a bunch of delivery instructions to his disbursing partners on the overseas on a daily basis, he will similarly despatch the funds he has collected from his local customer to a consolidating Hawaladar, either in cash, or by means of a bank transfer, together with a note to the consolidator as to what portions of that sum are owed to each of his disbursing hawaladars. A similar process of consolidation may go for one or more stages until it reaches the top of the tree, at which level meg-swaps, in which swaps of involving tranches made up of multiples of 100,000 are mixed and matched on a global basis. However consolidating is one thing, but mixing and matching – settlement in other words – is an entirely different operation. How, then, is that achieved?

Whilst a further diagram provides the best way of illustrating what goes on, and for simplicity's sake, I have not only assumed that the Hawala transfers in question are solely restricted to value transfers passing back and forth between in the UK and Pakistan, followed up by another even more unrealistic assumption, namely that shape of the (already greatly

simplified) the hierarchically ordered consolidation and deconsolidation networks in each location are identically structured. If that were so, then we can get a good grasp of how back to back swaps are implemented in this context by setting out a double mirror image of the diagram which we have just been considering, as I have done in the diagram below.

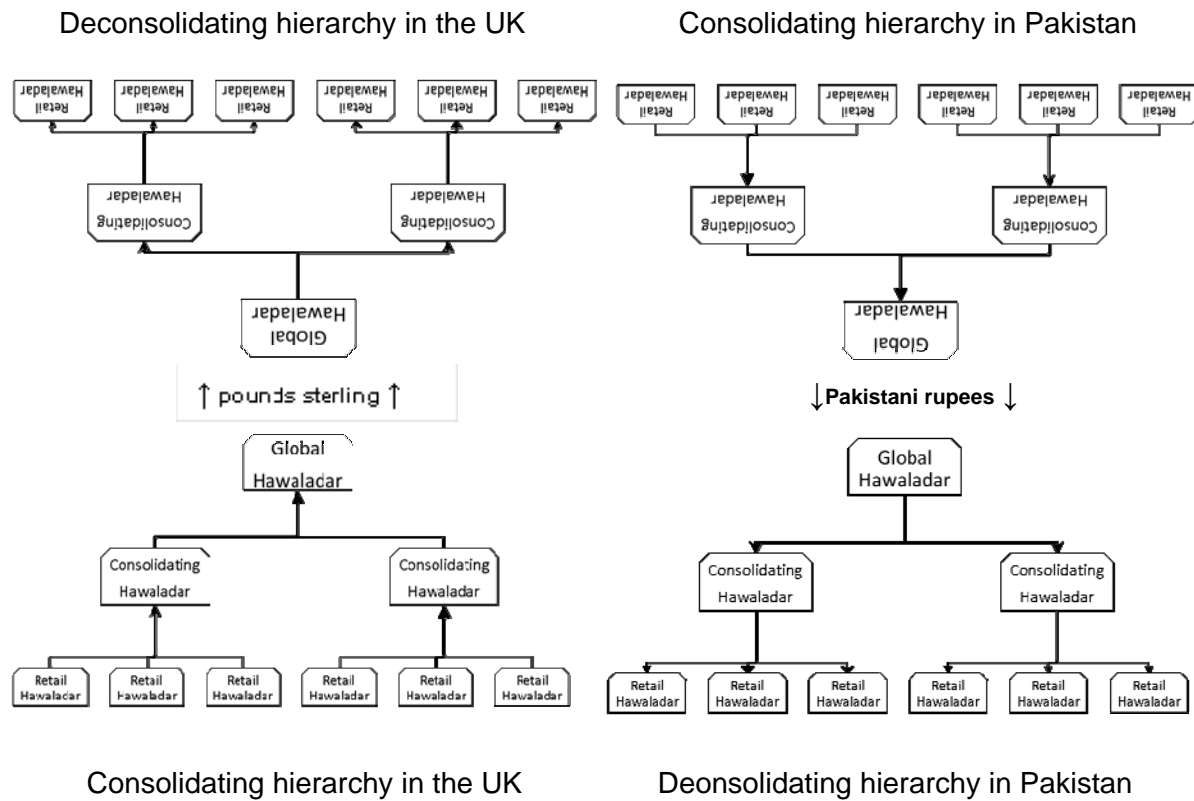


Figure 2 The logistics of contemporary Hawala-style transjurisdictional back to back value swaps

With structures with this kind in mind we can begin to unlock the highly sophisticated logistical basis on which contemporary Hawala-style transjurisdictional back-to-back value swaps are comprehensively implemented. So far we have only taken cognisance of how one relative small (but nevertheless vital) arm of the system works: namely the faxed transmission of disbursement instructions from retail Hawladars in the UK (in the bottom left) to their Pakistani counterparts (in the bottom right) who were about to implement disbursements to specified recipients as per their instructions from the UK.

Whilst all this was going on, another wholly independent set of Hawaladars in Pakistan (located in the top right) would have been busy preparing a series of reverse trades, typically orders from businessmen and SMEs based in Karachi, which lies the best part of a thousand miles away are from the villages in northern Pakistan, where H^b and his ilk is most likely to

be located as he sets about implementing H^a's disbursement instructions. To the extent that the orders received in Karachi would for the most part be placed in order to settle invoices for goods and services which had been imported from the UK, each transaction would on average be a great deal larger than the average size of a migrant's remittance, hence very many fewer orders would normally be required in Karachi to generate a similarly sized tranche of value to that available in the UK. What this would constitute, however, was a set of processes running in parallel to those which I have already outlined with respect to developments in the two lower quartiles, although in this case running in the reverse direction: hence in the two upper quartiles delivery instructions run systematically from right to left, rather than from left to right, whilst the processes of cash consolidation in rupees, as well as its subsequent deconsolidation as it is broken up in the process of being distributed to the disbursing Hawaladars (as well as instructions setting the precise basis on which this should be done) be done pass upwards in the two right hand quartiles, and in the reverse direction on the opposite side of diagram.

Once this process is complete, the way is at long last open to implement a comprehensive settlement swap, which as like as not will have been negotiated by a set of brokers in Dubai, such that the sending Global Hawaladar in the UK has assembled five tranches of value, each worth £100,000, which he wishes to swap for a consignment of Pakistani Rupees of similar value, on behalf of all the many customers in the hierarchy below him. Meanwhile a Global Hawaladar in Pakistan will have lined up a parallel set of customers who have placed series of orders for the purchase of a series of tranches of sterling in exchange for rupees, at a price of Rs 138 to the Pound, given that he has lodged a preliminary bid of Rs 135 to the pound for the purchase of the £500,000 on offer by the counterparty in London. On this basis a back-to-back swap can readily be brokered, but in a manner which no approach is made to the foreign exchange market *per se*, since no physical movement of currency into or out of either jurisdiction, or in other words along a horizontal vector of my diagrammatic model.

All this reveals the logic of carefully negotiated and implemented logistical steps which underpin the daily articulation of back-to-back Hawala-style value transfers, implemented, and constantly renegotiated, within the context of transjurisdictionally extended coalitions of reciprocity, and hence largely on the basis of a verbal say-so, since data transfer within the network is largely restricted to the delivery instructions required to accurately implement that portion of the overall swap which is being handled by any given pair of Hawaladars. Hence in formal terms this is very much a distributed system: no one participant in the exchange has

a comprehensive overview of the intricacies of the entire operation. It is also worth emphasising that the scale of such swaps can be expanded almost indefinitely, not just by drawing in more and more Hawaladars at every level of my simplified model, but also recruiting participants from a multiplicity of jurisdictions into any one such swap. Moreover at the top-most levels of consolidation swaps are rarely if ever implemented in terms of cash: instead most players at this level now have direct access to SWIFT, transfers of tranches of value in multiples of 100,000 – whether in denominated US dollars, euros, or pounds sterling – are routinely transferred electronically through the formal banking system; moreover since inter-bank transfers between these currencies are subject to there are few if any exchange controls, especially in offshore markets, such mega-transfers can readily be redenominated in any of these currencies, depending on which proves to be the most convenient delivery vehicle.

Last but not least it is also worth emphasising that whilst the coalitions of reciprocity which underpin this system are globally distributed, and are all ultimately held together by the collective need of all participants' to maintain their reputations for honesty, since a failure to fulfil their obligations, and indeed any other form of malfeasance, would promptly lead to their exclusion from the game, it most certainly does not mean that everyone involved in any given network has first-hand knowledge over everyone else. Rather one can only begin to participate in such networks if one has – or failing that, if one can establish relationship of trust with established participants in the system, no less along vertical as well as horizontal vectors. Hence system security effectively relies on established participants vouching for new entrants. But once a new entrant has been taken aboard, there is every prospect that he will in due course climb up through the hierarchy, as a result of expanding the range of those who are willing to trust him on the one hand, and of his skills in recruiting custom and cutting doing deals on the other. Moreover the personal, as well as constantly renegotiated character of each and every link in the system is a significant bastion against systemic failure. Since settlement is implemented on a daily basis, not only does failure to deliver become instantly apparent, but the source of failure is readily identifiable in *personal* terms. Likewise if there is no good reason for such a betrayal of trust, an instant sanction is available: given that there are a multiplicity of mutually competitive nodes at every level in the system, and a ready means of communication within the coalition as a whole, he can readily be frozen out of the system by all his partners the very next day, by the simple expedient of routing all future transactions through his competitors until such time as he makes good the deficiency, and

sets about the onerous task of rebuilding his reputation for trustworthiness. No get-out clause on the basis of limited liability is available in this context.

It follows that the structures to which the application of these principles give rise are anything but static behemoths: rather the whole system operates on a cut-throat business. It is not just that there are invariably a multiplicity of retail outlets in every market where Hawala facilities are on offer, such that customers to pick and choose between operators as they look out for the best deal; similar negotiations also occur at each and every level in the process of consolidation and deconsolidation take place, as those with tranches of value to transfer constantly search the market to establish which consolidator is prepared to offer them the best deal. Nor is it necessary for the settlement swap to be implemented by the top-level Hawaladars based in the current global hub in Dubai; there is nothing to stop consolidators at a lower level in the system cutting corners by negotiating a direct deal with a fellow lower-level consolidator who has assembled a tranche of value which neatly complements his own, thus cutting out the more senior middlemen. It follows that when such networks give rise to a settlement 'pulse' on a global basis every twenty-four hours, they can promise their customers something close to real-time delivery on an exceptionally lean and mean basis. From this perspective it is quite clear that a combination of competitive entrepreneurship and a consequent commitment to flexibility is a large part a secret of their success. In the context of distributed (and hence acephalous) system/market-place of this competitive advantage springs from players' capacity to perform their chosen role in the logistical network more adroitly, and above all more cheaply (but just as reliably) as their competitors. In doing so significant benefits arise from three main sources. Firstly from systematic efforts to render redundant (and hence remove the costs associated with) all forms of data storage and information transfer which are not specifically germane to the speedy and accurate implementation the logistical tasks immediately to hand; secondly from implementing information transfer more speedily than one's rival; and thirdly from the sophistication with which one is able to 'playing the market' viz-a-viz those Hawaladars with whom one does business.

Two further points follow from this. In the first place what has emerged on this basis is anything but a static behemoth. On the contrary ever element in the system is subject to constant renegotiation, with the result that the fine structure of each daily settlement 'pulse' differs significantly from that of its predecessor, since deals at every level in the system are constantly renegotiated, not least in response to changing demands in the local market.

Secondly, and just as significantly, all players in this arena have taken almost instant advantage of developments in communication technology, whether in the form of the introduction of international direct dialling facilities, the internet, mobile phones, and last but not least satellite phones for those operating in the back of beyond in failed states.

As a result, contemporary Hawala style transjurisdictional value transfer systems have become market leaders in a wide range of contexts where established Banking facilities are either unavailable or so tied up by regulatory restrictions that they are unable or unwilling to provide many of the financial services to which their customers wish to gain access – as is frequently still the case in the global South. However as we have also seen, Hawaladar-style services also enjoy a significant competitive niche in the global North, not so much as a result of the absence of banking services, but rather because the logistics of their transjurisdictional value transfer services are organised on such a sclerotic basis that their overhead costs for small-scale transfers – i.e. anything much less than £50,000 – are such that their fees for implementing the whole operation take a significant chunk out of the quantum of value being transferred. Hence the service on offer from Hawala-style networks are financially attractive to a wide range of customers, including migrant workers, long-distance travellers of all kinds, and SMEs with a significant tranche of customers in the global South.

Back-to-back swaps in a comparative perspective

Whilst contemporary Hawala networks are undoubtedly distinctive in the sense that they are as well able to facilitate the transjurisdictional transfer of value all the way scale from packets of £100 up to substantial multiples of £ 100,000, and also that they underpin their settlement networks in relationships of mutual trust, rather on the basis of formal contracts, which are in turn expensively archived at both ends of the process, and in all probability at many other way-stations in between, there is nothing unusual about their underlying logistical model. All transjurisdictional value transfer operations, whether conducted as between of within formally constituted Banks, by the Treasury Departments of multi-national corporations, or indeed Western Union are faced with much the same logistical challenge, and at least in broad structural terms, arrive at much the same solutions. Nor is there anything unique about their use of back-to-back swaps as a means of implementing transjurisdictional settlements. Although it is unlikely that Dubai's Hawaladars were aware of it, mainline bankers in London were simultaneous developing an almost identical strategy as a means of coping with the UK's recently instituted exchange control regime. As Coyle notes

Currency swaps originally were developed by banks in the UK to help large clients circumvent UK exchange controls in the 1970s. UK companies, seeking to finance their US subsidiaries' operations, were required to pay a premium (known as an exchange equalization premium or the investment premium) when obtaining dollar loans from their banks. The idea behind a swap was to avoid having to pay this premium for dollar borrowing. A bank would identify a UK-based organization that wanted to borrow in dollars and a US-based organization wanting a sterling loan. An agreement would then be negotiated whereby

- the UK organization borrowed sterling and lent it to the US company's UK subsidiary
- the US organization borrowed dollars and lent it to the UK company's US subsidiary.

No sterling left the UK and no dollars left the US, so there was no requirement to pay the exchange equalization premium on dollars. The benefit of cheaper borrowing could be shared between both parties to the swap, and both the US and UK organizations therefore could pay less for currency debt liabilities than if they had borrowed directly in the foreign currency (sterling or dollars). These early arrangements were called “back-to-back” or “parallel” loans, from which more sophisticated currency swaps were gradually developed over time. (Coyle 2000: 24 – 25)

However there is a striking difference between the swaps engineered across the Atlantic back in the 1970s – which in due course provided the foundation of the multi trillion dollar shadow banking crisis based on the exchange of Credit Default Swaps, and whose implosion in the latter part of 2008 came close to collapsing the greater part of Euro-American financial order – and the back-to-back swaps routinely implemented through Hawala networks. For whilst Hawala swaps are conceptually grounded in an exchange of debt, the ‘debts’ which Hawaladars owe each other are best understood as the obligations which they owe to one another as a result of their mutual transactions, and which are fulfilled in due course when a settlement takes place. However the packages which pass through the logistical system they have constructed, and which give rise to these obligations are manifestly packages of *positive* value, rather than packets of *negative*, but nevertheless interest-generating, value, as was manifestly the case with respected to the sliced, diced, and securitised sub-prime mortgages out of which the tangled, ever-growing, and ultimately unstable heap of Credit Default Swaps was constructed. But what is equally striking is that although those all those swaps were underpinned by carefully articulated contractual agreements which in principle provided the steel frame around the whole edifice was constructed, it was not so much breach of contract which brought the system to its knees, but lack of trust. Moreover the collapse in mutual trust proved to be *systemic* in character, as John Lanchester demonstrated in his wonderfully entitled exploration of the causes of the collapse, *Whoops!: Why everyone owes everyone and no one can pay out* (2010).

However financial institutions in the Asia-Pacific region, whether formally or informally constituted, remained largely untouched by the chaotic consequences of their counterparts' uncritical pursuit of ever more arcane forms of financial engineering. Besides being more far concerned with facilitating the production and distribution of goods than with the derivative profits which could potentially be derived from the application of the Gaussian Copula, Asian financiers also avoided the contemporary Euro-American commitment to the unbridled application of the principle *caveat emptor*, and their parallel practice of hiring expensive lawyers not just in an effort to ensure that their interests are adequately protected on a contractual basis, but also make maximum usage of regulatory loopholes, regardless of the potential systemic consequences.

All this stands in sharp contrast to those financial systems – of which contemporary Hawala-style value transfer networks are but a specific example – in which players prefer to conduct their operations on the basis of transfers of positive, rather than negative value, and to rely on the construction of coalitions of inter-personal reciprocity grounded in relationships of personal trust (rather contracts which can all too easily be worth no more than the paper on which they are written when a bubble explodes) as a their preferred source of systemic security. Could it also be the case that just as the engineering capabilities deployed by innumerable Asian enterprises have precipitated the emergence of rust-belts in many once-prosperous manufacturing centres throughout Euro-America, their financial counterparts' preferred organisational strategies are also facilitating the construction of transjurisdictional financial services initiatives which is far more fit for purpose in the current phase of globalisation than the unstable and sclerotic services currently provided by the august institutions based in Wall Street and the City (of London)? As the ancient Greeks were only too well aware, Hubris routinely invites the attention of Nemesis.

Considered from a global perspective, there appears to be every indication that as a result of its enthusiastic adoption of the mathematically grounded (but deeply myopic) premises of the Chicago School, the 'rationally' and 'scientifically' grounded debt manufacturing and manipulation exercises which has brought the Euro-American financial services sector such enormous profits during the past few decades is currently collapsing under the weight of its own internal contradictions? But if I am right in so concluding, this will certainly not the collapse of the global economic order: all we are currently witnessing is yet another dimension of the shift in the location of its centre of gravity from West to East, accompanied – as was equally inevitable, a severe revision of the conceptual premises in terms of which

the financial order is grounded. If that is what twentieth century is destined to be all about, parochially minded Euro-Americans will have to do a great deal of re-thinking, no less in this sphere than many others, if they are to have any chance of keeping up with the competition.

Bibliography

- Ballard, Roger (2010) "Three Phases of Globalisation: the significance of Dubai's emergence as a trading hub" in Pieterse, Jan (ed) *21st Century Globalisation: Perspectives from the Gulf* Dubai: Sheikh Zaidi University
- Ballard, Roger (2009) "The Dynamics of Translocal and Transjurisdictional Networks: A Diasporic Perspective" in *South Asian Diaspora* Vol 1, (2) 141-166
- Ballard, Roger (2006) "Hawala: criminal haven or vital financial network?" in *Newsletter of the International Institute for Asian Studies*, (September 2006 issue) University of Leiden, The Netherlands p. 8 - 9
- Ballard, Roger (2005) "Coalitions of Reciprocity and the Maintenance of Financial Integrity within Informal Value Transmission Systems: the operational dynamics of contemporary hawala networks" in *Journal of Banking Regulation* Vol. 6, 319 - 352
- Chaudhuri, K. N. (1985) *Trade and Civilisation in the Indian Ocean: An Economic History from the Rise of Islam to 1750* Cambridge: Cambridge University Press
- Coyle, Brian (2000) *Currency Swaps* London: Financial World Publishing
- Frank, A. G. (1998) *ReORIENT: Global Economy in an Asian Age*. Berkeley: University of California Press
- Greif, Avner (1989) "Reputation and Coalitions in Medieval Trade: Evidence on the Maghribi Traders" *The Journal of Economic History*, 49, 857-82.
- Ho, Enseng (2006) *The Graves of Tarim: Genealogy and Mobility across the Indian Ocean* Berkeley: University of California Press
- Lanchester, John (2010) *Whoops!: Why everyone owes everyone and no one can pay* London: Penguin
- Maimbo, S and Ratha, D eds. (2005) *Remittances: Development Impact and Future prospects* Washington DC: World Bank pages 103 - 118.
- Schacht, Joseph (1964) *An Introduction to Islamic Law* Oxford: Clarendon
- Smith, Adam (1993) [1776] *The Wealth of Nations* Oxford: Oxford University Press
- Udovitch, A (1970) *Partnership and Profit in Medieval Islam* Princeton: Princeton University Press