SANDPAPER POLITICS[1]: THE BLURRING OF THE SHARP EDGE OF ECONOMIC POLICYMAKING

Karl Cronin

Junior Sophister

The worlds of economics and politics are impossible to separate, but the information asymmetries involved in divesting the power to make economic decisions to our politicians often lead to disenchantment. Karl Cronin delves into the philosophical background of the social contract. He examines the role of interest groups, the media and bureaucracy and argues that ideology and party politics are essentially incompatible.

Introduction

> 'What is curious is not that we are content to delegate to others the hard business of taking decisions, but that we choose to give authority over our lives to people who so many of the population seem to think are a bunch of charlatans' (Paxman, 2003, p.13).

Statistical studies show that the western world's economic success can be somewhat attributed to the quality of its institutions. Knack and Keefer (1995) argue that the strength of civil liberties backs up property rights, which allows democracies to progress economically. Hall and Jones (1999) argue that social capital is an influence, whilst Acemoglu and Robinson (2005) point out that inequality was reduced to a certain level in western institutions, allowing democracy to consolidate. These studies point to the West as being a free and fair place, yet people in general feel disillusioned with the political and bureaucratic policy outcomes that determine the course of their lives.

This paper, by focusing on what the basis is for this disillusionment and whether it is rational, will posit an explanatory philosophical reasoning as to why both societies and individuals willingly adopt

[1] Based on the opinion "His first instinct when confronted with most choices was to blur them. Blair was a sandpaper politician. Sharp edges were to be smoothed" (Rawnsley, 2001, p.315).

such institutional control, before criticising the nature of policymaking. In order to achieve this, explanations of the political process will be offered with a critique of EU interest group representation and an analysis of both the media and the civil service.

Philosophical Background

Foucault (1991) argued that western, liberal neo-Christian democracies are self-regulating panopticons which grew from Bentham's utopian ideas of regulation, supervision and punishment. We are trained from birth to conform (through supervision) into bureaucracy which we allow to determine our lives. We contribute by and are rewarded for our conformity, and punishment trains us when we digress. Popper (2002, p.60) conjectured that we are born with the innate 'propensity to expect regularities and to search for them', thus such a regulated society suits our nature. If Foucault and Popper were correct in their observations it is little wonder that we delegate collectively and seem individually disillusioned with outcomes at the same time. Good governance then is what is best collectively for society's benefit, yet individuals in government are also influenced by what may be in their own individual self-interests.

The Political Process

The political principal-agent problem comes about as politicians are our agents yet we feel disillusionment because once politicians gain power the relationship seems reversed – the government (as our agency) ceded a tenancy of our secondary rights to arbitrate on our behalf within the spheres of reparation, redress and defence (Locke, cited in Rawls, 2007, pp.117-121). If they are deemed to have misused our rights an individual has only one vote in the next election to try to restore balance in the relationship. An individual may be correct in his analysis but this might be hard to argue or publicize due to the asymmetric nature of information. Although 'information' is a public good, beneficial for all, it can be costly to acquire (Hillman, 2009, p.74) and disseminate. Due to the nature of our single vote, which is unlikely to determine outcomes, it is irrational to seek complete information, based on the 'income utility' which a non-deterministic vote brings (Downs,1957, p.147). The cost here outweighs the lack of individual benefit. This is not to say that voting is not important, but that a certain amount of 'ignorance' is 'rational', most especially when information, such as taxation policy, requires expert analysis that is costly to gain (Hillman, 2009, pp.75-77).

Downs (1957, p.142) argues that 'each party invents an ideology in order to attract the votes of those citizens who wish to cut costs by voting ideologically'. Voters may also choose parties with local or family roots, or parties that are better for a particular industry or single issue. Parties are therefore judged on their ideological performance by voters whose own benefits don't necessarily pigeon-hole into that party's preferred outcomes. Similarly, politicians choose parties. Ideologies change as situations arise,

with party-members whipped to vote certain ways in parliamentary democracies. Politicians often blame the whip when individual ideologies digress from that of the party. Voting against the whip on an ideological basis can stem career progress.

In multiparty systems governments usually take the form of minority governments or coalitions and therefore 'a multiparty system offers voters an ostensible choice between definite, well-integrated policy-sets in each election, but only rarely does one of these sets actually govern' (Downs, 1957, p.144). If voters have voted to reflect an ideology or strong belief in certain issues, the resulting policy outcomes, however they are spun, might not reflect their individual preferences. After electoral punishment, politicians will still have to negotiate policy outcomes with similar results. In a two-party system, because choice is limited, deviations from expected policy may not bring as much electoral punishment, as within such systems choice will be limited at the next election (Ferejohn, 1986, pp.19-20). Voting for the opposition might be a step too far, given general ideologies.

> 'In political life, perhaps the most basic incentive comes from the need to be re-elected' (Besley, 2004, p.196).

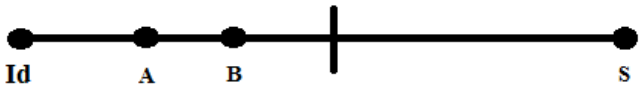
Besley (2004) shows (using data from U.S. states where Governors have two terms and different state dependent salaries) that salary levels only 'weakly' reflect performance in regard to policy outcomes, yet does acknowledge that a political career can affect future earnings. Politicians are faced with the dilemma that elections cost money and successful careers are longer than one term. Besley and Case (1995, p.793) show that 'incumbents who are eligible to run again care about building their reputations' and do so by forming policies more to electoral wishes. Zupan (1990) shows that long serving Congressional Representatives do change their ideological positions in their final term, but due to the successful nature of their careers 'natural selection' has proved them 'fitter', so there is no real difference between their final term voting behaviour and that 'displayed by non-retirers', yet compared to their previous voting patterns they do vote differently. What all three studies implicitly point out is that politicians often form policies for reasons other than their electorate's preferences, becoming more the principal than the agency of their electorate's ideology. Thus voters justifiably feel disillusioned with their chosen representatives.

Interest Groups in the EU

Within the EU policymaking machine there is a recognized democratic deficit, which has been filled by interest group representation. In theory, if one politician accepts support from an interest group,

and, whether through electoral donations (Hillman, 2009, p.78), technical information or electoral support, they are deemed to be at an electoral advantage, other politicians will rationally follow.

In Europe the Commission is not an elected body. The Burson-Marsteller Report (a professional lobbying firm) indicates that the Commission and parliamentarians rate business lobbyists as more successful than NGO and civic lobbying, with NGO lobbyists rated more successful by elected parliamentarians. The report points out that decision makers trust their unelected bureaucratic staff more than lobbyists, thus bureaucrats are targeted (Burson-Marsteller, 2005). Lobbying restrictions are the weakest regulations of any regulated political system (Chari, Hogan and Murphy, 2010, p.109). It can be argued that lobbying regulations, rather than being to guard the political process from influence, are in place to add legitimacy to a legislated process.



In the above illustration (based on Hillman, 2009, p.79), party or national ideologies are represented at point Id, whereas a special interest group's policy outcome needs are expressed at point S. If an interest group provides politicians with support on a policy issue, they will at least move to point A; the benefit gained is equal to the vote loss from marginally moving from party or indeed national ideology. Any further benefit gained would enable politicians to move closer to the interest group, at point B or further, depending on the policy.

> 'Article 8a in the Lisbon treaty highlights the functioning of the European Union as a representative democracy and Article 8b provides for the incorporation of representative interest associations and civil society organisations in EU policy-making' (Saurugger, 2008, p.1282).

> 'While most groups are only consulted by policy-makers, under certain circumstances they may formally take part in some government activities or be legitimised by some sort of subsidiary principle' (Beyers, Eising and Maloney, 2008, p.1107).

The subsidiary principle here, legislation with light-touch regulation, leads to the situation where EU policymakers in need of information and expertise begin towards point A, rather than at their own, their government's or their party's ideological point. Interest groups are a legitimized political

constituency. It is then rational to move to point B, as parliamentarians need access to electoral voting blocks, such as farmers' groups or union movements, whilst commissioners and bureaucrats need access to the information interest groups provide.

Due to the fact that interest groups are smaller than the population they can co-ordinate their actions in a focused way electorally and they can be seen to have more to gain within single policy outcomes (Hillman, 2009, p.80). In Europe, industry lobbyists do not 'bargain', as has been perceived but appeal to the 'better argument' process in negotiation (Naurin, 2007, p.224). Civic group lobbyists represent institutions that are not necessarily democratic in structure, especially 'insofar as the internal structures of these organisations are rarely internally democratically accountable' (Saurugger, 2008, p.1279). In policy forming, decision makers rely on their unelected bureaucratic staff, who in turn have been lobbied (Burson-Marsteller, 2005, p.12) by all sides.

What are left behind are the European electorate's ideologies, which are being represented by undemocratic forces of both the left and the right. Meanwhile policy decision makers start from a point closer to the centre due to the legislative legitimacy of the lobbying process which allows the agency of minorities an advantage over the agency of their overall principal, the electorate. The logical outcome of this is the disillusion of that electorate with the decision making process at EU level.

The Fourth Estate

'Once, MPs set the agenda. Since then, the journalists have taken over the political arena... Once, Parliament controlled the Executive. Today, the Government doesn't listen to us but to a fourth estate which obsesses it... My job has been deskilled. All the interesting bits have been taken by journalists... ability, style, originality... What counts in politics is the crawl up the ladder' (Mitchell, A., cited in Paxman, 2003, p.285-286)

Some asymmetry problems can be corrected by the media, who as a 'key branch' of civil society (Besley and Burgess, 2002, p.1446) can act as 'objective' (Hillman, 2009, p.82) information providers. Of course the media are lobbied by special interest groups as well as politicians and individuals, and as information sources can be legitimately kept confidential, it can be hard for an electorate to realize that journalists often use 'political opportunism' which arises for oppositions when policy outcomes are less ideologically placed than proportions of the electorate would prefer (Hillman, 2009: 82).

Yet the media does enable politicians to be lauded for correct decisions, such as in India where a correlation has been shown between states with more newspapers and famine responses (Besley & Burgess, 2002). Media freedoms also correlate positively with high levels of political knowledge and participation as well as voter turnouts (Leeson, 2008). Yet, due to the manipulative nature of journalism, its ability to be controlled by influential owners and the state's interest in controlling it 'when the

importance of information to democratic decision making [is] at its highest' (Gentzkow and Shapiro, 2008, p.149), the media perform at their best as information providers when competition is high, deregulated, and media channels diverse (Gentzkow and Shapiro, 2008). It is up to citizens to piece the information together, yet due to its asymmetric nature this is difficult to achieve. It can be difficult to ascertain on whose agency the media are acting. This is reflected in contradictory but mainly cynical attitudes of much of the electorate towards the media.

The Bureaucracy

In theory, bureaucrats are the agents of the electorate, yet within the electoral process they become the government's agency. As monitors of the social contract, bureaucrats monitor citizens in such areas as policing or tax regulations, which are policy outcomes. This essentially places them as principals, thus we reach another principal-agent problem. Again, due to the asymmetric nature of information, it is difficult for citizens to monitor these bureaucracies. The bureaucracy is run on a day to day basis by unelected officials. These officials have their own preferences and benefits that they can advance by influencing policy outcomes on an ideological basis. It is impossible for citizens to acquire and understand the information required to judge bureaucracy; this problem is backed up by the fact that opposition parties normally have more limited access to the machinations of bureaucracies than incumbents.

To complicate matters further, when monitoring bureaucracies within the setting of performance goals, 'it is important not to confuse a focused effort with a productive one' (Heckman, Heinrich and Smith, 1997, p.394). It is beneficial for managers to have large staffs and higher budgets in relation to their careers. While managers may be engaged by performance related issues, there is no guarantee that staff will follow. Whilst certain civil servants might be seen to be 'Thomas à Becket' characters, putting public interest over 'personal self-interest' (Hillman: 2009, p.107), their ideas on what is in the public interest will be influenced by their ideological self interest. Similarly, committee members that regulate bureaucratic policy outcomes suffer from the same natural inclinations. Civil servants are rational if they protect deficiencies, as their careers are on the line. Thus it is rational to be disillusioned with a bureaucratic layer and its decision makers.

Conclusion

In conclusion, individuals in society are born with an innate ability to seek and recognise regularity, which implies that regularity is a psychological positive whilst implicit dissimilarity may be a negative; therefore the individual conforms by nature. Society itself is a regulative institution which takes advantage of this conformity. Policy outcomes have gone through dilution stages, initially due to the

rights surrendering the structure of the social contract, backed up by the asymmetric nature of information which creates principal agent problems from the starting point of personal ideologies across the sphere of the democratic process. Political parties seek election by adopting ideologies not necessarily totally representative of their own politicians' preferences. The politicians are often whipped to adopt legislation which has been influenced by other political parties and undemocratic and unrepresentative interest groups from civil, social and business spheres. A corruptible media is used to assimilate asymmetric information which is also provided by a rationally selfish bureaucracy. Due to the fact that the electorate is rationally ignorant, and that singular policy outcomes can become diluted of ideology, the electorate can be deemed rational to feel disillusioned. In the circumstances where politicians or parties adopt political ideologies on which pre-election policy promises are based, the electorate, when weighing up outcomes on a policy by policy basis, can be deemed rational when viewing politicians as charlatans.

References

Acemoglu, D. and Robinson, J., 2005. *Economic Origins of Dictatorship and Democracy*. New York: Cambridge University Press.

Barro, R. J., 1973. 'The Control of Politicians: An Economic Model'. *Public Choice*, 14, pp. 19-42.

Besley, T., 2004. 'Joseph Schumpeter Lecture: Paying Politicians: Theory and Evidence'. *Journal of the European Economic Association,* 2(2-3), pp.193-215.

Besley, T. and Burgess, R., 2002. 'The Political Economy of Government Responsiveness: Theory and Evidence from India'. *The Quarterly Journal of Economics*, 117(4), pp.1415-1451.

Besley, T. and Case, A., 1995. 'Does Electoral Accountability Affect Economic Policy Choices? Evidence from Gubernatorial Term Limits'. *The Quarterly Journal of Economics*, 110(3), pp.769-798.

Beyers, J., Eising, R. and Maloney, W., 2008. 'Researching Interest Group Politics in Europe and Elsewhere: Much We Study, Little We Know?'. *West European Politics*, 31(6), pp.1103 -1128.

Burson-Marsteller. 2005. 'The Definitive Guide to Lobbying the European Institutions Based on a Survey of the European Parliament, the Council and the European Commission'. Brussels: Burson-Marsteller.

[Online] Spring 2005. Available at: http://www.sabrizain.org/traffic/library/lobbyingeu.pdf [Accessed: December 6, 2010]

Chari, R., Hogan, J. and Murphy, G., 2010. *Regulating Lobbying: a global comparison*. Manchester: Manchester University Press.

Downs, A., 1957. 'An Economic Theory of Political Action in a Democracy'. *The Journal of Political Economy*, 65(2), pp.135-150.

Ferejohn, J., 1986. 'Incumbent Performance and Electoral Control'. *Public Choice*, 50(1), pp.5-25.

Ferejohn, J. and Shipan, C., 1990. 'Congressional Influence on Bureaucracy'. *Journal of Law, Economics,& Organization*, Special Issue: Papers from the Organization of Political Institutions Conference, April 1990 (6), pp.1-20.

Foucault, M., 1991. *Discipline and Punish – The Birth of the Prison*. London: Penguin Books.

Gentzkow, M. and Shapiro, J. M., 2008. 'Competition and Truth in the Market for News'. *Journal of Economic Perspectives*, 22(2), 133-154.

Hall, R. E. and Jones, C. I., 1999. 'Why Do Some Countries Produce So Much More Output Per Worker Than Others?'. *The Quarterly Journal of Economics*, 114(1), pp.83-116.

Heckman, J., Heinrich, C. and Smith, J., 1997. 'Assessing the Performance of Performance Standards in Public Bureaucracies'. *The American Economic Review*, 87(2), pp.389-395.

Hillman, A. L., 1982. 'Declining Industries and Political-Support: Protectionist Motives'. *The American Economic Review*, 72(5), pp.1180-1187.

Hillman, A. L., 2009. *Public Finance and Public Policy.* 2nd ed. Cambridge: Cambridge University Press.

Knack, S. and Keefer, P., 1995. 'Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures'. United States Agency For International Development (USAID).

[Online] Available at: http://pdf.usaid.gov/pdf_docs/PNABU168.pdf

Leeson, P. T., 2008. 'Media Freedom, Political Knowledge, and Participation'. *Journal of Economic Perspectives*, 22(2), pp.155–169.

Naurin, D., 2007. 'Backstage Behavior: Lobbyists in Public and Private Settings in Sweden and the European Union'. *Comparative Politics*, 39(2), pp.209-228.

Paxman, J., 2003. *The Political Animal*. London: Penguin.

Peltzman, S., 1976. 'Toward a More General Theory of Regulation'. *Journal of Law and Economics*, 19(2), pp.211-240.

Popper, K., 2002. *Conjectures and Refutations*. New York: Routledge.

Rawls, J., 2007. *Lectures on the History of Political Philosophy*. Paperback ed. Harvard: First Harvard University Press.

Rawnsley, A., 2001. *Servants of the People – The Inside Story of New Labour*. London: Penguin.

Saurugger, S., 2008. 'Interest Groups and Democracy in the European Union'. *West European Politics,* 31(6), pp.1274-1291.

Tirole, J., 1994. 'The Internal Organization of Government'. *Oxford Economic Papers*, 46, pp.1-29.

Zupan, M. A., 1990. 'The Last Period Problem in Politics: Do Congressional Representatives Not Subject to a Reelection Constraint Alter Their Voting Behavior?' *Public Choice*, 65(2), pp.167-180.

AGENCY THEORY AND REMUNERATION POLICY: PERFORMANCE INCENTIVES IN THE FINANCIAL SECTOR

Mark Sykes

Senior Sophister

During a time when bankers' bonuses inspire such outrage, this paper uses the agency approach to objectively evaluate the implications of remuneration policy on employee behaviour in the financial sector. Mark Sykes outlines the theory of the principle agent problem and its relevance to managerial capitalism and the separation of ownership and control in organisations, using contemporary international and Irish examples to illustrate the problem and finally discusses the regulatory response to date.

Introduction

> 'The bonus-hunting bankers...stand charged with destroying wealth on an epic scale. Foolish, greedy, irresponsible behaviour and excessive risk-taking led to massive losses and the crisis in the banking system which is now costing millions their jobs and many their homes. Why should such failure be rewarded?' (Cable, 2009).

The recent financial crisis has shone a tremendous light upon the theory and question of remuneration policy in the financial sector. The level and type of remuneration provided in the run up to the crisis, and in some instances following it, has led to emotional and populist responses from policy makers and commentators alike. The terms 'banker' and 'bonus' have become tainted, and when used together, resemble a popular byword for excess and irresponsible behaviour.

This essay shall use the tools of agency theory to analyse remuneration in the financial sector. Specifically, it will attempt to ascertain whether remuneration policy can lead to excessive risk taking, and if this is the case, whether it should be controlled by regulation. Firstly, it will discuss the concept of agency theory and the principle agent problem, and how it inevitably arises once ownership and control are separated in organisations. It shall then show how the problem can be partially solved by performance based incentives, such as bonuses and share options, and argue that excessive risk taking

`

can be an unintended consequence of these solutions. It will also contend that the principle-agent problem is particularly difficult to solve in the financial sector. The essay will conclude by arguing that wholesale changes to the remuneration policies of the financial sector are necessary but that these changes are already underway.

Agency Theory and Organisational Structure

> 'The directors of such (joint stock) companies, however, being the managers rather of other people's money than of their own, it cannot well be expected that they should watch over it with the same anxious vigilance with which the partners in a private company frequently watch over their own. Like the stewards of a rich man, they are apt to consider attention to small matters as not for their master's honour, and very easily give themselves a dispensation from having it. Negligence and profusion, therefore, must always prevail, more or less, in the management of the affairs of such a company' (Smith, 1776 cited in Jensen and Meckling, 1976, p.305).

The field of economics lays at its foundation an emphasis on incentives; it is thus no surprise that Adam Smith, one of the fathers of the field, foresaw the applicability of the principle-agent problem to organisational structure. Agency theory treats individuals as rational utility maximisers, and shows that, in the presence of asymmetric information a principle-agent problem will arise whereby the agent does not act in the best interests of the principle (Stiglitz, 2008: 637-643).

Example: A Theoretical Justification of Agency Costs.

Consider for example a situation whereby an employer (C) engages an employee (L) to produce output (Q) in exchange for a wage (W). The employee's output is a function of effort (e), skill (s) and luck (l). Assume the worker dislikes work. Initially the employer can only observe output.

$$U(C) = f(Q,-W)$$
$$U(L) = f(-e,W)$$
$$Q = f(e,s,l)$$

For a given wage, the employee will minimise effort to maximise their utility. The employer, in seeking to maximise their own utility can then either incur supervision costs to observe and encourage effort, or can try to incentivise effort via making W a function of Q (carrot and stick). However, these actions will never be complete solutions; effort may not be perfectly observed (or the costs could be prohibitive) and given that output is a function of factors outside of the employee's

control, it is an imperfect basis to remunerate on (i.e. rewarding a lazy but lucky worker over a hardworking but unlucky worker).

Thus, the principle-agent problem is one of asymmetric information and divergent objectives leading to moral hazard. The pay-off to the agent will differ to that of the principle, the agent will not in general take the action which the principle would like him to take, or that they would contract for in the presence of perfect information (Stiglitz, 2008, p.638). The employee may not adjust his effort as the situation requires or may engage in too much or too little risk taking (Stiglitz, 2008, p.638). The solution is to look for the compensation scheme which maximises the expected utility of the principle, given that the agent will undertake the actions which maximise his own expected utility and that the agent must be willing to accept the contract (Stiglitz, 2008, p.639). We can thus quantify the totality of agency costs as consisting of: monitoring expenditures by the principal, bonding expenditures by the agent and a residual loss (resulting from the agent making decisions which do not maximise the welfare of the principal) (Jensen and Meckling, 1976, p.311).

Given these agency costs, one might reasonably wonder, as Smith did, whether a publicly owned firm can ever compete with owner run organisations. As Coase (1937) and Fama and Jensen (1983b, p.302) argued, the firm is simply a set of contracts among factors of production. In competitive markets, inefficient functional forms will not survive. Thus, the ex-post presence of publicly owned organisations in an industry is a sufficient condition for them to be the most efficient functional form. Specifically, Fama (1980, p.291) and Markowitz (1952) argued that the separation of ownership and control allows investors to diversify their portfolios, hence reducing risk in line with modern portfolio theory and enabling the use of specialised managers (Fama and Jensen, 1983a, p.330). Fama also argued that competitive managerial labour markets, along with internal and external monitoring devices would 'evolve to stimulate the on-going efficiency of the corporate firm, and with the market for outside takeovers providing discipline of last resort' (Fama, 1980, p.295).

As such, we can view the choice of organisational form as a cost benefit analysis. Agency costs are the biggest negative relating to the separation of ownership and control, but these can be somewhat mitigated via rational managerial labour markets which incorporate expected underperformance on an ex-ante basis (Fama, 1980, p.296). On the benefits side of the ledger, publicly owned corporations provide important diversification and specialisation, and facilitate the access of additional external capital[1].

Remuneration and Agency Costs in the Financial Sector

It is important to predicate our discussion of remuneration policy in the financial sector by highlighting relevant differences between many financial firms and non-financial firms. Firstly,

[1] Thus, for instance, we would expect industries which benefit greatly from economies of scale (telecommunications, energy) to be organised as joint stock companies. Conversely, we expect professional service firms, with high risks from agency costs and limited capital requirements to be organised as partnerships.

although all corporations in which ownership and control are separated face moral hazard arising from increased agency costs, financial firms can also face moral hazard via their being 'too big to fail'. That is, in many cases, the negative externalities associated with their failure outweigh the costs to the exchequer of socialisation of losses via a public 'bail-out'. This is essentially an extension of the moral hazard created by having limited liability in the first place, and can incentivise additional risk taking on behalf of debt and other providers of capital as their downside risk is limited. As such, the perceived presence of moral hazard and excessive risk taking ex-post in a financial firm does not prove that remuneration policies were to blame.

Financial firms are generally more highly levered and thus their returns more volatile than non-financial firms. As such, the effects of remuneration policies which encourage risk taking are likely to be more severe. Furthermore, in many areas of the financial sector, the full effects on profitability of an employee's actions are not known within a single financial year and thus remunerating on this basis is inefficient. For example, profits accruing from a loan or insurance policy will not be known with certainty until it is fully repaid in the case of a loan or elapses in the case of an insurance policy (unless they are securitised and sold, in which case the information problem falls to the purchaser of the security). In fact, there will always be an information asymmetry between management and employees in these instances, that is, given that the financial product was sold under conditions of uncertainty, their ex-post profitability does not indicate whether the employee adhered to the appropriate lending/credit standards when initiating the contract.

It is also useful to note, given the diversified nature of many financial institutions, that the ability of individual employees to take risks, even if their remuneration policies implicitly incentivise them to, varies widely between business units and is often constrained by internal controls. That said, it remains plausible that senior management, whose actions, given their importance, could have a meaningful influence on share price value, could be capable of taking excessive risks were their remuneration policy incentivising such activities.

Lastly, it is worth acknowledging the argument that there exists a 'bonus culture' in financial institutions. If true, in isolation, this is not a negative attribute which would engender risk taking. As explained in 'A Theoretical Justification of Agency Costs', above, relating remuneration to output or performance is a useful mitigant of agency costs and serves to align shareholder and employee's utility functions. However, these policies could plausibly lead to risk taking by management in excess of that demanded by investors, an issue which we will shortly discuss in more detail.

The Global and Local Financial Crisis and Remuneration Policy

As alluded to already, notwithstanding that created by remuneration policy, there exists a significant amount of moral hazard in the financial sector, which plausibly encourages risk taking in excess of that which maximises shareholder value. As such, we cannot conclusively attribute any excessive risk taking in the run up to the financial crisis to misaligned incentives driven by

remuneration policy. That said, we can still gain meaningful insights by examining what management should do in theory - maximise the present discounted value of the firm and as such the share price Figure 1 - and what their remuneration policy incentivised them to do.

Figure 1

$$P_t = \sum_{k=0}^{\infty} \left(\frac{1}{1+r} \right)^{k+1} E_t d_{t+k}$$

(Sheffrin, 1996, p.122)

Using the examples of senior management remuneration in Bear Stearns, Lehman Brothers, Allied Irish Banks (AIB) and Anglo Irish Bank ('Anglo'), four institutions which either failed or required state support during the recent financial crisis, shows that senior management remuneration was primarily composed of base salary, performance related bonuses and share options (Bebchuk, Cohen and Spamann, 2009) (Anglo, 2007, pp.124-130) (AIB, 2007, pp.160-165).

In all cases, though the proportion varied widely, *base salaries* were the smallest portion of executive pay and in the case of Lehman Brothers and Bear Stearns was almost negligible in terms of total compensation. Taken on its own, this is a positive in terms of principle-agent problem mitigation, returning to 'A Theoretical Justification of Agency Costs' above; fixed salaries provide limited incentive for effort[2] on the part of management.

In all cases, *performance bonuses* were a multiple of base salary and on an ex-post basis were the most valuable form of compensation. They were paid in cash and on the basis of annual performance[3], including on metrics such as share price and earnings per share. For example, the average Bear Stearns senior executive received an annual cash bonus of just over $10m between 2000 and 2006 with the equivalent for Lehman being $5m (Bebchuk, Cohen and Spamann, 2009, p.12). Executive directors in AIB received an average bonus of €1m in 2006 and €700k in 2007 (AIB, 2007, pp.161-162). In Anglo the equivalent figures were just under €1m[4] in both 2006 and 2007 (Anglo, 2007, pp.126-127). As discussed previously, remuneration based on annual results may be inefficient for a financial firm, given the multi-annual implications of decisions made during the year. Similarly, the fact that these bonuses were in cash meant that employees futures were not tied to the long term performance of the firm and using the agency approach as always, incentivised short term profits at the expense of long term performance, contrary to the shareholders objective function given in

[2] However, the 'stick' side of incentivisation would remain, with unemployment plausibly being a risk of underperformance. Similarly, as with the rest of the paper, we take the narrow agency view, ignoring potential non-monetary motivations for work: the importance of self-esteem, peer approval, etc.
[3] Not explicitly stated that bonus was in the form of cash for AIB.
[4] Figures exclude John Rowan who retired on 31 Dec 2005, and received a €1.1m retirement bonus (Anglo, 2007, p.127).

equation Figure 1. That said, this misalignment of incentives is somewhat mitigated by the presence of share price performance in the evaluation criteria, which in the presence of rational markets should evaluate company performance based upon the same criteria in Figure 1, i.e. should not reward short term profit taking at the expense of long term growth.

At first glance, *share options* appear the form of remuneration which would best align management incentives to the objectives of shareholders, given that you are broadly connecting their pay-offs. Furthermore, the use of options rather than the issuing of shares directly is somewhat equivalent to giving management a levered position in the company, giving them far larger incentive to maximise returns relative to the same number of shares. Given this, it is understandable that these options, although now worthless in most cases, were a large part of compensation in each company. On December 31st 2007 each of the Executive Directors of AIB held in excess of 90,000 share options, with their equivalents in Anglo all holding in more than 500,000 on September 30th of the same year (AIB, 2007, p.164) (Anglo, 2007, p.129). In the cases of Lehman Brothers and Bear Stearns many awarded options were cashed out over the period 2000-2007 but in 2008 the top 5 Executive Directors still held 18.8m and 6.8m shares respectively (Bebchuk, Cohen and Spamann, 2009, p.17).

That said, it is important to note the obvious intuitive fact that the value of a share option is not solely a function of the underlying share price. It is also a function of the volatility of the share price (Black and Scholes, 1973). That is, ceteris paribus, the value of an option generally increases as the 'riskiness' or expected variation in the stock price increases. Secondly, the very fact that downside risk is limited in relation to a call option generates incentives for risk taking as the pay-off from success in an endeavour is far in excess of the potential losses from failure.

Thus, we can summarise that base salaries, because of their small proportion of overall remuneration, had only a limited impact on management behaviour. Performance bonuses and share option schemes certainly provided incentives for management to be short-sighted and take risks in excess of that which maximised the net present value of the firm, as per the shareholder view of the firm. However, the extent to which this behaviour actually occurred is difficult to quantify and disentangle from the wider causes of the financial crisis.

Lastly in relation to this portion of the essay it is useful to make a further note, that although remuneration policies on both sides of the Atlantic have been shown to encourage excessive risk taking and that large amounts of remuneration were cashed out in the run up to the financial crisis, Chief Executives still retained a large amount of 'skin in the game' and made significant personal losses when their respective firms share prices collapsed. In September 2008, when Lehman Brothers collapsed, their CEO, Richard Fuld, held almost 11m of their worthless shares, a loss of almost $1bn from peak valuation (Bebchuk, Cohen and Spamann, 2009, p.17; Yahoo!Finance). Similarly, as Figure 1 estimates, Irish Chief Executives made large losses from their institutions failures.

Although the losses made by the AIB CEO were smaller, they in fact represented a larger percentage of overall awarded compensation.

Figure 2

	AIB CEO	Anglo CEO
Shares Held	255,845	510,899
Peak Share Price	€21.57	€17.31
Current/Final Share Price	€0.27	€0.00
Loss from Peak	€5,499,499	€8,843,662
Options Held	120,000	1,201,834
Weighted Average Exercise Price	€13.78	€7.43
Loss From Peak	€934,800	€11,874,120
Accumuated Loss	€6,384,299	€20,717,782

(Anglo Annual Report 2007; AIB Annual Report 2007; Yahoo! Finance)
Note: AIB CEO also conditionally held 251,163 shares which would have been awarded between 2008 and 2010 based on certain performance targets been met. Those were not met.

Moving Forward: Incentive Alignment and Regulatory Response

> '[There] is agreement by supervisors and regulatory bodies…that the inappropriate remuneration structures of some financial institutions have been a contributory factor (to the financial crisis). Remuneration policies which give incentives to take risks that exceed the general level of risk tolerated by the institution can undermine sound and effective risk management and exacerbate excessive risk-taking behaviour' (CEBS, 2010, p.6).

As should be evident at this stage, agency costs are a very real and important issue, particularly in the financial sector when they are added to the additional moral hazard created by the issue of 'too big to fail'. The only perfect solution is to have the owners run the business, which is not feasible or desirable given the many benefits of the separation of ownership and control. In fact, even in this instance, agency issues would occur in relation to debt holder's objectives (Jensen and Meckling, 1976, p.350). As such, any response will only be a partial solution, but it is certainly

plausible that better efforts can be made to align incentives than those described previously in relation to the financial crisis.

The CEBS (2010) has made many wide ranging recommendations in relation to remuneration policy, its composition and oversight in the financial sector. For instance, it has emphasised the need for remuneration to be consistent with effective risk management (CEBS, 2010, p.48), has set tough guidelines for the composition of remuneration committees (CEBS, 2010, p.32), has proposed the outlaw of variable remuneration not linked to performance such as 'golden parachutes' and 'minimum bonuses' (CEBS, 2010, p.41-42) and has argued for increased deferral of remuneration and the use of non-cash instruments (CEBS, 2010, p.18). The Irish Financial Regulator has echoed these sentiments, particularly in relation to risk management (Central Bank of Ireland, 2010).

These responses, in general, appear reasonable, and should more closely align management and shareholder objectives, somewhat mitigating management's incentive to be short sighted. Furthermore, they are in line with those being taken by major financial institutions following shareholder consultation. For instance, Credit Suisse (2011) has announced remuneration structures in excess of the minimum stringency set out by the CEBS. Variable compensation will vest over 4 years, be linked to the firm's return on equity and be linked to the firm's share price with no element of leverage (Credit Suisse, 2011). Similarly, Barclays has pioneered the use of contingent convertible bonds or 'CoCos' as part of senior executive variable remuneration. These awards will convert to equity and thus be capable of taking losses should the bank's core capital fall below 7%[5] (Barclays, 2011, p.19).

It is important to distinguish between reforms which relate to how remuneration is paid versus the level of remuneration itself. The discussion thus far has seeked to explain how moral hazard may arise in a given remuneration structure and what methods could be used to mitigate it. It has made no value judgment as to the determination of the level of remuneration. While it may be politically popular to impose pay restrictions or ceilings on the level of remuneration in financial firms, particularly those whose failures have let to public expenses, these would likely lead to efficiency losses. The need for regulation to reduce the probability of future socialisation of losses in the financial sector is inarguable, but this does not imply the state should have any say in the battle between capital and talent[6] in an individual institution.

Conclusion

While impossible to prove outright, there appears no doubt that the incentives created by remuneration policy in the financial sector contributed to the financial crisis which is still ongoing in

[5] For 2010 bonuses these awards were synthetic, pending regulatory approval of the use of traded 'CoCos' (Jones, 2011).

[6] As discussed generally in Martin and Moldoveanu (2003). Essentially this is an issue between the relative returns to labour and capital in financial institutions. Reducing remuneration levels would essentially be a transfer of wealth from labour to capital (ignoring the potential for deadweight loss).

Ireland. Using the tools of agency theory we have set up the problem, rationalised the failure and applauded the majority of the regulatory response.

However, we must caution that regulatory interventions are prone to unforeseeable consequences. The incentives created by remuneration policy will never perfectly align employee's incentives with the objectives of shareholders. Similarly, we must caution against the maximisation of shareholder value as the only objective for financial firms. Given the large positive and negative externalities they are capable of generating, a wider stakeholder view may be more appropriate. The narrow shareholder theory ignores the importance of the credit and lending channels to the functioning of the economy and equally disregards the potential socialisation of losses if a systematically important private institution becomes illiquid or insolvent.

Finally, it is worth noting the imperfections of agency theory itself. It leaves no room for non-financial motivations and assumes perfect rationality. Nevertheless, this author feels that the approach has merit, and although its assumptions are incorrect it has useful predictive power, which is all we require of a model. If economics has succeeded in anything, it has shown that incentives matter - and the financial sector is no exception.

References

AIB (Allied Irish Banks), 2008. 'Annual Report 2007'. *AIB*, [Online]. Available at: http://www.aib.ie/servlet/ContentServer?pagename=AIB_Investor_Relations/AIB_Download/aib_d_download&c=AIB_Download&cid=1205398774668&channel=IRSI

Anglo (Anglo Irish Bank), 2008. 'Annual Report 2007'. *Anglo Irish Bank*, [Online]. Available at: http://www.angloirishbank.com/About-Us/Reports/Previous_Reports/Annual_Report_2006_.pdf

Barclays PLC, 2011. '2010 Results Presentation, 15 February 2011'. *Barclays* [Online] 15 February. Available at: http://group.barclays.com/Investor-Relations/Financial-results-and-publications/Results-announcements

Bebchuk, L. A., Cohen A. and Spamann, H., 2009. 'The Wages of Failure: Executive Compensation at Bear Stearns and Lehman 2000-2008'. Harvard Law School, *Discussion Paper No. 657*.

Black, F., and Scholes, M., 1973. 'The Pricing of Options and Corporate Liabilities'. *The Journal of Political Economy*, 81(3), pp.637-654.

Cable, V., 2009. 'Bring Back the Guillotine...for Bankers'. *The Daily Mail* [Online] 9 February. Available at: http://www.dailymail.co.uk/debate/article-1138673/VINCE-CABLE-Bring-guillotine--bankers.html [Accessed: January 22, 2011]

Central Bank of Ireland, 2010. 'Review of Remuneration Policies and Practices in Irish Retail Banks and Building Societies'. *Central Bank of Ireland*, [Online] 1 December. Available at: http://www.financialregulator.ie/press-area/press-releases/Pages/ReviewofRemunerationPoliciesand Practices.aspx. [Accessed: January 22, 2011]

Coase, Ronald H, 1937. 'The Nature of the Firm'. *Economica*, 4(16), pp.386-405.

CEBS (Committee of European Banking Supervisors), 2010. 'Guidelines on Remuneration Policies and Practices'. *CEBS*, [Online]10 December. Available at: http://www.eba.europa.eu/cebs/media/ Publications/Standards%20and%20Guidelines/2010/Remuneration/Guidelines.pdf [Accessed: January 22, 2011]

Credit Suisse, 2011. 'Credit Suisse Announces Its Compensation Structure'. Credit Suisse [Online] 10 January. Available at:https://www.credit-suisse.com/news/en/media_release.jsp?ns=41668 [Accessed: January 22, 2011]

Fama, E. F., 1980. 'Agency Problems and the Theory of the Firm'. *Journal of Political Economy* 88(2), pp.288-307.

Fama, E. F. and Jensen, M. C., 1983a. 'Agency Problems and Residual Claims'. *The Journal of Law and Economics,* 26(2), pp.327-349.

Fama, E. F. and Jensen, M. C., 1983b. 'Separation of Ownership and Control'. *The Journal of Law and Economics*, 26(2), pp.301-325.

Jensen, M. C. and Meckling, W.H., 1976. 'Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure'. *Journal of Financial Economics*, 3(4), pp.305-360.

Jones, A., 2011. 'Barclays Beats Forecasts with 32% Rise'. *The Financial Times*, [Online] 15 February.Availableathttp://www.ft.com/cms/s/0/979c8886-38d4-11e0-b0f6-00144feabdc0.html#axzz1E1EfWcap

Markowitz, H. M., 1952. 'Portfolio Selection'. *Journal of Finance*, 7(1), pp.77-91.

Martin, R. L. and Moldoveanu, M. C., 2003. 'Capital Versus Talent: The Battle That's Reshaping Business'. *Harvard Business* Review, 81(7), pp.36-41.

Sheffrin, S., 1996. *Rational Expectations*. Cambridge: Cambridge University Press.

Stiglitz, Joseph E., 2008. 'Principle and Agent (ii)'. *The New Palgrave Dictionary of Economics*. 2nd ed. pp.637-643.

Yahoo! Finance, 2011. *'*Business Finance, Stock Market, Quotes, News*'*. *Yahoo Finance*, [Online] 22 January. Available at: http://finance.yahoo.com/.

AN ECONOMETRIC ANALYSIS OF THE RELATIONSHIP BETWEEN FEMALE PARTICIPATION IN THIRD LEVEL EDUCATION AND THE BIRTH RATE

Anne Talbot

Senior Sophister

Studies have shown that women who have a higher education tend to choose to have less children. Anne Talbot tests this theory with a well-considered econometric analysis of the relationship between female participation in third level education and the birth rate, using other interesting variables such ashealth and unemployment.

Introduction

High educational attainment levels are an imperative indicator of the standards of living of a country, but the further development of a country relies on its future generations to improve these standards. In a famous quote, US President Herbert Hoover stated that "children are our most valuable natural resource".

In this project I have completed a regression analysis of the relationship between female participation in third level education and the birth rate. Other independent variables include mothers' average age at first birth, infant mortality rate, the percentage of GDP spent on health by the government, legality of abortion, unemployment rate of women and the marriage rate. This analysis is based on panel data over the time period 2000-2007 for the EU-15 countries[1].

Women's role in society is no longer to be the bread maker, but often the bread winner. Births outside of marriage are more common than ever before, people's attitudes towards women's roles in society, towards education and towards having children have all changed.

Observation and analysis shows that the birth rate has declined drastically in the recent past; in Ireland alone the birth rate has shrunk from 4.1 in 1970 to 1.95 in 2000 (Central Statistics Office). Coinciding with this decrease has been an increase in educational attainments, more notably those of women. As education participation at third and higher levels continues to soar, what effect does this impose on future generations, or lack thereof? Intuition suggests that there is a direct and indirect relationship present. By constructing an econometric analysis of birth rates I aim to identify, analyse and describe this relationship.

Literature Review

There are few studies analysing the effect of education on the child birth rate, and I was unable to find specific studies for third level education. There are however countless studies on the child birth rate and the influences on it. Studies used cross-sectional data, time series data, and panel data. In almost all studies of the child birth rate, it has been proven that a higher level of schooling leads to a lower birth rate and many examples are shown by contrasting the Western World with the Developing World where education is not as widely available. I am more interested in the effects of education on the birth rate in EU-15 countries. According to a study by the National Centre for Health Statistics, Centres for Disease Control and Prevention in America (NCHS, 1997):

"A woman's educational level is the best predictor of how many children she will have". Many studies are done in the United States, and a paper by Mathews and Ventura's (1994) implies that racial background is an extremely influencing factor on the number of children women had. However, in the US, where third level education is estimated to cost $22,000 per child per year[4], many Hispanic Americans and African Americans (the groups with the highest birth rates) were in lower paid jobs than whites (the group with the lowest birth rate), which would imply that fewer Hispanics attended college due to the higher costs and so it is hard to distinguish results.

Studies are also often done at a national level and aim to analyse the effect of migration on the birth rate. Countries such as France and Ireland have seen huge net immigration figures in the early 2000s, whereas many British have moved permanently to Spain and other southern countries. In 2005 alone, Eurostat reported figures of net international immigration to the EU of close to 1.8 million. In Ireland the immigration of women of child bearing age in the boom years lead to an increase in the birth rate.

I aim to show whether tertiary education has or does not have an impact on the birth rate that is worth recognising or disregarding in its own right.

Empirical Approach

$$Birth = \beta_0 + \beta_1 educ + \beta_2 mar + \beta_3 age + \beta_4 health + \beta_5 unemp + \beta_6 infmort + D_1 abort + u$$

Dependent Variable:

Y *(Birth)*: The dependent variable Y represents the live birth rate represented by the total live births in a country for a given year. *A live birth* is defined as "the complete expulsion or extraction from its mother of a product of conception, irrespective of the duration of pregnancy, which after such separation breathes or shows any other evidence of life such as beating of the heart, pulsation of the umbilical cord or definite movement of voluntary muscles, whether or not the umbilical cord has been cut or the placenta is attached" (UNECE). The data for this variable was obtained from the UNECE

(United Nations Economic Commission of Europe) and is assumed to be adequate and accurate for the sample years.

Independent Variables X_i:

X_1 *(educ)*: Third level education of women. The first explanatory variable is the percentage of women in third level education, expressed as a percentage of total persons in third level education. This data was obtained from the UNECE Statistical Division Database, so is assumed to be adequate and accurate. I would expect there to be a negative relationship between my dependent variable, the birth rate, and my first independent variable.

X_2 *(mar)*: Marriage. The second independent variable is the number of new marriages in a country per year. Previous studies have shown that women who are married are more likely to have children than those who aren't and so it is an important variable to include. I would expect a strong positive relationship between the number of births and number of marriages. This data was obtained from UNECE Statistical Division Database and is assumed to be adequate and accurate.
A plot of birth, education and marriage rate showed similar trends, implying that they change with each other.

X_3 *(age)*: Age. The third independent variable is the mean age of women at the time of birth of their first child. I expect that age will have a negative effect on the child birth rate since older women are less likely to have multiple children ,due to the natural cycle of life. This data was obtained from the UNECE Statistical Division Database and the CSO, thus is assumed to be adequate and accurate.

X_4 *(health)*: Health. The fourth independent variable represents the percentage of GDP spent on the healthcare system by the government. Since it is reasonable to assume that more money spent on the healthcare system improves the system, I expect that this variable will positively affect my dependent variable. This data was obtained from the Eurostat statistical database and is assumed to be adequate and accurate.

X_5 *(unemp)*: Unemployment. The fifth independent variable represents the number of women unemployed as a percentage of the labour force. The unemployment rate is defined as "the share (in per cent) of the unemployed in the labour force" (UNECE). The data I have collected is from UNECE and is assumed accurate and adequate. I expect unemployment rate to positively affect the birth rate.

X_6 *(infmort)*: Infant mortality Rate. The sixth independent variable represents the infant mortality rate, defined as "the number of deaths of infants under one year of age per 1000 live births in a given year" (UNECE). I expect a high infant mortality rate to negatively affect the birth rate. This data was obtained from the Eurostat statistical database and is assumed to be adequate and accurate. A plot of the Infant mortality rate showed it to be normally distributed.

X_7 *(abort)*: Abortion. I included the seventh variable as a dummy variable since abortion is not legal in all EU-15 countries. I assigned a value of one if abortion was legal in the country for the

given year and a value of zero if abortion was illegal in the country for the given year. I expect the legality of abortion to decrease the birth rate, this is reasonable to assume based on the purpose of abortion. Table one provides the breakdown of abortion as a dummy variable and the number of observations where it is legal (1) and illegal (0).

Table 1

	Overall		Between		Within
abort	Freq	Percent	Freq	Percent	Percent
0	19	15.83	3	20	79.17
1	101	84.17	13	86.67	97.12
Total	120	100	16	106.67	93.75

U: Error term. The Error term represents the unobservable or uncountable variables such as natural ability to have children, or demands for contraception. This must be included as my econometric analysis, based on my independent variables, may be incomplete. This term is assumed to have a mean of zero and to be normally distributed.

Approach

"Panel data (also known as longitudinal or cross-sectional time-series data) is a dataset in which the behaviour of entities are observed across time" (Data Statistical Services).

This project's entity is countries. With panel data there are various techniques used to test the data. I must use the Hausman test to determine whether fixed or random effects describe our data more appropriately. If the data is found to be fixed then we must test for heteroscedasticity and the least square dummy variable model (LSDV) provides a good means to understand fixed effects. If the data is found to be random then the "variation across entities is assumed to be random and uncorrelated with the independent variables included in the model" (Data Statistical Services). We must use the Breusch and Pagan Lagrangian multiplier test. The Pasaran CD test is used to test whether the residuals are correlated across entities since cross-sectional dependence can lead to bias in tests results. Autocorrelation is hard to test for with panel data and so I will not test for autocorrelation in my model.

Description of data set

I collected my data from three main sources. The first being UNECE, the second being Eurostat and the third being CSO Ireland. These data sources provided the most comprehensive and up to date data for my desired countries. I decided to base my model on the EU-15 countries since I feel that they are comparable in terms of living standards, employment levels and GDP per capita

levels. They also have free or relatively cheap third level education fees, along with a fully affiliated ECTS university credit system. My data is base on the time period 2000-2007. This is due to the amount of comparable data available for these years. Initially I had planned to have a wider time period however data availability restrictions impeded this.

I chose my variables carefully and based on resources available. I included the marriage rate as marriage is often associated with families and children. I included the average age of a woman at first birth because a woman's ability to have children is restricted by age so the older a woman is the more likely it is she will have fewer children. I decided to include the percentage of GDP spent on health by the government as in all of my chosen countries, child birth and post natal care are free and so the quality of such service would be reflected in the expenditure, which would impact on the birth rate. I included the unemployment rate of women because women who are long-term employed are often career driven and having children may not be a priority. I included the infant mortality rate since a high mortality rate could be a deterrent to having children. I decided to include abortion since, if abortion is legal and women chose to have abortions, the birth rate decreases.

Table 2 contains the variables, their means, standard deviations, minimum and maximum values. Bir, ed, marr and ump are the logs of birth, educ, mar and unemp respectively. I took the logs of these variables because they varied greatly between countries and over time.

Table 2

Variable	Obs	Mean	Std. Dev.	Min	Max
year	120	2003.5	2.300895	2000	2007
country	120	8	4.338609	1	15
birth	120	274385.2	276267	5303	830288
educ	91	498664.5	446513.2	91158	1352421
mar	108	90533.03	95781.9	1472	305637
age	89	28.26292	0.864237	26.4	30
health	118	8.918644	1.213454	5.8	11.1
unemp	120	7.145	2.893367	1.9	18.4
abort	120	0.841667	0.366584	0	1
infmort	117	4.122222	0.790473	1.8	6.2
bir	120	11.88175	1.293377	8.576028	13.62953
ed	91	12.70731	0.908923	11.42035	14.11741
marr	108	10.73635	1.341818	7.294377	12.63015
ump	120	1.87893	.435091	.6418539	2.912351

Empirical Results

After taking the log of birth, education, unemployment and marriage (which was conducted to improve the functional form) I ran the general multiple regression model in Stata. Since my model consists of panel data I had to ascertain whether I had fixed or random effects in my model; this was done by using a Hausman test and my results showed I had fixed effects. Table 3 contains the results of my fixed effects model.

Table 3

Log Birth			
Log Education	-4.81e-07**	Observations	68
	(2.33-e-07)		
Log Marriage	0.0321***	Adjusted R-squared	0.62
	(0.071)		
Age	-0.001	Sigma_u	0.915
	(0.023)		
Health	0.0226	Sigma_e	0.033
	(0.014)		
Log Unemployment	-0.049**	RHO (fraction of	0.999
	(0.02)	variance due to u_i)	
Abortion	-0.025		
	(0.032)		
Infant Mortality	-0.014		
	(0.015)		
Constant	8.59***		
	(0.915)		
F test (all u_i=0)	F(12,48) = 74.54	Prob>F=0.0000	

Robust standard errors in parenthesis. *significant at 10%; **significant at 5%; ***significant at 1%.

We can see that education, marriage, unemployment and the intercept are significant at the 5% significance level while marriage and the intercept are significant at the 1% significance level. At the 15% significance level health would be significant also. Contrasting to this is age, which has a p-value of almost one and so is highly insignificant in my model. Looking at the signs of the coefficients, we see that education, age, unemployment, abortion and infant mortality negatively affect the birth rate, which concurs with my expectations. The relatively large coefficient of marriage implies that marriage has the greatest influence on the birth rate.

The R-squared value is a goodness of fit test for the model. The overall R-squared value of 0.62 implies that 62% of the variation in the birth rate is explained by this model.

The F-test is "used to test the overall significance of the regression" (Data Statistical Services). Since the p-value for the F-test is <0.05, I can reject the null hypothesis that all of the Beta coefficients are zero and conclude that at least one of my independent variables has a significant effect on the birth rate.

The rho value of 0.99866553, also called the intra-class correlation, represents the "variance not explained by differences across entities" (Data Statistical Services).

Table 4 shows the contrasting coefficients when different techniques are used. Here Dummy represents the regression line where we are controlling for the unobserved heterogeneity. Fixed and Dummy are very similar, reinforcing the decision to use the fixed effects model.

Table 4

Variable	Fixed	Random	Dummy
Education	-0.127	0.255	-0.128
Marriage	0.332	0.578	0.341
Age	0.0004	-0.0191	-0.005
Health	0.0172	0.024	0.024
Employment	-0.009	-0.006	
Abortion	-0.027	-0.006	-0.025
Infant Mortality	-0.0116	-0.01	-0.152
Unemployment			-0.055
Country 2			.42120689
Country 3			-.13495557
Country 4			-.10887974
Country 5			1.9388043
Country 6			1.7276863
Country 7			.27509024
Country 8			-.11282649
Country 9			0
Country 10			0
Country 11			.80203807
Country 12			.26888844
Country 13			1.4287954
Country 14			.29237861
Country 15			1.8710752
Constant	9.8470606	2.8464006	9.3335514

Since my model is a fixed effects model I can ignore the Breusch and Pagan Lagrangian multiplier test. The Pasaran CD (cross-sectional dependence) test is used to test whether the residuals are correlated across countries. After running the test we see that there is no cross-sectional dependence. To test for heterscedasticity, we use the function *xtest3*. We obtain a p-value of 0.000 and since the null hypothesis states we have homoscedasticity we reject the null and conclude that our model is heterscedastic, which must be controlled for using the 'robust' option. Tables 2, 3 and 4 used the robust option.

Extensions

Since my model does not account for all variables that affect the birth rate, such as contraception use, extensions to my model would be possible. The data available was limited and I was unable to find data on a large time scale. Also, contraception data was not available to me. Such data may be available to certain organisations, or may be available in the future and so further analysis of my topic would be appropriate. My model looks at the EU-15 countries. Perhaps extensions to this set could be to include all EU countries, the US or Japan to account for a cultural impact on the birth rate.

Conclusion

With regards to the original data set, the main conclusion is that the most important factor affecting the birth rate is the marriage rate. Due to the reasoning for marriage and the social ideas of marriage, this makes sound and logical sense. Other studies found marriage to have a high impact on the birth rate also, both in developed and less developed countries.

Looking in particular at education, which was my variable of interest, it does, as expected, have a negative effect on the birth rate. The main reasons for this negative effect is due to women who complete third level education having more awareness of contraception and how to use it properly – this may not come from the education received in lectures, but as a result of college life and the availability of information on such matters. A direct and indirect relationship exists.

In various other papers on birth rates, it is often concluded that, while third level education does not always reduce the first-birth rates, "the effect on higher order birth rates was found to be significant" (Demographic-Research), implying that education does reduce the birth rate, but with regards to the number of children women have, not the number of women who have children.

The purpose of this project was to analyse the effect female participation had on the child birth rate and to provide suggestions for this effect, and so, I concur with John Dewey, who famously said "Education is a social process. Education is growth. Education is, not a preparation for life; education is life itself". Not only do we obtain a third level education, we learn valuable life lessons and general knowledge that improve our lives and it is the collective result of female participation in third level education that negatively impacts on the birth rate.

References

1: The EU-15 countries are Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxemburg, Netherlands, Portugal, Spain, Sweden, United Kingdom.

2: $22,000 based on the average fees for one academic year in a college in the USA (America.gov).

America.gov, [Online]. Available at: www.america.gov

Baum, C. F., 2006. An Introduction to Modern Econometrics Using Stata. 1st ed.: Stata Press.

Data and Statistical Services, Princeton University, [Online]. Available at: http://dss.princeton.edu

Demographic-Research, [Online]. Available at: www.demographic-research.org/Volumes/Vol3/3/html/4.html

Gujarati, D., 2004. *Basic Econometrics*. 5th ed.: McGraw-Hill.

Mathews, T.J. and Ventura S. J., 1994. 'Birth and Fertility Rates by Educational Attainment: United States, 1994'. National Centre for Health Statistics, Monthly Vital Statistics Report, 45(10).

UNECE (United Nations Economic Commission for Europe), [Online]. Available at: http://w3.unece.org/pxweb/Dialog/Concepts.pdf

Wooldridge, J. M.,2009. *Introductory Econometrics, A Modern Approach*. 4th ed.: Thomson.

Data Sources

America.gov, [Online]. Available at: www.america.gov

Central Statistics Office (CSO), [Online]. Available at: www.cso.ie

Data and Statistical Services, Princeton University,[Online]. Available at: http://dss.princeton.edu

Demographic-Research, [Online]. Available at: www.demographic-research.org/Volumes/Vol3/3/html/4.html

Europa, [Online]. Available at: www.europa.eu

Eurostat, [Online]. Available at: www.epp.eurostat.ec.europa.eu

Stata, [Online]. Available at: http://www.stata.com/meeting/1nasug/drukker.pdf

UNECE (United Nations Economic Commission for Europe), [Online]. Available at: www.unece.org

WHY DO WOMEN END UP IN 'FEMALE' JOBS?

Josie O'Reilly

Senior Sophister

Josie O'Reilly seeks to address the fundamental differences that arise between male and female unemployment. She scrutinises the institutional influence of different types of welfare states, coming to some surprising conclusions. She also examines how socio-cultural norms – often ignored by economists – affect our view of women and the role that they play in the economy.

Introduction

If you slipped on the icy pavement today and had to go to hospital you would expect a female nurse to tend to you while you waited to see a doctor. Similarly, if your kitchen appliances stopped working or you suddenly found yourself in the dark, you would call an electrician and expect a man to come and fix it. In our industrially advanced societies these expectations usually prove to be accurate, yet in the last century women's place in society has been revolutionized. As they have earned the right to vote, to study, to work, women's outlook on life has evolved; no longer the bland wallflowers defined by their husbands and families, they are now active agents with wider horizons, who make decisions independently (Goldin, 2006) and are an important part of the labour force (Muhlau, 2010a). So why are women still confined to 'female' jobs? This essay will begin by defining what we mean by 'female' jobs, it will then review the economic theories and non-economic theories as to why women end up in them, and finally it will conclude by briefly examining the societal effects of occupational segregation.

What Makes a Job 'Female'?

The different theories to explain occupational segregation by sex all concur that it is a defining feature of our post-industrial societies (Charles, 2005; Estevez-Abe, 2005; Anker, 2001), but from then on they diverge drastically. Before we begin our review of these theories, it is important to clarify certain aspects of our subject. Firstly, although occupational segregation by sex is a phenomenon which exists all over the world (Anker, 2001), we will limit ourselves to western post-industrial societies. Secondly occupational segregation by sex can be vertical, which implies that

women are less likely to be in hierarchically powerful positions (i.e. managers), or it can be horizontal, which means women are clustered in certain sectors of the labour market (i.e. teaching jobs). If we consider both of these aspects we can draw up a rough sketch of what a 'female' job is likely to be in a developed nation. A woman's job is typically within the public sector and usually in the service industry (Turner and McMahon, 2010; Datta Gupta, Smith and Verner, 2008), her job is likely to hold less power and responsibility than her male counterparts (Mandel and Semyonov, 2006) and she will receive lower wages (OECD, 2008; Muhlau, 2010b; Belgian Presidency Report, 2010). Female jobs are also more likely to be part-time or reduced-time and they are more likely to be salaried jobs rather than self-employed positions (Elder and Johnson, 2001). The skills needed for them are those that mirror stereotypical female attributes, such as interpersonal skills, manual dexterity and a caring disposition, as well as jobs which correspond to women's traditional tasks in the household, such as cleaning, cooking and looking after children. (Anker, 2001; Hakim, 2002; Belgian Presidency Report, 2010). These are generalisations and it is important to realise that there are large disparities between countries. For example, women in liberal market economies such as Ireland, the UK or the US are more likely to be in a managerial position than women from social-democratic regimes (Turner and McMahon, 2010; Mandel and Semyonov, 2006) and are more likely to be employed in the private sector (Estevez-Abe, 2005; Mandel and Semyonov, 2006). These discrepancies have fuelled occupational segregation investigation and from them have emerged the various theories we will now review.

Polacheck's Human Capital theory is built on the corner stones of Neoclassical economic theory. Put simply, these foundations state that the individual is a rational being who evaluates costs and benefits in order to defend his self-interest and the sum of all these egotistical calculations equates to the common good. Taken in context of the labour market, this means that workers choose jobs which will maximise their abilities and preferences while minimising their 'handicaps' (Anker, 2001). The argument is the same for employers; they will choose employees who are likely to make them profits from the wages they are paying them and the 'on-the-job' training they may provide. Applied to the gender question the argument runs as follows: from an employer's point of view women have less human capital than men in certain areas, they may not have the qualifications needed or they may not have enough on-the-job training and experience, all of which make them less productive than their male counterparts. Employers being rational individuals they will consequently employ men over women (Estevez-Abe, 2005). Other theories similarly based on neo-classical economics claim that additional costs of hiring women include: their lack of flexibility, their high turnover rate and their high level of lateness and absenteeism (Anker, 2001). Equally, legislation requirements such as separate toilets or regulation concerning night work or physically taxing occupations (Anker, 2001), act as disincentives for employers as they increase the cost of hiring women. These theories highlight the importance of human capital in the gender occupational divide; they do not however delve into the

reasons women are disadvantaged in this area and by failing to do so they offer little hope to alleviate occupational segregation.

The Institutional Context

Other theories thankfully do go further. One approach focuses on the institutional context in which women evolve. For instance, research has shown that countries with stronger welfare systems suffered higher occupational gender segregation, both vertical and horizontal (Estevez-Abe, 2005; Korpi, Ferrarini and Englund, 2009; Datta Gupta, Smith and Verner, 2008; Mandel and Semyonov, 2006; Mandel, 2009). The reason for this lies in the welfare state itself. On the one hand, social policy aims to encourage female employment by providing lengthily paid parental leave and institutional arrangements for part-time work or reduced working weeks (such as the 35 hour week in France) (Mandel and Semyonov, 2006). Although these measures have been found to increase female work participation, they also have certain adverse consequences. Encouraging career breaks and part-time work ultimately emphasises women's roles as mothers and carers over their role as employees (Mandel and Semyonov, 2006). It also reduces the time women spend on the job, which means that they gain less experience and miss out on vital training, reducing their human capital and leaving them unable to compete with their male colleagues. The consequence is that women are depicted as 'risky investments' for employers, especially for jobs which require specialisation (typically 'male' jobs) and for which the cost of training or finding a replacement in case of maternity leave would be crippling (Mandel and Semyonov, 2006). Whereas private sector employers, slaves to market competition, are unable to assume these costs, the public sector is; this brings us to the other half of the argument. As the welfare state grew, it removed women's traditional tasks from the family sphere and absorbed them into the state's service sector. Jobs such as childcare, education, nursing or caring for the elderly became paid occupations, but this move into the formal economy did not ease gender segregation in these areas. Like in the home they remained distinctively 'female' jobs. The main reason for this is that, as we mentioned above, they provide the employment protection for women that the private sector cannot offer. At the same time, not all women have the burden of family responsibilities, hence they do not all need the protection that state employment provides them, nor do they all represent a cost to their employers (Datta Gupta, Smith and Verner, 2008), so why are they still predominantly clustered in the public sector?

Statistical Discrimination Theory provides us with some measure of the answer. Although employers know that not all women are less experienced, less qualified and less committed to their career, it is easier and cheaper to rely on these group characteristics when recruiting rather than evaluating each individual applicant (Anker, 2001). As we have seen, welfare states' family friendly policies increase the likelihood of women taking career breaks and prioritising family over career, making these group characteristics seem closer to the truth. This increases the scope of statistical

discrimination in these countries and with it reduces women's career ambitions (Korpi, Ferrarini and Englund, 2009). Women don't only end up in female jobs because they are less qualified or because other occupations are less suited to their responsibilities, they also make a rational decision to apply for 'female' jobs because they know they are unlikely to be employed in 'male' jobs. It follows that their choice of education reflects this assessment; it would be pointless to invest years of study into engineering to then find oneself disadvantaged when looking for a job, which partially explains why women can be less endowed in human capital.

Another author to highlight the importance of institutional context is Estevez-Abe. In order to explain occupational gender segregation, her skill-based theory focuses on two types of welfare state. On one side countries with Coordinated Market Economies (CMEs), such as Sweden or Germany, in which the state partially controls the market, encouraging long term partnerships between social partners and providing protection for jobs, unemployment and wages. This leads to a more stable labour market, in which companies work together rather than against each other and in which workers are more likely to enjoy longer term work contracts (Estevez-Abe, 2005). The education system behind this type of society is one which centres on specific non-transferable skills, which correspond to a very limited number of occupations. But because the state provides unemployment benefits, retraining in case of job loss and a general stability which prevents jobs losses in the first place (Charles, 2005), the risk of selecting this type of education is attenuated. Equally, employers are willing to invest in their employees' skill development as they will reap the benefits. Liberal Market Economies (LMEs) on the other hand have a very low level of state intervention in the market, which creates a much more elastic labour market in which the employee's only real employment protection is his qualifications. For this reason education in these countries (e.g. UK, USA or Ireland) are more likely to be based on general skills, which will be portable from one job to another. Due to the greater insecurity of jobs, employers are less likely to invest in firm-specific skills for their workers, just as much as their workers are unlikely to want to learn them. Estevez-Abe finds that CMEs were more likely to suffer from high occupational segregation than LMEs (Estevez-Abe, 2005); these findings corroborate the previous thesis which equally attributes higher occupational segregation to stronger welfare states. However, rather than placing the blame directly on family-friendly policies, this theory shifts the blame to the skill-regime. As women are more likely to have to interrupt their careers in order to care for their families (Estevez-Abe, 2005; Stier, Lewin-Epstein and Braun, 2001) they are less likely to benefit from skill-specific education or firm specific training programs. As a consequence they are more exposed to dismissal, wage depreciation or loss of position within the work place (Charles, 2005). To guard against this, CMEs offer extensive social policy (obligatory paid parental leave), which as we have seen above induces higher occupational segregation. On the other hand, in general skill regimes moving in and out of jobs is acceptable; qualification and training is more likely to be transportable, so that if upon return from maternity leave a worker is forced to change jobs, her skills will allow her to do so without submitting to wage depreciation or loss of

human capital. This means that men and women are ultimately on a more level playing field in LMEs and so suffer less occupational segregation.

The Socio-Cultural Context

The most important criticism voiced against these institutional theories is that they do not take into account the socio-cultural context of occupational gender segregation. Women do not simply choose certain occupations because they weigh up the positive and negative aspects of the question (be it in relation to family friendly policies or skill regimes). The cultural norms and values which surround them and form part of their identity dictate the roles and occupations which are acceptable and those which, because they are deemed 'male', must be avoided (Charles, 2005). Women, for example, are considered to be more interpersonal, better at communicating, more caring, more docile, more honest but also less proficient at maths and sciences (Anker, 2001), while those attributes which are inherently male, such as greater physical strength, natural ability to supervise and give orders, are of higher value and deserve higher status than those of women (Charles, 2005). Economic restructuring happened within this socio-cultural context. According to Charles (2005) this key mix of factors has led to greater occupational segregation. The expansion of the service sector has meant more jobs which reflect most of the characteristics attributed to women (sales jobs, receptionists, home helps). Economic rationalization, another feature of structural change, brought more jobs into the formal economy. Demand for workers, especially female service sector workers, rose, making it necessary to hire women that may not be career centred. One of the consequences of more women with family responsibilities entering the workforce is that the service sector has been forced to adapt to women's needs to a certain extent (Charles, 2005), with part-time work and flexi-time having become an integral part of the service sector and also of women's jobs. Although structural theory is effective in explaining horizontal and vertical segregation, some have argued that to ignore the institutional context is to overlook the fact that in many cases of structural change, growing service sectors are the fruit of social policy – this is particularly true for Scandinavian countries (Estevez-Abe, 2005).

A variation of cultural-structural theory focuses on work organization blaming the new post-Fordist work model for ongoing occupational segregation, specifically vertical segregation (Van Echtelt et al., 2009). The old 'time-greedy' (Van Echtelt et al., 2009, p.189) model of work was based on the premise that workers had no other obligations than their job; women, who carried the bigger burden of domestic work, were clearly at a disadvantage. It was hoped, that a new flexible and autonomous way of organising workers would change this (Appelbaum, E. and Bailyn, L., 2009 cited in Van Echtelt et al., 2009), but post-fordist models have shown themselves to be even more demanding than their predecessor, creating a highly competitive work environment in which flexibility and autonomy just push employees to work longer and harder in order to get to the top. As socio-cultural norms haven't evolved enough for men to shoulder their fair share of domestic

responsibilities, women are still not competing on a level playing field and so cannot hope to break through the glass ceiling of vertical segregation.

Gender or feminist theory has a completely different take on the question; for these theorists gender occupational segregation is product of patriarchal values and notions of female subordination (Anker, 2001) which we assimilate as children and keep on assimilating throughout our lives through our parents, our peers, the education system and the media. These socio-cultural norms give birth to a whole range of stereotypical ideas, which although they do not necessarily reflect reality, are too deeply embedded in our collective consciousness for us to refute them. Because of this there can be no such thing as completely individual choice. These powerful stereotypes not only dictate the type of jobs women will end up in, but also the hours they will work, and the length of their employment. One such stereotype is the idea that women have a more caring nature. This qualifies them for jobs as nurses, teachers, midwives and social workers, but interestingly we find fewer women in caring occupations of higher ranking, such as doctors or university professors because of another gender stereotype which dictates that women are naturally docile and disinclined to supervise or give orders (Anker, 2001). As their work force is predominantly female, these occupations then adapt to women's needs and become more flexible or prone to part-time organisation (Anker, 2001). Another similarly 'fuzzy' factor used to explain gender occupational segregation is discrimination. In its report on gender discrimination, the OECD (2008) defines it as 'the unequal treatment of equally productive individuals only because they belong to a specific group'. By this logic women are limited to certain occupations because employers feel a personal dislike towards their gender in general. Discrimination being a notoriously difficult factor to measure, it can be most easily defined as what is left over after we take into account all other reasons for segregation: education, experience, motivation and expectation (Muhlau, 2010b). Using this method, it was found that discrimination accounts for 8% of gender occupational segregation (OECD, 2008). Of course, the reasons for such discrimination can be traced back to gender stereotyping, making these two theories congruous.

Conclusion

To conclude this review, it seems fair to say that there is a wealth of different explanations for gender occupational segregation. The reasons offered as to why women end up in 'female jobs' range from their personal choice to the cultural values that surround them, from the type of economy they evolve in to the kind of welfare state they live in. Yet the consequences of the phenomenon are quite straightforward. On the one hand, systematically excluding half of the active population from certain occupations is clearly underutilizing the workforce (Anker, 2001). In the global competitive market that we live in today, submitting your economy to such inefficiency seems imprudent. On the other hand, such wasteful management of human resources, when most European countries are struggling with the costs of a rapidly aging population, seems like a glaring oversight (OECD, 2008). Boosting fertility rates is an obvious way to counter these costs, but if women continue to have to make a

choice between a career and starting a family they are unlikely to show any signs of equalising participation in the workforce (Van Echtelt et al., 2009; Datta Gupta, Smith and Verner, 2008). Equally, occupational gender segregation affects women's status and income and consequently how we, as societies, perceive women. This has a huge effect on income inequality, poverty, economic dependency and mortality, not only for the present generation but also for those to come (Anker, 2001). As our segregated vision of gender occupation is passed down through the education system, the media and our collective consciousness, we are perpetuating a cycle of gender inequality and submitting our children to the same fate.

References

Anker, R., 2001. 'Theories of Occupational Segregation by Sex: An Overview'. In: Fetherolf Loufti, M., ed., 2001. *Women, Gender and Work: What is Equality and How Do We Get There?*. Geneva: International Labour Office.

Belgian Presidency Report, 2010. 'The Gender Pay Gap in Member States of the European Union: Quantitative and Qualitative Indicators'. In: *Council of the European Union*. Brussels, Belgium. 26 November 2010.

Charles, M., 2005. 'National Skill Regimes, Postindustrialism, and Sex Segregation'. *Social Politics: International Studies in Gender, State and Society,* 12(2), pp.289-316

Datta Gupta, N., Smith, N. and Verner M., 2008. 'The impact of Nordic Countries' Family Friendly Policies on Employment, Wages, and Children'. *The Review of Economic Households*, 6(1), pp.65-89.

Elder, S. and Johnson, L.J., 2001. 'Sex Specific Labour Market Indicators: What They Show'. In Fetherolf Loufti, M., ed., 2001. *Women, Gender and Work: What is Equality and how do we get there?*. Geneva: International Labour Office.

Estevez-Abe, M., 2005. 'Gender Bias in Skills and Social Policies: The Varieties of Capitalism Perspective on Sex Segregation'. *Social Politics: International Studies in Gender, State and Society*, 12(2), pp.180-215.

Goldin, C., 2006. 'The Quiet Revolution that Transformed Women's Employment, Education and Family'. National Bureau of Economic Research, Cambridge, MA. *Working Paper no.11953*.

Hakim, C., 2002. ‘Lifestyle Preferences as Determinants of Women's Differentiated Labor Market Careers’. *Work and Occupations*, 29(4), pp.428-459.

Korpi, W., Ferrarini, T. and Englund, S., 2009. ‘Egalitarian Gender Paradise Lost? Re-Examining Gender Inequalities in Different Types of Welfare States’. In: EMPLOY-FAMNET, *Workshop*. Berlin, Germany,11-12 May 2009.

Mandel, H. and Semyonov, M., 2006. ‘A Welfare State Paradox: State Interventions and Women’s Employment Opportunities in 22 Countries’. *American Journal of Sociology,* 111(6), pp.1910-1949.

Mandel, H., 2009. ‘Configurations of Gender Inequality: The Consequences of Ideology and Public Policy’. *British Journal of Sociology*, 60(4), pp.693-716.

Muhlau, P., 2010a. *Lecture*. 23 November 2010. Trinity College Dublin, unpublished.

Muhlau, P., 2010b. *Lecture*. 30 November 2010 Trinity College Dublin, 2010, unpublished.

OECD, 2008. ‘The Price of Prejudice: Labour Market Discrimination on the Grounds of Gender and Ethnicity’. In: *OECD Employment Outlook*. Ch. 3.

Stier, H., Lewin-Epstein N., Braun, M., 2001. ‘Welfare Regimes, Family-Supportive Policies, and Women's Employment Along the Life-Course’. *The American Journal of Sociology*, 106(6), pp.1731-1760.

Turner, T. and McMahon, J., 2010. ‘Women’s Occupational Trends in the Irish Economy: Moving Towards High-skilled Occupations or Evidence of Deskilling?’. *Gender, Work and Organization*. Available through: Wiley Online Library.

Van Echtelt P., Glebbeek A., Lewis S. and Lindenberg, S., 2009. ‘Post-Fordist Work: A Man's World? Gender and Working Overtime in the Netherlands’. *Gender and Society*, 23(2), pp.188-214.

REVIEW OF THE EMPIRICAL EVIDENCE ON THE EFFICIENT MARKET HYPOTHESIS IN THE FOREIGN EXCHANGE MARKETS

Róisín Donnelly

Senior Sophister

The efficient markets hypothesis is an often-criticised theory whose practical value is questionable. Roisin Donnelly provides an extremely well researched account of the EMH and the empirical evidence for its applicability to the foreign exchange markets, ultimately arguing that it seems to have taken hold over time in more mature markets.

Introduction

An old economics joke has been employed to demonstrate the efficient markets hypothesis (EMH): An economist is walking down the street with a companion when they see a $100 bill on the ground. As the companion attempts to reach down and pick it up, the companion stops him and says 'Don't bother – if it were a genuine $100 bill, someone would have already picked it up' (Lo and MacKinlay, 1999, p.6).

The EMH has been a subject of debate since its inception in the 1960s, dividing academics and professionals. In general, academics have believed that the EMH holds, but have been proven wrong in applied finance (LeBaron, 1999). This essay will first discuss the EMH and apply its theory to the foreign exchange market. The second section will then examine evidence of efficiency in the market, in terms of forward and spot rates and the profitability of technical rules. This section will track the evolution of research on the EMH, particularly noting recent evidence of the EMH. The final section will consider the implications of the EMH in the current market, given unprecedented central bank interventions, emerging markets and the aftermath of the global financial crisis.

The Efficient Market Hypothesis

Fama (1970) stated that an efficient market is where prices fully reflect all available information. If this is true, no profit opportunities can be left unexploited. The market is therefore considered to be a 'sensitive process of all new information', and responds with price fluctuations instantaneously. (Baillie and McMahon, 1989, p.40)

Fama (1970, p.383) further developed the EMH into three forms indicating different levels of available information: weak form efficiency reflects past historical prices; semi-strong form efficiency reflects both historical prices and information that is 'obviously publicly available'; strong form efficiency questions whether some information is private and therefore, if some investors have 'monopolistic access' to price relevant information. If all public and private information is available to all investors then strong form efficiency holds.

Interestingly, for the EMH to hold, there must also be belief that the market can be in disequilibrium, which is an incentive to gather information to uncover arbitrage. Therefore arbitrageurs play a key role in the existence of the EMH (Grossman and Stiglitz, 1980). It has been suggested that the degree of efficiency can also vary across markets (e.g. markets in emerging countries may be less efficient than the US market due to a less intensive analysis) (Bodie, Kane, and Marcus, 2008).

Simple EMH assumes that investors are both 'endowed with rational expectations' and risk neutral (although the model can be modified to adjust for risk) (Sarno and Taylor, 2002, p.5).

Specifically for the foreign exchange market, semi-strong efficiency can be divided into two categories: single market efficiency and multimarket efficiency (Geweke and Feige, 1979). Single market efficiency implies that all public information about a single exchange rate is part of available information. Multimarket efficiency implies that information on all exchange rates is part of available information (Sarno and Taylor, 2002).

Under the EMH, forward prices follow a martingale, which is 'a statistical process in which the expected value of successive changes are independent of previous changes' (Dooley and Shafer, 1984, p.50). This suggests that the spot rate today is the best indicator of the spot rate in the future and therefore changes in the spot rate are serially uncorrelated, appearing random, with equal probabilities of both appreciation and depreciation (Baillie and McMahon, 1989).

A key insider in terms of private information in the foreign exchange market is the central bank. Baillie and McMahon (1989) suggest that strong form efficiency is unlikely to hold, due to non random interventions by central banks. If a central bank makes zero expected profit on intervention, then the market is strong form efficient (Sweeney, 2000).

Empirical Evidence on the EMH

Generally, tests for market efficiency involve determining whether the current spot and forward rates have a long term relationship with information available (Levich, 1978). There are several tests for spot market efficiency. One method is to test the null hypothesis that changes in spot rates are serially uncorrelated. Another method is to test technical trading rules; whether profit can be made from technical analysis such as filter rules[1] (Levich, 1978) and moving-average rules[2].

[1] A filter rule can be defined as generating 'buy and sell signals by the following: buy a

Tests of forward market efficiency include regression models to determine whether the forward rate is an unbiased predictor for the future spot rate. If this is true, then the market is weakly efficient (Longworth, 1981). Furthermore, a semi-strong form test is whether the spot rate change is on average equal to the lagged forward premium (Longworth, 1981).

It is suggested that if investors' expectations are part of available information, then they are reflected in market prices, and hence the forward rate is an unbiased predictor of the future spot rate (Levich, 1979). Many multinational companies would use this as a forecast because it reflects the 'collective wisdom of many well informed, profit seeking traders' (Levich, 1979, p.50). Furthermore, it is useful as it can be revised rapidly to reflect new information and is inexpensive to use as a method of forecasting.

Theoretically, the merits of the current forward rate predicting the spot rate are obvious. However is it actually that accurate in practice? Levich (1979, p.54) found that, across a wide range of currencies, the forward rate is an unbiased predictor of the level of the future spot rate: the 'level of the forward rate will explain a high percentage of the variation in the level of the future spot rate'. However the forward rate premium is not very accurate in predicting exchange rate changes and hence Levich (1979) concludes that while the forward rate is an unbiased indicator of the future spot rate, it is a poor one.

Kohlhagen (1979) found that, in the long run, the forward rate is an unbiased predictor of the future spot rate. However, he identified short run periods in which the current forward rate was considerably different to its related subsequent spot rate. Furthermore, Geweke and Feige (1979) found inconclusive evidence for single market efficiency but firmly rejected multimarket efficiency using econometric testing of the correlation between the spot and forward rate. They found inefficiencies varied over time, due to both risk aversion (predominantly in the 1970s) and transaction costs (predominantly in the 1960s).

Longworth (1981) also rejected both weak and semi-strong form tests of the forward rate as a predictor of the future spot rate in the US Dollar (USD)-Canadian Dollar (CAD) currency pairing. Moreover, it was found that the current spot rate is a better predictor of the future spot rate, echoing earlier findings by Fama (1984). Later research indicated that spot and forward rate are non-stationary with unit roots and are therefore biased and inconsistent (Bakshi and Naka, 1997). What became clear was that spot and forward rates were cointegrated[3]. It was found that the seven largest exchange rates were tied together in a long term relationship (Baillie and Bollerslev, 1989). Therefore exchange rate determination is not based on individual currency fundamentals as previously thought. Furthermore,

---

currency whenever it rises x percent above its most recent trough; sell the currency and take a short position whenever the currency falls x percent below its most recent peak' (Nguyen, 2004, p. 3).

[2] A moving-average rule can be defined as 'when the short-term moving average penetrates the long-term moving-average from below (above) a buy (sell) signal is generated' (Nguyen, 2004, p. 3).

[3] Cointegration violates efficiency as if 'exchange rates are cointegrated, then it is possible to predict one on the basis of another' (Wua and& Chen, 1998, p. 831).

as these relationships determine future prices, there is evidence of a time varying risk premium, which is a violation of weak-form efficiency.

Testing for Profits

Under the EMH, there should be no trading rule found that can beat the market (Copeland, 2008). Invariably, examining whether the efficient market hypothesis holds often tests if technical analysis holds, which would disprove the EMH. Therefore a review of the profitability of technical trading is necessary in this discussion of efficiency. These rules are usually tested against a simple buy and hold strategy (Copeland, 2008). This is the primary technique used for testing efficiency in the foreign exchange market (Nguyen, 2004).

Dooley and Shafer (1984) found that using filter rules of 1%, 3% and 5% for nine currencies from 1973 to 1981 was profitable for the sample period. However, they also found elements of risk in the trading rules, as there was sub-periods of losses in each currency (Nguyen, 2004).

Sweeney (1986) examined the profitability of filter rules against a benchmark of a buy and hold strategy for nine currencies from 1973 to 1990. It was found that all except the 10% filter beat buy and hold, and that the 0.5% and 1% filters were statistically significant, suggesting inefficiency in the market. However, taking into account transaction costs, they found that only the 1% filter was still statistically significant. Nevertheless, this suggests that profits can be made using a 1% filter after transaction costs, which is a blatant contradiction of efficiency.

Schulmeister (1988) tested the weak form efficiency of the USD versus the Deutsch Mark (DM) between 1973 and 1986. Testing both the moving-average and momentum rule, it was found that the average annual return over the entire period was 15%. As such strategies exploit previous information to make profit Schulmeister (1988) concluded that foreign exchange markets are not even weak form efficient.

However, more recent evidence suggests that profits have been on the decline. Levich and Thomas (1991) analysed both the filter rule and the moving-average rule over the 15 year period from 1976 to 1990 for the British Pound (GBP), CAD, DM, Japanese Yen (JPY) and Swiss Franc (CHF). They tested the filter rule from 0.5% to 5% and three moving average rules. Significantly, although they found profitability in these trading results and therefore market inefficiency, when they divided the sample period into three, the period from 1986 to 1990 showed declining (albeit positive on average) profitability.

Interestingly, Curcio et al. (1997), using filter rules on the intra-daily foreign exchange market, found evidence against profitability of trading strategies; therefore in favour of market efficiency. Studying two periods - April to June in 1989 and January to June in 1999 - they tested the DM, JPY and GBP against the USD. In the first period, one third of strategies generated significant profit. However, once trading costs were accounted for, most cases of profits were eliminated. In the later period there were only a very small number of significant profits found, which were matched by

the same number of significant losses, suggesting market efficiency. Incorporating transaction costs into the later test sample eliminates all significant returns and therefore endorses market efficiency.

Rubio (2004) further developed evidence of market efficiency from the period 1975 to 2004. 48 strategies for each currency pairing were tested; that of the USD versus the Austrialian Dollar (AUS), CAD, JPY, CHF and GBP. These pairings represented significant percentages of the foreign exchange market turnover (4%, 4%, 20%, 5% and 11% respectively). The strategies tested were based on rules such as the moving-average rule and filter rules. Out of 48 strategies for each of the currencies, they found only one or two were profitable after taxation and commissions. Therefore, in general, they found the market did not produce excess profits and was efficient.

Pukthuanthong-Le and Thomas III (2008) examined weak form efficiency for the period from 1975 to 2006, testing the moving-average and momentum rules for the GBP, CHF, JPY, and CAD based on the USD. It was found that trading rules had worked in the past but have been worthless since 2000, signifying that the market has become weak form efficient. They suggest that reasons for prior inefficiency could be due to immature markets. Consequently, they tested emerging market currencies for inefficiency using the Brazilian Real (BRL), Mexican Peso (MXN), Russian Rouble (RUB), New Zealand Dollar (NZD), and the South African Rand (ZAR). The results mirror the early inefficiencies in the major currency markets. From the period between 2000 and 2006 trading rules were profitable (except for the MXN). If these emerging currencies follow the path of the major currencies, it should be expected that they become more efficient (Pukthanthong-Le and Thomas III, 2008). Oh, Kim and Eom (2007) found that, on average, markets with large liquidity and trading volumes (such as in Europe and North America) showed higher market efficiency than those with smaller liquidity (such as in the Asian and African markets). Overall, it appears that mature markets have become efficient over the last 10-15 years, following decades of inefficiency. However, emerging markets remain inefficient.

LeBaron (1999) found that exchange rate predictability is dramatically reduced in periods in which there is no central bank intervention. He found that, by examining the period from 1979 to 1992 and testing the moving-average rule of the USD with the JPY and DM there was significant forecast ability, generated by large Sharpe ratios (0.6 to 0.9). This is compared with buy and hold strategy Sharpe ratios of 0.3-0.4. However, removing periods of intervention by the Federal Reserve and repeating the tests produced results that were not statistically significant (LeBaron, 1999). This supports Dooley and Shafer's (1984, p.48) notion that 'at worst, central bank intervention would introduce noticeable trends into the evolution of exchange rates and create opportunities for alert private market participants to profit from speculating against the central bank'.

The Current Situation

As has been described, central bank intervention can often disturb market efficiency. Therefore, the recent trend in intervention should prove interesting in terms of measuring efficiency.

Japan recently intervened in the Yen for the first time since 2004, selling the Yen against the Dollar to stem gains and protect its exports (Bloomberg, 2010a). The use of Quantitative Easing (QE) by the Federal Reserve has been criticised as being ineffective due to information uncertainty (Fell, 2010). Furthermore, Fell (2010) reports that the Federal Reserve has not made specific pledges about accommodative monetary policy and inflation (unlike when Japan used QE). Therefore, investors 'have no clear-cut policy target upon which to anchor the likely path of forward rates' (Fell, 2010, p.1). This suggests that the actions of the Federal Reserve are futile; although they want their currency to remain weak, they have not given enough signals to investors to make informed decisions based on new information such as inflation.

A possible impact of the Fed's latest round of quantitaitve easing (QE2) is that the excess liquidity in the US market will flow to emerging markets, resulting in inflation and bubbles(Saunders, 2010; The Economist, 2010). This could lead to quantitative tightening, where emerging market governments introduce capital controls to reduce inflows into their countries. Both Brazil and Thailand have already imposed high taxes on bonds, with Indonesia and South Korea possibly following suit. Such measures by emerging markets would inhibit efficiency due to greater transaction costs (Pilling, 2010).

A recent report has highlighted the changing nature of currency traders' behaviour as a result of the financial crisis. Recent developments suggest that the risks of currency trading are greater; traders are more likely to be fired if they make big losses and may not be paid the traditional large bonuses for big gains (Melvin and Taylor, 2009). Moreover, a recent survey of foreign exchange professionals showed that 80% were worried about the impact of recent regulations on profitability (Bloomberg, 2010b). One possible effect of both increased regulation and decreased bonuses could be a disincentive to seek arbitrage opportunities. As arbitrage is vital to the EMH, this could hinder the future efficiency of the market.

Conclusion

Similarly to the foreign exchange market, recent research on the stock market by Sullivan, Timmermann and White (1999) suggests that it has become more efficient, thanks to cheaper computing power, lower transaction costs and increased liquidity. These can also be applied to foreign exchange markets. Aggarwal, Lucey and Mohanty suggest that the foreign exchange market boasts high liquidity and low trading costs (2006).

Malkiel (2003) argues that although inefficiencies exist in the stock market, such as persistent predictability of prices, they always eventually disappear. This essay has shown that this appears true in the foreign exchange market; trading strategies that once were profitable have been deemed redundant in the last 10 years. The focus will now shift to inefficient emerging markets and how globalisation will assist the rapid growth in liquidity and efficiency.

References

Aggarwal, R., Lucey, B. M., and Mohanty, S. K., 2006. 'The Forward Exchange Rate Bias Puzzle: Evidence from New Cointegration Tests'. Dublin, *IIIS Discussion Paper No. 123*.

Baillie, R. T., and Bollerslev, T., 1989. 'Common Stochastic Trends in a System of Exchange Rates'. *Journal of Finance* , 44, pp.167-181.

Baillie, R. T., and McMahon, P. C., 1989. *The Foreign Exchange Market - Theory and Econometric Evidence.* Cambridge, Cambridge University Press.

Bakshi, G. S., and Naka, A., 1997. 'Unbiasedness of the Forward Exchange Rates'. *The Financial Review ,* 32 (1), pp.145-162.

Bloomberg, 2010a. 'Japan May Sell Yen for Second Day to Protect Economy'. *Bloomberg*, [Online] 15 September. Available at: http://www.bloomberg.com/news/2010-09-15/japan-intervenes-for-first-time-since-2004-as-yen-surge-threatens-recovery.html [Accessed: December 12, 2010]

Bloomberg, 2010b. 'Bloomberg FX Survey Reveals Anxiety on Market Regulations'. *Bloomberg*, [Online] 18 November. Available at: http://www.bloomberg.com/news/2010-11-18/bloomberg-fx-survey-reveals-anxiety-on-market-regulations.html [Accessed: December 3, 2010]

Bodie, Z., Kane, A., and Marcus, A. J., 2008. *Investments.* Singapore, McGraw-Hill/Irwin.

Copeland, L., 2008. *Exchange Rates and International Finance.* Essex, Addison-Wesley Publishers.

Curcio, R., Goodhart, C., Guillaume, D., and Payne, R., 1997. 'Do Technical Trading Rules Generate Profits? Conclusions from the Intra-Day Foreign Exchange Market'. *International Journal of Financial Economics*, 2, pp.267-280.

Dooley, M. P., and Shafer, J., 1984. 'Analysis of Short-Run Exchange Rate Behavior: March 1973 to November 1981'. In: Bigman, D. and Taya, T., *Floating Exchange Rates and the State of World Trade and Payments*. Beard Books. Ch.3, pp. 43-72.

Fama, E. F., 1970. 'Capital Markets: A Review of Theory and Empirical Work'. *Journal of Finance ,* 25(2), pp.383-417.

Fama, E., 1984. 'Forward and Spot Exchange Rates'. *Journal of Monetary Economics* , 14 (3), pp.319-338.

Fell, C., 2010. *QE2:* 'A Shot in the Dark'.*Charlie Fell*, [Online] 5 November. Available at: http://www.charliefell.com/index.php/articles/2-general/84-qe2-a-shot-in-the-dark [Accessed: November 30, 2010]

Geweke, J. F. and Feige, E. L., 1979. 'Some Joint Tests of the Efficiency of Markets for Forward Foreign Exchange'. *Review of Economics and Statistics*, 61, pp.1-32.

Grossman, S. J. and Stiglitz, J. E., 1980. 'On the Impossibility of Informationally Efficient Markets'. *American Economic Review* , 70, pp.393-409.

Kohlhagen, S. W.,1979. 'The Forward Rate as an Unbiased Predictor of the Future Spot Rate'. *Columbia Journal of World Business ,* 14, pp.77-85.

LeBaron, B., 1999. 'Technical Trading Rule Profitability and Foreign Exchange Intervention'. *Journal of International Economics ,* 49(1), pp.125-143.

Levich, R. M., 1979. 'Are Forward Exchange Rates Unbiased Predictors of Future Spot Rates?' *Columbia Journal of World Business* , Winter, pp.49-61.

Levich, R. M., 1978. 'Further Results on the Efficiency of Markets for Foreign Exchange'. Federal Reserve Bank of Boston. In: *Conference Series 20*. Boston, October 1978. pp.58-80.

Levich, R. M. and Thomas, L. R., 1991. 'The Significance of Technical Trading-Rule Profits in the Foreign Exchange Market: A Bootstrap Approach'. National Bureau of Economic Research, Cambridge, MA, *Working Paper Series no.3818*.

Lo, A. W. and MacKinlay, C., 1999. *A Non-Random Walk Down Wall Street.* Princeton: Princeton University Press.

Longworth, D., 1981. 'Testing the Efficiency of the Canadian-U.S. Exchange Market under the Assumption of no Risk Premium'. *Journal of Finance*, 36(1), pp.43-49.

Malkiel, B. G., 2003. 'The Efficient Market Hypothesis and Its Critics'. *The Journal of Economic Perspectives ,* 17 (1), pp.59-82.

Melvin, M. and Taylor, M., 2009. 'The Crisis in the Foreign Exchange Market'. Centre for Economic Policy Research, London. *CESifo Working Paper Series no.2707*.

Nguyen, J., 2004. 'The Efficient Market Hypothesis: Is It Applicable to the Foreign Exchange Market?'. University of Wollongong, New South Wales, Australia. Economics Working Paper no.04-20.

Oh, G., Kim, S. and Eom, C., 2007. 'Market Efficiency in Foreign Exchange Markets'. *Physica A: Statistical Mechanics and its Applications* , 382 (1), pp.209-212.

Pilling, D., 2010. 'Emerging Markets to counter QE2'. *Financial Times*, November 3. p. 2.

Pukthuanthong-Le, K. and Thomas III, L. R., 2008. 'Weak-Form Efficiency in Currency Markets'. *Financial Analysts Journal* , 64 (3), pp.31-52.

Sarno, L., and Taylor, M., 2002. *The Economics of Exchange Rates.* Cambridge: Cambridge University Press.

Saunders, E., 2010. 'QE2 Bubble Fears'. *Financial Times*, November 10. p.10.

Schulmeister, S., 1988. 'Currency Speculation and Dollar Fluctuations'. *Quartely Review, National Bank of Lavoro*, 167, pp.343-365.

Sullivan, R., Timmermann, A., and White, H., 1999. 'Data-Snooping, Technical Trading Rule Performance, and the Bootstrap'. *Journal of Finance*, 54, pp.1649-1691.

Sweeney, R. J., 1986. 'Beating the Foreign Exchange Market'. *Journal of Finance*, 41, pp.163-182.

Sweeney, R. J., 2000. 'Does the Fed beat the Foreign-Exchange Market'. *Journal of Banking & Finance*, 24, pp.665-694.

The Economist, 2010.'Beyond Bretton Woods 2 Is there a better way to organise the world's currencies?' . *The Economist*, [Online] 4November. Available at: http://www.economist.com/node/17414511?story_id=17414511 [Accessed: December 4, 2010]

Wua, J.-L. and Chen, S.-L., 1998. 'Foreign exchange market efficiency revisited'. *Journal of International Money and Finance*, 17(5), pp.831-838.

THE PARADOX OF GDP/GNP AS DETERMINANTS OF HUMAN PROGRESS – EFFECTS ON WELFARE AND EQUALITY

Marc Morgan

Senior Freshman

Economic growth is often seen as a panacea that every country must strive towards. Marc Morgan challenges this obsession with economic output, pointing to its many failings in terms of assessing our quality of life. He argues that satisfying materialistic desires does not necessarily lead to human progress, and should not be the primary objective at the expense of equality.

The Paradox of Growth and the Misunderstanding of Progress

The economic model that has presided over most of the western world over the last century has been built on an incessant obsession for output growth. Throughout the century this growth has been viewed as the appropriate measure of society's material living standards, almost without question. This paper posits that it is time to seriously reassess our affinity to GDP/GNP statistics for the very reason that while output growth has been upwardly linear in direction over time growth in basic societal wellbeing has not followed that same direction. It is no wonder that in our schools we are taught as 'consumers' rather than as human beings. This paradox of production going in one direction and human welfare going in another, arouses the suspicion that the production (of goods and services) is no longer in the hands of the people it is meant to satisfy. More worryingly, the case can be made that individuals are not being considered the best judges of their own welfare; that right being reserved for the GDP/GNP demands of competing nations, at the helm of economic 'experts'.

Undoubtedly what should be of interest in any economic model is the welfare of each individual and how it can be maximised to reflect the optimal welfare of society. The notion of 'human progress', used widely in economic and social literature, begs a qualitative framework. Progress, as history should indicate, is less a matter of fulfilling future desires than of eliminating recurring social and economic problems of the past like persisting unemployment, closures, wars and disease. History does again show that growth in GDP/GNP tends to affiliate itself more with the former aim than with the latter. As a result there must be mutual understanding between the needs of society and the needs of the individual, a change from competition to cooperation. Growth of material

output should be replaced by 'growth of humanity' (Ridoux, 2009: 11). Ridoux tells us that our focus should not be fixed on purchasing power but rather on '*living power*' (Ibid: 16). In a model of such importance the notion of equality between peoples, how we understand and evaluate it, will be crucial to the development of any future model in the years to come. Misrepresentations of equality, especially in relation to liberty, all boil down to what people themselves want these terms to signify. In other words it all comes down to what type of society we want to live in; under what moral principles. Here I would like to just analyse some of the irrational disputes between proponents of different moral societies, after having dwelled on the issue of which definitions of equality best define the progress we should strive for.

GDP/GNP: Sins of Omission and Alternatives

Since the Great Depression of the 1930's the economic indicators: Gross Domestic Product (GDP) and Gross National Product (GNP) have established themselves as the standard bearers for neoclassical macroeconomics. The former measures material welfare in terms of the value of output (goods and services) produced by domestic residents within an economy during the year, while the latter measures this welfare as of the value of what is produced, during the same timeframe, in output by all national individuals and enterprises located domestically or overseas. The most obvious criticisms in relation to these indicators as sole measurements of material welfare are criticisms of omission – what the indicators fail to include. These are numerous. For example: 'the work of volunteers, domestic housework, and the transfer of knowledge or the birth of a human being' (Ridoux, 2009: 53) are not accounted for in GDP/GNP calculations, simply because they do not necessarily command a price. The same can be said for working hours, length of annual holidays, literacy rates, mortality rates, life expectancy, and environmental pollution among other countless examples. But yet these are all key determinants of material welfare (without a healthy, well educated environment with sufficient free time one could not possibly enjoy the material goods in one's possession). Furthermore, while the value of output can be measured per head of population ('per capita') it discriminates between the ideal parameter and the real parameter in relation to incomes per capita. We should ask ourselves why more attention is given to an average allocation of incomes over the true allocation as measured by the spread of incomes? The well known fact that large variances in income allocation exist requires explanation into why some individuals 'are wealthier than others – and that depends on factors to do with the structure and dynamic of capitalist society' (Harmon, 2009:43) which GDP/GNP completely ignores. Since the Great Depression global and national levels of wealth have increased exponentially and yet on average employment and real wage levels have been more or less constant, amid large volatility (Harman, 2009). Ridoux comments how in France, since 1980, 'unemployment has risen 50 per cent, while, during the same period, GDP has incremented 156 per cent!' (Ridoux, 2009: 56). This apparent paradox continues to persist because we remain attentive to GDP/GNP statistics empty of any meaningful content.

In the end, given all the data that national accounts provide us with, it would be wise to note what type of materialism we are left with. Ideally, national accounts would also reveal the nature of what is produced in an economy during the year, but they of course do not. This in itself is a major worry when we have an economic model which heavily relies on material welfare as a measure of an individual's personal welfare. It is not always the case that when material welfare increases, overall well being will increase. As John Kenneth Galbraith put it quite bluntly, it is 'the weapons culture which underlies the macroeconomic stabilization of the economy'[1], but contradictory to the notion of human welfare, a culture 'thoughtfully designed to destroy all life' (Galbraith, 1958: 257). It is of no surprise then that a distortion exists between the happiest/healthiest countries in the world and the richest. According to the Happy Planet Index (HPI) the countries topping the global HPI list come largely from Central America, South America and the Caribbean – parts of the world where GDP/GNP levels would show otherwise. Criticism of output growth as counting for individual well-being is evident from proposed alternative indicators to GDP/GNP. For example under the criterion of 'life satisfaction' the HPI has found that 'the averages for countries tend to be higher where people within that country enjoy higher levels of social capital, better climate, richer natural resources, higher life expectancy, better standards of living and more voice within government' (Happy Planet Index Report 2.0, 2009). No explicit preference is given to material wellbeing or (abstract) output growth. Other indicators like the Human Development Index (HDI) or the Genuine Progress Indicator (GPI) alongside the HPI represent real progress in attempting to address problematic distortions in human welfare evaluation. While not identical in their respective measurement procedures these indicators do all presuppose that individuals themselves are the best judges of their own welfare.

Judges of Wellbeing: Problems with Equality and the GDP/GNP-Inequality Paradox

This is an issue not often debated but crucially important to understand. It concerns the problem of free will. If we can classify self management as a good in itself, as I see it, then once people are granted control of their own economic lives a freedom to judge themselves should also logically follow. Nevertheless defining the individual within a collective, resulting in society becoming the best judge of an individual's welfare might be a more desired approach. As John Stuart Mill states (echoing Jesus of Nazareth): '...laws and social arrangements should place the happiness or the interest of every individual as nearly as possible in harmony with the interest of the whole' (Mill, 2001: 17). So the optimal quest for humanity is to ensure that society as a whole becomes the best judge of an individual's welfare, as determined by the individual judgement of the welfare others, allowing for the harmony Mill speaks about between the subject and the collective.

Such a scheme would require the condition of *equality* amongst its subjects. Once this term is mentioned, however, difficulties tend to arise in what it actually means to be equal, and how its

[1] Here Galbraith is making particular reference to the US economy, but the same line of reasoning could hold for countries of the Middle East, Russia and China at present, as well as much of Europe during the 1940's.

attainment is supposed to be measured? Are we concerning ourselves with equality of opportunity, equality of outcome, equality of contribution or an equality of effort? What can be pointed out is that the first three heavily rely on factors outside of an individual's control, while the latter notion of equality 'is something people have control over' and hence it is believed that only 'greater sacrifice deserves greater reward...' (Hahnel and Albert, 2002b: 28). But given the lack of a quantitative measurement for equality, how can we ever assign rewards to characteristics of a notion that we can't come to commonly define? If we strive for income levels as being the measure then we face the problem of utility because 'different people derive different amounts of satisfaction from the same income levels' (McAleese 2008: 42). Then there arises the issue with quantifying such utility, is the satisfaction of purchasing power measurable?

Equality of contribution, considered by many right wing liberals to be the essence of human freedom (Nozick, 1974) turns out to face similar obstacles. The fact is that 'people will always have different abilities to benefit others' (Hahnel and Albert, 2002a: 9) not to mention different financial positions. This in turn causes major difficulties for any equality of welfare. Equality of outcome encounters the discrimination of different levels of effort, simply transferring injustices from the sphere of unequal opportunity to unequal effort and uniform outcome. Tying in neatly with the goal of creating a harmony between the individual and the collective, whether in the workplace or within the larger society, in the words of Hahnel & Albert 'there is no better way to judge efforts than by a jury of fellow workers who serve on an effort rating committee on a rotating basis' (Ibid, 2002a).

Those who are not yet convinced of this procedure might think the easiest route to take is to turn our back on achieving equality and attempt to avoid *inequality*. But only by knowing its causes can inequality be eroded, and it seems that throughout moral scholarship inequality stems from a denial of an equality set by the original state of nature (Rousseau, 1984). This is caused by society's urge for material accumulation which can arguably be encouraged and shaped by the infinite demands of GDP/GNP measurement. If all the output produced under GDP/GNP measurement is not consumed then the danger of an overproducing economy forces the domestic population to consume this output under the false image that it will increase an individual's material welfare. But with its disregard for an individual's initial economic endowment a market economy only recognises those individuals in possession of the only true free entity in the economic world – capital. Global inequalities then are essentially enhanced by the concentration of capital among (certain) competitive individuals with concentrated goals to accumulate material goods. So as the production of these goods increases the gap between those controlling production and those doing the producing widens. Nevertheless an economy under such conditions would be *richer* according to GDP/GNP figures. The reality and impact of capital's dynamism as warned by Marx (1867), as well as its paradoxes might return those who sought efforts to avoid inequality back into the realm of promoting equality. For these renewed souls this might constitute a case for supporting an equality of opportunity, albeit that these circular movements of opinion could make a case for questioning society's actual desire for equality.

Is Equality Part of Human Progress?

Any form of equality requires an objective stance in its defence if human progress is to be truly representative. The liberty principle (Rawls, 2003) associated with most branches of equality, it is argued, would be one certainly chosen by individuals if they were not going to know their place in society; in the words of John Rawls from the 'original position' (Ibid: 104) While under this 'veil of ignorance' (Ibid: 129) any individual would choose an outcome that would minimise personal liability, in whatever sense. No one, argues Rawls, would run the risk of possibly being given 'the life of a slave'. (Cryan and Shatil, 2009: 151) From this original standpoint equality is argued as a means to achieve greater social justice, which if absent can eventually result in a barbaric state of war of man against man. This argument, highlighting the universal threat of a destructive society over an equitable society can only be viewed from historical accounts of mass protests and demonstrations even in so called 'developed' countries. The Paris riots in 1968 only help to frame the argument supporting a morally developed country over any 'economically efficient' alternative, relying solely on GDP/GNP statistics. This in itself would constitute progress.

Robert Nozick, a Harvard colleague of Rawls, looked at liberty, or equality of choice, from a different perspective altogether. In his book *Anarchy, State and Utopia* (1974) Nozick acknowledges an individual liberty associated with the right to life, a right universally shared but individually determined. This latter point being critical for Nozick since any intervention to the natural liberty assigned to life would in itself be a constraint on freedom. What this goes on to suggest in correlated arguments is that any measure taken by a government or otherwise to ensure a distributive equality will only serve as a disincentive to be productive, since this action would appear to curtail any notion of equality of effort or contribution, under instruments like progressive taxation, as argued similarly above. Nozick's argument in fact only helps to advance the notion of equality of effort, and not equality of choice, since it fails to explain why in all developed economies the preferences of some individuals seem to matter more than the preferences of others, when it comes to economic choice. This development has to do with the dynamism of capital and the privileged position of GDP/GNP statistics, as has been discussed throughout.

Rewarding people according to sacrifice, rather than contribution or end results which are components of GDP/GNP measurement would constitute human progress since it resides in factors within the control of human beings. The question begs: do we want a morally developed society or an 'efficient' but empty economy? This being the ultimate paradox: morality and efficiency not having a common origin.

Conclusion

For any desired future society it is clear from the above portrayal that the economic model that is to be proposed must be subject to the rights and needs of humankind. A model which is built on saturating the infinite demands of output growth and materialistic desires will only result in us dealing once again with problematic notions regarding the rights of man which should already be presupposed in the development of any society. The paradox surrounding GDP/GNP growth should illustrate the dangerous road towards their approval. Between liberalism and equality, we may go through many more years in dispute until a change of mentality is finally attained. We might start by promoting equality of effort and reward individuals accordingly which would be more motivational and satisfactory for humanity at large.

One source of inspiration might come from John Maynard Keynes when he advocates, contrary to GDP/GNP demands, that we should work 'three-hour shifts or a fifteen hour week...' increasing solidarity and permitting the sharing of a production already achieved (Keynes, 1972: 329). Human progress could only then be liberated from the constraints of GDP/GNP measurement and begin its upward journey. Looking in the direction of the Caribbean might provide another source.

References

Cryan, D and Shatil, S. 2009. *Introducing Capitalism: A Graphic Guide*. London: Icon Books Ltd.

Galbraith, J. K., 1999. *The Affluent Society*, London: Penguin Books.

Hahnel, R and Albert, M. 2002a. 'In Defence of Participatory Economics'. *Science & Society* 66: 1: 7-21.

Hahnel, R and Albert, M. 2002b. 'Reply'. *Science & Society* 66: 1: 26-28.

Harman, C. 2009. *Zombie Capitalism: Global Crisis and the Relevance of Marx*. London: Bookmarks Publications.

Happy Planet Index Report 2.0, 2009: http://www.happyplanetindex.org/explore/global/

Keynes, J.M. 1972. 'Economic Possibilities for our Grandchildren'. *Essays In Persuasion: The Collected Writings of John Maynard Keynes* IX: 321-332.

Marx, K. 2008. *Capital*. Oxford: Oxford University Press.

McAleese, D. 2008. ‘Policy Objectives and Issues for a Regional Economy’ in O’Hagan, J and Newman, C (ed.) *The Economy of Ireland*. Dublin: Gill & Macmillan Ltd.

Mill, J, S. 2001. *Utilitarianism*. Indianapolis: Hackett Publishing Company.

Nozick, R. 1971. *Anarchy, State and Utopia*. Oxford: Basic Books Inc.

Ridoux, N. 2009. *Menos Es Más: Introduccion a la filosoia del decrecimiento*. Barcelona: Los Libros Del Lince.

Rousseau, J,J. 1984. *A Discourse on Inequality*. London: Penguin Books.

Rawls, J. 2003. *A Theory of Justice*. Boston: Harvard University Press.

IS THERE AN ALTERNATIVE TO THE TAYLOR RULE WHEN ESTIMATING A NEUTRAL REAL RATE OF INTEREST FOR THE EURO ZONE?

Sophie Ward

Senior Sophister

In this paper Sophie Ward seeks to find a policy neutral interest rate which may be used for policy decisions. The issue of monetary policy is currently at its most important of the last few decades. The idea of having a 'neutral rate' around which policy can be based is an interesting one. It could have use as a benchmark to determine whether a chosen policy rate is likely to be expansionary or contractionary and be a judge of the extent. It defines an innovative type of economic policy whereby base rates may be set hoping to converge to this rate for maximal price stability. This might be particularly interesting now as there has been much debate over whether lax monetary policy was partially to blame for the crisis that we are currently trying to recover from. As this essay will consider, this neutral rate is not observable, so it must be estimated.

Introduction & Motivation

> " Government actions and interventions, not any inherent failure or instability of the private economy, government actions, caused, prolonged, and worsened the crisis" - John Taylor

The neutral real interest rate is, in essence, a medium term concept that indicates the level of real interest that is analogous with monetary policy that is cycle-neutral. It seems logical that such a 'real' rate be used as rational economic agents base decisions on real variables, not nominal ones. Our task will be to break down the nominal interest rate into parts that allow this neutral real rate (NRR) to be determined. There are many factors that may affect the economy other than monetary policy.

This question is best approached by estimating a model for the neutral rate, which I shall call the 'NRR Model', and also a more general model for the determination of the nominal interest rate, which I shall call the 'Determination Model'.

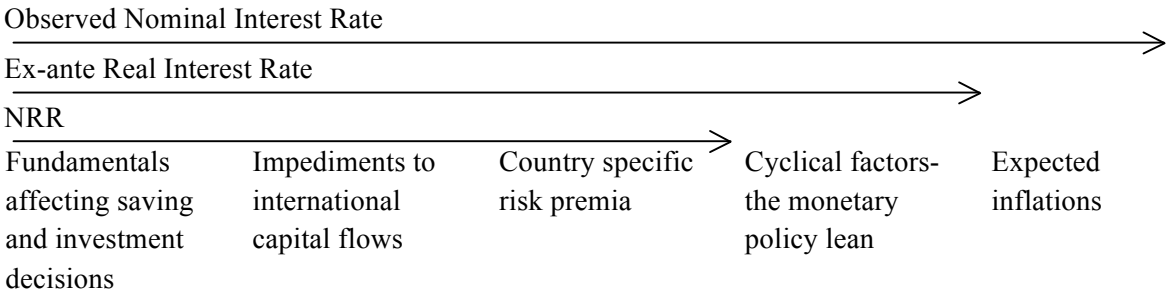
Literature thus far estimates this rate using a standard 'Taylor' framework, and for single countries, rather than a zone of monetary union such as the European Union. Such papers relate especially to the Federal Reserve, the Bank of Canada and the Reserve Bank of New Zealand.

The Theoretical Concept of NRR

The grounds of a theory begin by thinking about how we can tell if the base rate set in monetary policy is too high or too low. We have already mentioned that such a judgement may be made by virtue of a 'real interest rate' deducted for all nominal and cyclical factors. The simplest decomposition of a nominal interest rate comes from Fisher's Equation: $i = r + \pi^e$ *where i is the nominal interest rate, r is the real interest rate and* π^e *is expected inflation.*

We will find this expression to be useful in specifying our model. However, the *r* in the Fisher equation is cyclical, which will not answer our question. The nominal interest rate is one that the central bank, in this case the European Central Bank, may directly control. We will be using two nominal interest rates in our discussion, each for a different use. The three-month interbank rate will be used as a proxy for the nominal interest rate (dependent variable) in our Determination Model. The ECB's base money market interest rate will act as the dependent variable in the NRR Model. This rate is used for leaning in monetary policy.

The intuition behind this research is that the nominal interest rate may be decomposed thus:



Our aim will be to estimate the NRR by finding variables to fill the 'cyclical factors' mentioned in the scheme above.

Variables and Empirics

We estimate Taylor's Rule for the euro zone before trying to estimate a more 'all-encompassing' model for the NRR. Taylor estimated, without the use of econometrics, the following: $i = 2.5 + \pi + 0.5(\pi - \pi^*) + 0.5(y - y^*)$ *where 2.5 is the NRR,* $(\pi - \pi^*)$ *is the inflation gap between*

actual inflation and the target inflation rate set by monetary policy and (y – y) is the output gap between actual output and potential output.*

Given that the ECB is known to be hawkish in its monetary policy we will expect the coefficient on the inflation gap to be higher than 0.5.

Models: In estimating the NRR Model, we will consider the following:

$$i = r + \beta_1(INFGAP) + \beta_2(GDPGAP) + \beta_3(STOXX) + \beta_4(SDR) + \beta_5(EMP) + u_i$$

Where i is the base rate, r is the NRR, INFGAP is the inflation expectation gap, GDP is the GDP growth gap, STOXX is a proxy for the performance of financial markets and SDR is a proxy for the exchange rate.

The mean of the constant and residual that result from such a regression may be taken as an indication of the neutral real rate. This is a reasonable assumption to make given that we are effectively stripping out the 'scheme-superfluous' nominal and cyclical components. The remainder of the model should be 'unexplained', and thus an indication of the NRR. In order to form a Determination Model, we will consider this:

$$i = \beta_0 + \beta_1(INF) + \beta_2(GDP) + \beta_3(STOXX) + \beta_4(SDR) + \beta_5(EMP) + \beta_6(logBANKRATE) + u_{ii}$$

The models themselves are to be run as a longitudinal cross-section. When estimating this rate, we are not especially concerned about the impact of a time series. We are more interested in defining a post-currency-union measure of the NRR for use in judging monetary policy. The years sampled are 1999 – 2007, and only the countries in the union from 1999 are included for simplicity's sake.

Variables: Each of the variables mentioned in the models above has a related discussion.

i (Bank Rate) - This is the rate that the ECB sets as its baseline money market rate. It will be useful in the NRR Model.

i (Nominal Rate) - This is the three month interbank rate for respective countries. It can be used when a Determination model is being estimated.

INF - This is the inflation for each country of the euro zone in terms of the HCIP, the harmonised consumer price index, as available from EuroStat. We would expect inflation to have an upward effect on the nominal interest rate. There should be no effect on the NRR.

INFGAP - This is the inflation expectation gap. It is taken as the difference between the actual level of inflation and the expected rate. We assume a basic 'static expectations' model, whereby the expected level of inflation for a given date is the actual inflation rate from the one before it.

GDP - This is taken as the growth in GDP from the previous year. When GDP increases we would expect the nominal rate to increase due to upward pressure on the cyclical component of the model.

GDPGAP - This is taken as the difference between the actual GDP growth rate, and an average of growth rates over the periods used.

STOXX - This is the percentage change in domestic stock indices from the previous year to get an idea of how financial markets have performed. We would expect a negative relationship between increasing stock indices and the nominal interest rate, as investment demand decreases when the interest rate increases.

SDR - This is the exchange rate against the IMF's special drawing rights. It allows a fair comparison. It is not clear whether there will be any relationship here, other than that if the exchange rate is favourable to inward capital flows we would expect the interest rate to adjust downwards in an arbitrage fashion.

EMP - This is the level of unemployment for respective countries. We would expect that if employment is high, the economy is operating with a positive output gap and that the interest rate would be raised in order to prevent over-heating in the economy.

All macroeconomic data (GDP, EMP & SDR) was obtained from the EuroStat website. The exchange rate data against special drawing rights came from the International Monetary Fund. Historical stock indices were obtained separately for each country from their respective central banks' websites and from Yahoo Finance.

Specifications and Results

The Taylor Rule was estimated first. The model that Taylor presented was not estimated econometrically, and it was found that to use both the inflation rate and the policy inflation gap was detrimental for the model owing to issues of collinearity. We find, in fact, that for the Euro Zone, Taylor's idea of the inflation gap is insignificant. This leaves us with a good lesson for producing a better NRR Model. We find that the best dependent variable is a log transformed bank set rate, given tests for heteroskedasticity.

$$i = 2.594 + 1.14(\pi - \pi^*) + 3.93(y - y^*)$$

After performing the 'mean' calculation for NRR.

Log Bank Rate raw

Inflation expectations gap	1.143	Observations	99
	(2.171)		
Output gap	3.933 ***	Adjusted R-squared	0.139
	(0.999)		

Standard errors in parenthesis. * significant at 10%; **significant at 5%; ***significant at 1%.

Given the hawkish reputation of the ECB, we are not surprised that our coefficient for the inflation gap does not match Taylor's 0.5. The coefficient for the output gap is not especially interesting given that contemporary monetary policy is not primarily focussed on GDP growth rates. This model contends that if the inflation gap were to increase by 1 unit, then we would expect the base rate to be raised by 1.14%. We do agree, however, that the NRR for the Euro Zone is around Taylor's 2.5%.

The NRR Model

Having deduced that Taylor's Rule cannot come up with a reliable indicator of the NRR for significance reasons, we attempt to find a statistically sound model for the NRR by decomposition of the cyclical gap.

The NRR Model was estimated repeatedly using both *i(Bank Rate)* and *i(Nominal Rate)* and various combinations of the explanatory variables until the best model was obtained, in terms of both the goodness of fit and significance. Each variable underwent tests for linearity and normality and it was found that decimalised rates produced the best specification for the model, and that the only variables that were better transformed were the two nominal interest rates, which were logged.

$$\text{logi(BANKRATE)} = 2.34 + 8.57(\text{INFGAP}) + 2.89(\text{GDP})$$
$$\textit{(2.49)} \qquad \textit{(0.98)}$$

Log Bank Rate raw

Inflation expectations gap	8.575***	Observations	99
	(2.491)		
GDP	2.89 ***	Adjusted R-squared	0.231
	(0.977)		

Standard errors in parenthesis.* significant at 10%; **significant at 5%; ***significant at 1%.

We learn from this model that the NRR for the Euro Zone is 2.34%. I would contend that this is a better model for the estimation of the NRR due to a higher adjusted R squared and much better significance of the variables and indeed the model as a whole. These results were confirmed by the t and F tests, with P values less than 0.05, therefore strongly rejecting the null hypothesis that coefficients are no different from zero. The coefficient for the inflation expectation gap cannot be

interpreted in the same way, to imply the level of hawkishness, as the Taylor Rule. However, we are not surprised to learn that if the inflation expectation gap increases by one unit, we expect the bank set rate to increase by 8%, in an attempt to realign expectation. It is important to note that this does not mean an 800 basis point rise in the rate, but an 8% increase on its current absolute percentage value. This seems reasonable. Again, I would not overstate the importance of the GDP growth rate, aside from the fact that we expect its increasing one unit has a 2.89% upwards effect on the nominal interest rate set.

The Determinant Model

The Determination Model for the Euro Zone was stumbled across when trying to define the NRR Model. It was found that the exchange rate and unemployment level were insignificant to the model. The following initial specification was obtained:

$$\text{logi} = -0.62 + 2.56(\text{INF}) - 0.22(\text{STOXX}) - 1.41(\text{GDP}) + 1.28(\text{logBANKRATE}) - 0.63$$
$$\textit{(0.07)} \quad \textit{(1.14)} \quad \textit{(0.04)} \quad \textit{(0.58)} \quad \textit{(0.05)} \quad \textit{(0.07)}$$

Log Nominal Rate raw

HCIP	2.561**	Observations	92
	(1.144)		
Stoxx	-0.22***	Adjusted R-squared	0.87
	(0.44)		
GDP	-1.42**		
	(0.576)		
Log Bank Rate	1.284***		
	(0.054)		

Standard errors in parenthesis.* significant at 10%; **significant at 5%; ***significant at 1%.

The intercept of this model cannot be interpreted as the NRR due to the fact that the dependent variable is no longer the bank rate. We learn that for a one-unit increase in the inflation rate, the nominal interest rate is expected to increase by 2.56%. For a one unit increase in the stock indices, the nominal interest rate trends downwards 0.2%. We are concerned to learn that for a one-unit increase in the GDP growth rate, the nominal interest rate is expected to decrease by 1.42%. If the *absolute* bank set rate increases by 1 basis point, then the nominal interest rate experiences a related increase of 128 basis points.

It was found that the model was statistically improved upon with the addition of a dummy variable, GDPINC (for GDP) that captures the upward or downward pressure on the nominal interest rate of a positive or negative output gap. If a positive output gap is present then upward pressure should be placed on the interest rate, where ‘1’ is set to be positive and ‘0’ otherwise.

Log Nominal Rate raw			
HCIP	2.048*	Observations	92
	(1.093)		
Stoxx	-0.22***	Adjusted R-squared	0.88
	(0.042)		
GDPinc	0.07***		
	(0.023)		
Log Bank Rate	1.315***		
	(0.053)		

Standard errors in parenthesis.* significant at 10%; **significant at 5%; ***significant at 1%.

Our Adjusted R squared has improved, as has the significance of the GDP variable. The direction of the GDP coefficient has improved to act with the expected sign, whereby increasing GDP (and hence a one unit positive output gap) is manifested in an increase of 0.07% of the nominal interest rate. We see that monetary policy remains the majority determinant of nominal interest rates.

Specific Analysis of the Models

We know that for an OLS model to be accepted, a number of conditions must be met and tested to follow the Gauss-Markov framework.

Linearity and Normality: The relationships between the predictors and the outcome variables should be linear. The variables and the residuals resulting from the regression should all follow normal distributions.

Linearity tests were performed on each explanatory variable. When a deviation from linearity was found, a 'gladder test' was run again to ensure that the identity variable follows a closely normal distribution. Kernel density estimates were also generated on the models.

Running Shapiro-Wilk tests for normal data for these models yield the following output:

Variable	Obs'	W	V	Z	Prob>z
r	92	0.987	0.993	-0.015	0.506

We can deduce that, since the null hypothesis cannot be rejected at the 5% level, the model must be normally distributed.

Hetroskedasticity: We require the residuals of the model to hold a homogenous variance in order to make any inferences. A graph of the residuals was plotted against fitted values and despite an ambiguous looking plot there appeared to be no convergence at either end. Statistical tests were run in order to draw a conclusion.

The NRR Model contains no heteroskedasticity, since the Breusch-Pagan test fails to reject the null hypothesis that homogenous variance exists.

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
HO: Constant Variance	
Variables: fitted values of logbankraw	
Chi2 (1) =	0.64
Prob > Chi2	0.423

White's test and the Breusch-Pagan test seem to indicate that some form of heteroskedasticity may be present in our Determinant Model:

Cameron & Trivedi's decomposition of IM-test

	Chi2	Df	P
Heteroskedasticity	28.22	13	0.008
Skewness	3.9	4	0.42
Kurtosis	0.47	1	0.494
Total	32.59	18	0.019

Breusch-Pagan / Cook-Weisberg test for Heteroskedasticity	
HO: Constant Variance	
Variables: fitted values of WHAT	
Chi2 (1) =	6.98
Prob > Chi2	0.008

We correct by running a robust regression.

Log Nominal Rate raw

HCIP	2.048**	Observations	92
	(1.032)		
Stoxx	-0.22***	Adjusted R-squared	0.88
	(0.036)		
GDPinc	0.07***		
	(0.023)		
Log Bank Rate	1.315***		
	(0.048)		

Robust standard errors in parenthesis.* significant at 10%; **significant at 5%; ***significant at 1%.

Nothing has changed in the statistics of our model, other than that the coefficient for inflation has now become significant, suggesting that this homoskedastic regression has generated a far better model for the determination of the nominal interest rate.

Colllinearity

Successful models should be free of multicollinearity, in that variables should not be correlated with one another. The presence of such a problem could lead to overinflated R squares, making the model seem better than it is. By finding the *variance inflation factor* we can deduce whether this is an issue for us in our models.

The final NRR Model yields acceptable variance inflation factors:

Variable	VIF	1/VIF
GDP	1.14	0.879
Inflation expectations gap	1.14	0.879
Mean Vif	1.14	

The final Determinant Model, post correction for heteroskedasticity, carries the follows variance inflation factors:

Variable	VIF	1/VIF
GDPinc	1.23	0.815
Log Bank Rate raw	1.2	0.832
Stoxx	1.11	0.9
HCIP	1.03	0.97
Mean VIF	1.14	

We should accept multicollinearity at a VIF threshold of 10, and an inverse threshold of 0.1. For each model, we are within both, and can deduce that no multicollinearity is present.

A weak correlation test consolidates our discussion on heteroskedasticity.

PW Correlations	Bank Rate	HCIP	Inflation ex' gap	Employment	Stoxx	GDP	SDR
Bank Rate	1.0000						

HCIP	0.1507	1.0000					
Inflation ex' gap	0.4461	0.3860	1.0000				
Employment	-0.1069	-0.2713	-0.0392	1.0000			
Stoxx	0.2530	-0.0347	0.1918	-0.0045	1.0000		
GDP	0.3983	0.2410	0.3476	-0.2988	0.3454	1.0000	
SDR	0.1598	0.0046	0.0383	0.1665	-0.983	-0.0370	1.0000
Inflation ex' gap raw	0.1507	1.0000	0.3860	-0.2713	-0.0347	0.2410	0.0046
Output gap raw	0.3983	0.2410	0.3476	-0.2988	0.3454	1.0000	-0.0370
Log Bank Rate raw	0.9948	0.1425	0.4223	-0.1096	-.2342	0.3924	0.1607
Log Nominal Rate	0.8806	0.2200	0.3047	-0.1469	-0.0153	-0.2281	0.1721

	Inflation ex' gap raw	Output gap raw	Log Bank Rate raw	Log Nominal Rate raw
Inflation ex' gap raw	1.0000			
Output gap raw	0.2410	1.0000		
Log Bank Rate raw	0.1425	0.3924		
Log Nominal Rate	0.2200	0.2281	0.9004	1.0000

Model Specification

In order to make inferences about whether the models have been correctly specified we may run a linktest, which tests the overall model specification strength. We can also run the Ramsey RESET test for omitted variables.

For the NRR Model:

Log Bank Rate raw

Hat	2.635	Observations	99
	(2.629)		
Hatsq	-0.568	Adjusted R-squared	0.23
	(0.911)		

Standard errors in parenthesis.* significant at 10%; **significant at 5%; ***significant at 1%.

Ramsey RESET test using powers of the fitted values of Log Bank Rate raw

HO: model has no omitted variables

F (3,93) =	0.27
Prob > F	0.847

For the linktest we were unable to reject the null hypothesis that the model was correctly specified, at the 5% level. For Ramsey's RESET test, we were unable to reject the null hypothesis that the model contains omitted variable bias, at the 5% level.

Tests for the Determinant Model lead to ambiguous results:

Log Nominal Rate

Hat	-0.417	Observations	92
	(1.009)		
Hatsq	-0.203	Adjusted R-squared	0.884
	(1.755)		

Standard errors in parenthesis. * significant at 10%; **significant at 5%; ***significant at 1%.

Ramsey RESET test using powers of the fitted values of logiraw

HO: model has no omitted variables

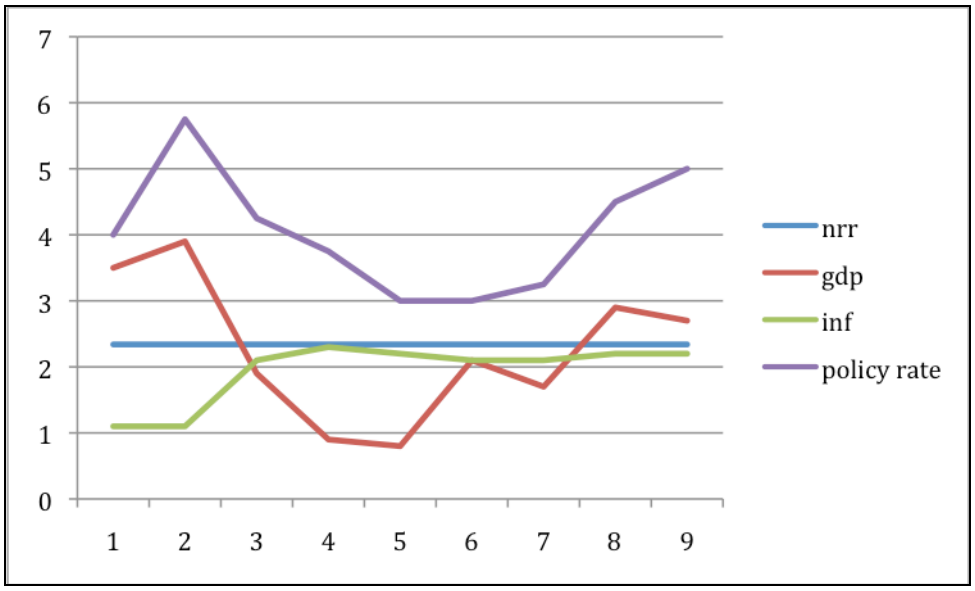
F (3,84) =	13.62
Prob > F	0.0000

It seems that the model is specified correctly, yet seems to be missing variables. We are not surprised to learn this knowing that there are many more determinants of the interest rate on top of

these. Some measure of either velocity of money in the respective economies or a measure of money supply might be very useful to this model if the question were to be asked again.

CONCLUSIONS AND EXTENSIONS

Using the models specified we have managed to deduce that the neutral real rate for the Euro Zone in the decade following monetary union on 1999 was 2.34%. The GDP growth rate for this period generally signified 'positive' output, so it is not surprising that the policy rate remained above the estimated NRR for the period as an attempt to control the possible overheating that was in progress. We can see also that the inflation rate needed control. The following plot begins with 1 = 1999, with the respective rates on the y axis.



The models that were generated have undergone a series of tests of their integrity. We have deduced that a lower R square is acceptable in our NRR Model. Through corrections for heteroskedasticity and the introduction of a dummy variable for GDP we have created a significant model with strong explanatory power for determining the nominal interest rate in the Euro Zone. The determinant model has too many variables for NRR determination, so we found the happy balance in our NRR Model where the NRR is not eaten into by non-parsimonious variables.

With further testing and data the accuracy of this research could perhaps be boosted. Firstly the assumption of static expectations is far reaching. However if inflation expectations could be generated, based upon a true maximisation problem and hence rational expectations, we would be better off.

Furthermore, I would propose that in reality, the NRR does shift over time and that a more in depth study of the period could be done by looking at the evolution of a rate over time using monthly data for each year.

References

Archibald, J. and Hunter, L. 2001. 'What is the Neutral Interest Rate and How Can We Use It?'. Federal Reserve Bank of New Zealand. *The Reserve Bank of New Zealand Bulletin*, 64(3).

Basdevant, O. Bjorksten, N. and Karagedikli, O., 2004. 'Estimating a Time Varying Neutral Real Interest Rate for New Zealand'. Federal Reserve Bank of New Zealand, *Discussion Paper Series.*

Heinemann, F. and Ullrich, K., 2006. 'The Impact of EMU on Inflation Expectations'. *Open Economy Review*, 17, pp.175-195.

Gujarati, D. 2009. *Basic Econometrics,* 5th ed. McGraw-Hill Higher Education, New York.

Lam, J. and Tkacz, G. 2004. 'Estimating Policy-Neutral Interest Rates for Canada Using a Dynamic Stochastic General-Equilibrium Framework'. *Swiss Journal of Economics and Statistics*, 140(1), pp.89-126.

Shostak, F., 2005. 'The Myth of the Neutral Interest Rate Policy', *Brookes News,* [Online]. Available at: http://www.brookesnews.com/051403interest.html

Trichet, J. C. et al., 2004. 'The Monetary Policy of the ECB'. *The European Central Bank Publication.*

Wooldridge, J., 2008. *Introductory Econometrics: A Modern Approach.* United States: South Western College Pub.

Data Sources

EuroStat Statistics: http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/

IMF Statistics: http://www.imfstatistics.org/imf/

World Bank Country Data: http://web.worldbank.org/WBSITE/EXTERNAL/DATASTATISTICS/0,,menuPK:232599~pagePK:64133170~piPK:64133498~theSitePK:239419,00.html

Yahoo Finance

DORNBUSCH'S OVERSHOOTING MODEL: A REVIEW

Christoph Walsh

Senior Sophister

Dornbusch's influential Overshooting Model aims to explain why floating exchange rates have such a high variance. Christoph Walsh provides an extremely well researched account of the model in detail, while examining the empirical evidence for uncovered interest rate parity and purchasing power parity.

Introduction

Dornbusch's (1976) overshooting model was path-breaking, used not only to describe exchange rate overshooting but also the 'Dutch disease', exchange rate regime choice and commodity price volatility. Dornbusch's model was highly influential because, at the time of writing, the world had only recently switched from the Bretton Woods system to flexible exchange rates and very little was understood of them and their volatility. In a study on graduate courses in international finance, Dornbusch's article was the only article to make it on to more than half of the reading lists – and it made it on to every one (Rogoff, 2002)!

The first section of this essay will describe Dornbusch's model in detail. The second section will then give empirical evidence on the two main assumptions of the Dornbusch model: UIP and long-run PPP, and will review the findings for and against the Dornbusch model in general.

The Model

Dornbusch's model makes the following assumptions:

Uncovered interest-rate parity (UIP) states that the interest rate differential is equal to the expected change in the spot rate: $i_t - i_t^* = E_t\left[\Delta s_{t+1}\right]$, *where* i_t *and* i_t**are the domestic and foreign interest rates at time t respectively and* $E_t[\Delta s_{t+1}]$ *is the expected change in the logarithm of the exchange rate*[1]. The model assumes that UIP holds at all times.

Note that in the following a rise in S_t is denoted as a depreciation of the exchange and a fall in S_t is an appreciation.

1 Referenced with permission.

UIP assumes that capital is perfectly mobile, investors are risk neutral and domestic and foreign assets are perfect substitutes. The model assumes that the capital market adjusts instantaneously to shocks and that investors have rational expectations (Arnold, 2009).

Absolute purchasing power parity (PPP) states that a basket of goods, when expressed in a common currency, should have the same cost across countries. PPP states that the real exchange rate, defined by $Q_t = S_t P_t^*/P_t$ is equal to unity, *where S_t is the spot rate and P_t and P_t^* are the domestic and foreign price levels respectively at time t.* Setting $Q_t = 1$ gives: $P_t = SP_t^*$. The price levels, expressed in a common currency, are equal (Pilbeam, 2006).

The model assumes that deviations from PPP are permitted in the short run, but the exchange rate will return to PPP in the long run (Arnold, 2009). In the short run, the real exchange rate can rise which increases the competitiveness of the economy and increases net exports (Frankel, 1979).

The assumption of long-run PPP is made because prices are 'sticky' in the short run due to nominal and real rigidities as in Keynesian theory (Romer, 2006). A main feature of the model is that it allows a distinction between sluggishly-adjusting goods markets and hyperactive asset markets (Rogoff, 2002).

The money market is in equilibrium when real money supply equals real money demand where real money demand is a rising function of output and a falling function of the interest rate: $M_s^t/P_t = L(Y_t, i_t)$ (Arnold, 2009).

Output is given by the standard IS curve, which is rising in the real exchange rate and falling in the interest rate (Copeland, 2008).

The inflation rate is a function of the output gap, given by the Phillips curve: $\Delta P_{t+1} = \psi(y_t$ - ybar) *where Δp_{t+1} is the change in the logarithm of the price level, y_t and ybar* are the logarithms of output and potential output respectively and ψ is a measure of the price flexibility in the economy (Copeland, 2008).

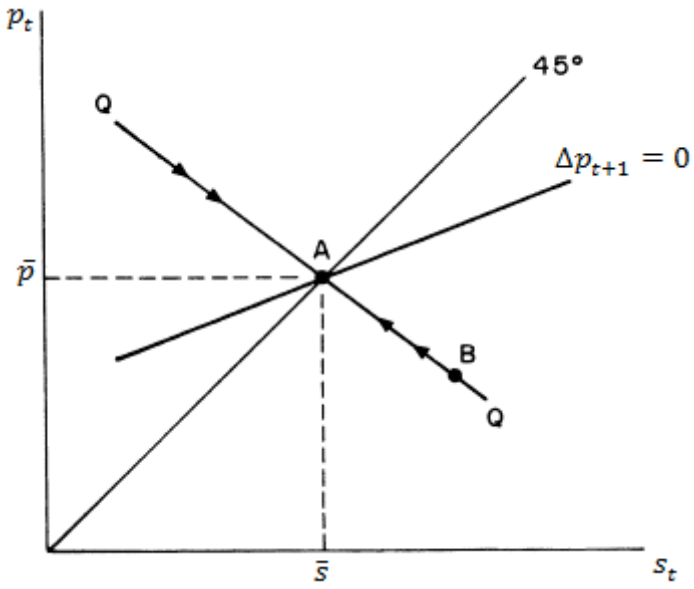
The economy under scrutiny is a small open economy so that it cannot affect foreign interest rates, prices and output.

The economy begins at a stationary state where:

(i) output is at potential,
(ii) PPP holds,
(iii) domestic and foreign interest rates are equal,
(iv) prices are constant, and
(v) the exchange rate is at its equilibrium level.

The stationary state, (Sbar, Pbar) is shown by point A in the Figure 1 below.

Figure 1



(Dornbusch, 1976)

The 45° line represents equilibrium PPP. Below the line, the economy is more competitive and net exports will rise. Above the line, the opposite is true.

The $\Delta p_{t+1} = 0$ line is the goods market when output is at potential. Below the line output is above potential and there is an upward pressure on prices (Pilbeam, 2006).

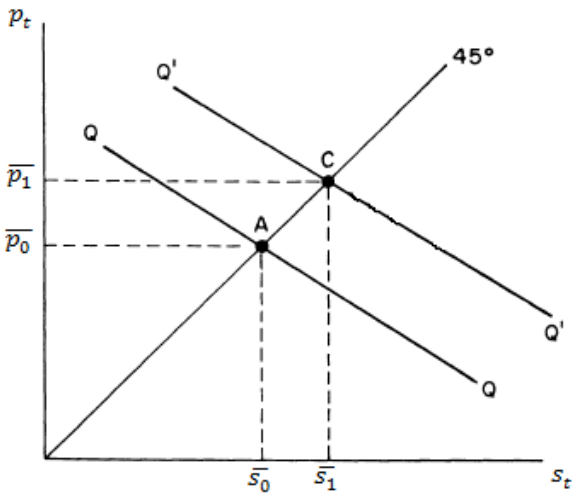
The QQ schedule combines money market equilibrium and UIP. At point A, the money market is clear, UIP holds and the domestic and foreign interest rates are equal. As the money market clears instantly and UIP holds at all times, the economy must be on the QQ schedule at all times. Point A is the model's stationary state (Dornbusch, 1976).

The model analyses the dynamics of the economy after an unanticipated rise in the money supply. It is convenient to first analyse what will happen in the long run after the change in money supply before studying the short-run dynamics.

The Long Run

By the quantity theory of money, a 1% rise in the money supply leads to a 1% rise in the price level. All else constant, a 1% rise in the price level must be matched by a 1% depreciation in the exchange rate to maintain PPP. This is as in the monetary model (Rosenberg, 1996). The rise in the money supply is given by a shift of the QQ schedule to Q'Q'. The new stationary state, is given by point C in Figure 2.

Figure 2

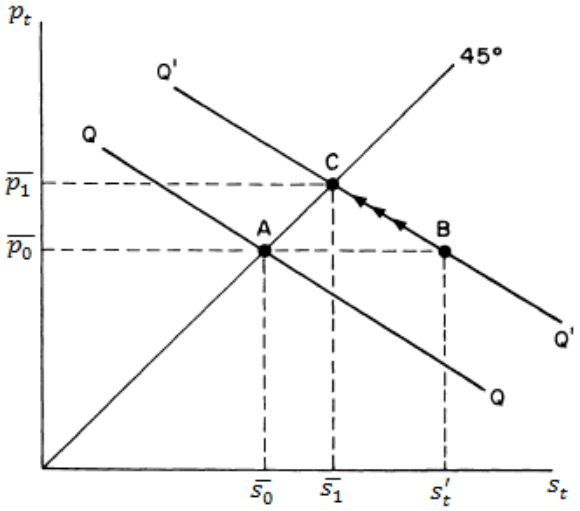


(Dornbusch, 1976)

The Short Run

After the increase in money supply, the money market is in disequilibrium. The interest rate falls by the liquidity effect to increase money demand. As $i_t < i_t^*$ foreign assets have become more attractive to investors. By UIP, individuals will only invest in domestic assets if they expect an appreciation in the exchange rate. But in the long run, as shown above, the exchange rate must depreciate overall. The exchange rate must therefore depreciate so much after the shock that it 'overshoots' its long-run equilibrium level and appreciate thereafter. The exchange rate jumps to point B in Figure 3.

Figure 3

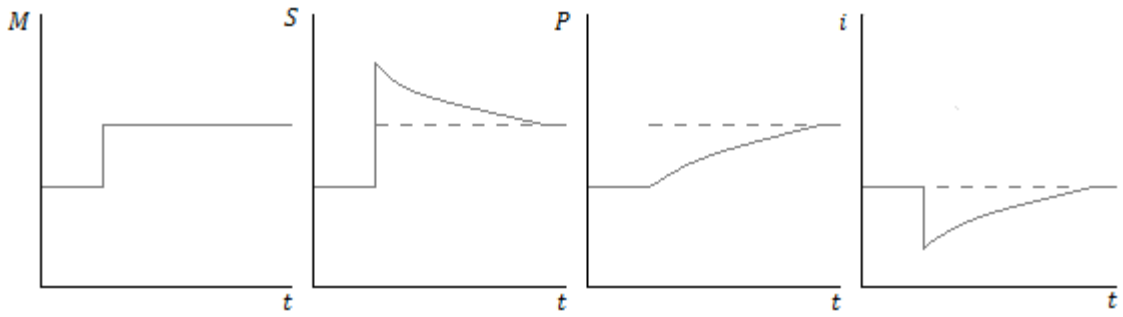


(Dornbusch, 1976)

Because of a rise in the real exchange rate and a fall in the interest rate, output is above potential. By the Phillip's curve, this gives an upward pressure on prices. The gradual appreciation of the exchange rate and inflation will lower the economy's competitiveness until it reaches the new stationary state C in Figure 3. Also, the fall in real money balances (due to the rising price level) will raise the interest rate back to its original level of $i_t = i_t^*$ so there are no more expected changes in the exchange rate (Copeland, 2008; Dornbusch, 1976).

The time paths of the money supply, the exchange rate, the price level and the domestic interest rate are shown below in Figure 4.

Figure 4



The extent of the overshoot depends on the interest-sensitivity of the demand for money and how sensitive the market is about over-valuations and under-valuations of the currency (Dornbusch, 1976; Copeland 2008).

Wilson (1979) analyses the case where the money supply shock is anticipated, rather than unanticipated in Dornbusch (1976). He demonstrates that the announcement or expectation of an expansionary policy alone will cause the exchange rate to jump – even before the policy is implemented.

Empirical Evidence

Evidence on UIP: Fama's beta coefficient, β, describes the relationship between the percentage change in the spot rate and the one period lagged value of the interest rate differential between two countries (Davis, Miller and Prodan, 2009): $\Delta S_t = \beta(i_{t-1} - i^*_{t-1})$

Using OLS, β would be expected to equal plus unity, which would be consistent with UIP. However, empirical tests usually generate negative β coefficients. Froot and Thaler (1990) found that in over 75 published estimates, the average value was -0.88. Some estimates were positive but not one was equal to or greater than plus unity. This is called the *UIP puzzle*. These findings undermine the credibility of the Dornbusch model.

Fama (1984) suggests that there may be an omitted variable in the equation. If this is so, because of omitted variable bias, β will be biased away from plus unity. The omitted variable could be exchange rate prediction errors or a risk premium. UIP assumes risk-neutrality when in fact investors are generally risk averse and demand a risk premium.

UIP is tied to foreign exchange (FX) speculation in the carry trade. It is suggested that either the volume of carry trade is insufficient to generate UIP or that the risk premium associated with carry trade is very large. Excess returns from the carry trade are found to range from 5% to 6.5% (Davis, Miller and Prodan, 2009). Scholl and Uhlig (2008) measured the Sharpe ratio for speculating on violations of UIP and found that it lay between 1 and 1.5, showing large excess returns in the carry trade. Other suggestions are that there are significant transaction costs and shortages of liquidity in the carry trade which can lead to deviations from UIP. The puzzle may also exist due to central bank reaction functions, which create a simultaneous equation bias when estimating β (Davis, Miller and Prodan, 2009).

Evidence on PPP

Views on PPP have changed over time. Before the breakdown of Bretton Woods, economists clung to PPP's existence (Isard, 1995). In the first phase of research on PPP in the 1970s and 1980s, the emerging consensus was that the real exchange rate followed a random walk, even in the long run, thus denying the existence of PPP. Over the past decade however, more advanced econometric methods have given new findings on PPP (Copeland, 2008). The main consensus now is that long-run PPP has some validity and holds better than in the short run (Sarno and Taylor, 2002). For example, Lothian (1997) used panel data of 23 OECD countries over 1974-91 and found that, although there were large and prolonged deviations from PPP in the short run, he couldn't reject it over the long run of 3 to 6 years. This supports the assumption of PPP in the Dornbusch model.

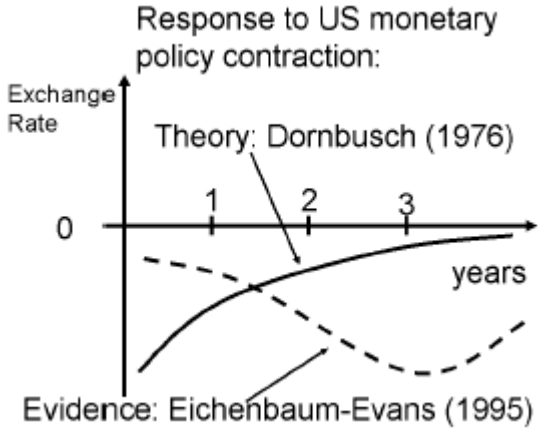
However, much of the other evidence on PPP is disenchanting. The *PPP puzzle* is that the real exchange rate does not equal unity across countries (Rogoff, 1996). Rogoff (1996) explains this by noting that exchange rates are highly volatile and the prices of goods are rigid. A deviation of the real exchange rate from unity has a half-life of 3-5 years. For example, if the real exchange rate jumped from 1 to 1.5, it would take 3-5 years for it to return to 1.25. Rogoff (1996) identifies the frictions in the goods market as transport costs, information costs, tariffs and the lack of labour mobility. Other reasons identified for the poor performance of PPP are imperfect competition, statistical problems in measurement, productivity differentials and home bias in trade (Arnold, 2009; Obstfeld and Rogoff, 2000; Pilbeam, 2006).

Evidence on the Dornbusch Model

Eichenbaum and Evans (1995) find that in studying the US economy in the period 1974-1990 with a value at risk (VAR) based identification scheme, contractionary monetary policy shocks are followed by a sharp rise in interest rates and lead to persistent, significant appreciations in the nominal and real exchange rate. The maximal effect of the shocks was not contemporaneous – the dollar continued to appreciate for a substantial period of time. This is inconsistent with the Dornbusch model which predicts that the exchange rate jumps instantaneously. This is called the *delayed overshooting puzzle*. They also found significant and persistent deviations from UIP: the returns on higher US interest rates were magnified by the future expected appreciations in the dollar.

Figure 5 below shows the differences in the time path of the exchange rate in Dornbusch's overshooting theory and evidence found by Eichenbaum and Evans (1995) which shows 'delayed overshooting'.

Figure 5



(Scholl and Uhlig, 2008)

Heinlein and Krolzig (2010) studied the *delayed overshooting puzzle* in the $/£ exchange rate using VAR over 1972-2009. They found strong evidence of delayed overshooting and a violation of UIP as a result of excess returns due to delayed overshooting. Exchange rate jumps after monetary shocks are only significant at 10% and, with 95% certainty, the jumps are not large enough for UIP to hold.

Scholl and Uhlig (2008) also investigated the *delayed overshooting puzzle* for the non-US G-7 countries over 1977-2001. They use a VAR procedure, imposing sign restrictions on the impulse responses for key monetary variables, for example, by assuming contractionary monetary policy shocks do not lead to decreases in domestic short-term interest rates. This narrowed down the range of possible shocks considerably. They find that after a monetary contraction, there is a persistent appreciation for periods of up to three years, thus finding robust evidence of delayed overshooting.

Bjørnland (2009) criticises the findings of Scholl and Uhlig (2008). He argues that they disregard the strong contemporaneous interaction between monetary policy and exchange rate movements and notes that they place zero restrictions on them. He finds that, after imposing a long-run neutrality restriction on the real exchange rate (long-run PPP), the puzzle disappears. The maximal impact of overshooting occurs in the first two quarters and his findings are also consistent with UIP. He made a study on four small open economies: Australia, Canada, New Zealand and Sweden over the period 1983-2004. He only used data from periods with less volatility and shocks and where the economy was open.

Pippenger1 (2009) finds that evidence of overshooting in the literature inspired by Eichenbaum and Evans (1995), such as Scholl and Uhlig (2008), is flawed. Eichenbaum and Evans (1995) did not claim to find 'delayed overshooting', as is widely interpreted. Rather, their impulse response functions suggest a gradual response to monetary shocks or 'undershooting'. The problem with the impulse response functions in the literature is that they describe impulse responses to 'innovations' in a variable rather than the variable itself. His findings are that since in the absence of intervention, exchange rates are essentially martingales (which is consistent with efficient FX markets), it does not seem appropriate to interpret the impulse response functions in the literature as clear evidence of overshooting.

Voss and Willard (2009) studied the US and Australia with a structural VAR model over the period 1984-2007. They examined the behaviour of the Australian exchange rate with monetary policy. They find that Australian monetary policy innovations give rise to an exchange rate response as expected under the Dornbusch model; the exchange rate changed immediately and slowly returned to its equilibrium level. They also did not observe significant deviations from UIP during the adjustment. They found, however, that changes in the US interest rate have no significant effect on the exchange rate and significant deviations from UIP were found. This is because Australian interest rates moved in tandem with US interest rates, perhaps because Australia is a small open economy while the US is not.

Frenkel (1976) examined the determinants of the exchange rate during the German hyperinflation of 1922-1923. The disturbances during this period were clearly monetary and dominate any other disturbances, thus it is possible to examine the effect of monetary variables in virtually complete isolation. He regressed the rate of change of the exchange rate against current and lagged values of the rates of change of the money supply. He found the elasticity of the exchange rate with respect to the money supply exceeded unity (1.57); this magnification effect of money on the exchange rate is consistent with overshooting. However, that the model fits in an environment of

---

Referenced with permission.

hyperinflation does not imply it will fit in an environment of stable prices. In a hyperinflation, the liquidity effect tends to break down and the Fisher effect dominates, even in the very short run.

Meese and Rogoff (1983) found that a random walk model performs as well as any estimated exchange rate model. The Dornbusch model was one of the models tested. They made out-of-sample forecasts with OLS over one-month, three-month, six-month and twelve-month periods for various currency pairs, using GLS, instrumental variables and lag specifications where appropriate. Rolling regressions were used to update the parameters each period. They compared the models using root mean squared error (RMSE). None of the models achieved a significantly lower RMSE than the random walk model at any horizon, even though forecasts were based on actual realised values of explanatory variables. Different estimations using first differences, GLS, corrections for serial correlation or different proxies for money supply and expected inflation didn't give significantly better results. They give possible reasons for the poor out-of-sample fit as sampling errors, stochastic movements in the true underlying parameters, misspecifications, OVB, or possible non-linearities.

Meese and Rogoff (1988) found that the hypothesis of 'real shocks' in technology like those in real business cycle theory are more consistent in their findings in explaining exchange rate volatility than monetary shocks as in the Dornbusch model.

Conclusion

While Dornbusch's model is highly elegant (Rogoff, 2002) the empirical evidence on the model is mixed. While there is some evidence of long-run PPP, UIP is rejected by most academics. Academics' empirical results on the model range from overshooting, delayed overshooting, undershooting to no overshooting whatsoever. So while being a highly influential model, it does not serve to give accurate forecasts of exchange rate movements.

References

Arnold, L., 2010. 'International Finance'. Regensburg University, [Online]. Available at: http://www-wiwi-cms.uni-regensburg.de/images/institute/vwl/arnold/IFSkript.pdf

Bjørnland, H., 2009. 'Monetary Policy and Exchange Rate Overshooting: Dornbusch was right after all'. *Journal of International Economics*, 79, pp.64-77.

Copeland, L., 2008. *Exchange Rates and International Finance*. 5th ed. Essex: Pearson Education Limited.

Davis, G., Miller, N. and Prodan, R., 2009. 'Toward a Solution to the Uncovered Interest Rate Parity Puzzle'. Miami University, Oxford, OH. *Working Paper*.

Dornbusch, R., 1976. 'Expectations and Exchange Rate Dynamics'. *Journal of Political Economy*, 84, pp.1161-1176.

Eichenbaum, M. and Evans C., 1995. 'Some Empirical Evidence on the Effects of Shocks to Monetary Policy on Exchange Rates'. *Quarterly Journal of Economics*, 110(4), pp. 975-1009.

Fama, E., 1984. 'Forward and Spot Exchange Rates'. *Journal of Monetary Economics*, 14(3), pp.319-338.

Frankel, J., 1979. 'On the Mark: A Theory of Floating Exchange Rates Based on Real Interest Differentials'. *American Economic Review*, 69(4), pp.610-620.

Frenkel, J., 1976. 'A Monetary Approach to the Exchange Rate: Doctrinal Aspects and Empirical Evidence'. *Scandinavian Journal of Economics*, 78(2), pp.200-224.

Froot, K. and Thaler, R., 1990. 'Anomalies: Foreign Exchange'. *Journal of Economic Perspectives*, 2(2), pp.179-192.

Heinlein, R. and Krolzig, H., 2010. 'Effects of monetary policy on the $/£ exchange rate. Is there a 'delayed overshooting puzzle'?'. School of Economics, University of Kent, Canterbury.

Isard, P., 1995. *Exchange Rate Economics*. Cambridge: Cambridge University Press.

Lothian, J., 1997. 'Multi-Country Evidence on the Behaviour of Purchasing Power Parity Under the Current Float'. *Journal of International Money and Finance*. 16(1), pp. 19-35.

Meese, R. and Rogoff K., 1983. 'Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?'. *Journal of International Economic*, 14, pp.3-24.

Meese, R. and Rogoff K., 1988. 'Was It Real? The Exchange Rate-Interest Differential Relation over the Modern Floating-Rate Period'. *Journal of Finance*, (43)4, pp.933-948.

Obstfeld, M. and Rogoff, K., 2000. 'The Six Major Puzzles in International Macroeconomics: Is There a Common Cause?'. National Bureau of Economic Research, *Macroeconomics Journal*, 15, pp.339-390.

Pilbeam, K., 2006. *International Finance*. 3rd ed. Hampshire: Palgrave MacMillan.

Pippenger, J., 2009. 'Dornbusch Was Wrong: There is no Convincing Evidence of Overshooting, Delayed or Otherwise'. Departmental Working Papers, Department of Economics, UC Santa Barbara, *Working Paper*.

Rogoff, K., 1996. 'The Purchasing Power Parity Puzzle'. *Journal of Economic Literature,* 34(2), pp. 647-668.

Rogoff, K., 2002. 'Dornbusch's Overshooting Model After Twenty-Five Years'. In: IMF (International Monetary Fund), *Second Annual Research Conference: Mundell-Fleming Lecture*.Washington, DC.

Romer, D., 2006. *Advanced Macroeconomics,* 3rd ed. New York: McGraw-Hill/Irwin.

Rosenberg, R., 1996. 'Currency Forecasting'. USA: Times Mirror Higher Education Group, Inc.

Scholl, A., and Uhlig, H., 2008. 'New Evidence on the Puzzles: Results from Agnostic Identification on Monetary Policy and Exchange Rates'. *Journal of International Economics*,76(1), pp.1-13.

Sarno, L. and Taylor, M., 2002. 'Purchasing Power Parity and the Real Exchange Rate'. IMF, *Staff Papers*, 49, pp.65-105.

Voss, G. and Willard, L., 2009. 'Monetary Policy and the Exchange Rate: Evidence from a Two-Country Model'. *Journal of Macroeconomics*, 31, pp.708-720.

Wilson, C., 1979. 'Anticipated Shocks and Exchange Rate Dynamics'. *Journal of Political Economy*, 87(3), pp.639-647.

A SWEET DEAL: SUGAR PLANTATIONS IN THE HAWAIIAN ECONOMY

Jason Somerville

Senior Sophister

Hawaii was once one of the world's most important sources of sugar – a fascinating and unusual part of the world. Jason Somerville's extensive research examines how this highly specialised economy developed – from its first contacts with the West to its accession into the United States – and addresses the famous Lucas Paradox of why capital fails to flow from rich countries to poor.

Introduction

The defining characteristic of the Hawaiian economy since western contact has been overspecialisation. Initially, sandalwood, Hawaii's naturally most abundant resource, became the backbone of the economy. As supplies diminished and whaling spread to the Pacific, Hawaiian output was largely reliant on docked whalers. The discovery of petroleum sent whaling into decline, only to be replaced by a more dominant industry; sugar. Pineapple too has played a crucial role in the overspecialisation of the Hawaiian economy, but as Hitch (1992, p.108) notes: 'Pineapple was never the symbol of concentrated economic, political, and social power in Hawaii that sugar was'. Therefore, this essay will explore the emergence of the sugar industry and its place within the Hawaiian economy. Furthermore, the reason behind the absence of capital flows to this industry during the late 19th century will be explored.

The Rise of Sugar

Sugarcane had been grown in Hawaii for some time. The Polynesians grew the crop for centuries, but the first commercial attempt at growing sugar came in 1835 when Ladd & Co., an American merchant house in Honolulu, leased several thousand acres of land from the King which was then planted in sugar, coffee, banana and taro (Cushing, 1985). This was a huge operation and so considerable labour was required. However, hiring labour was the biggest difficulty the company faced, although the pay was reasonable, as the local chiefs refused to allow their tenants to work on the plantation; eventually, the King was forced to intervene (Hitch, 1992). Once labour was acquired, the plantation owners realised that native labourers were completely unaccustomed to the type of

work that was required. To add to the difficulties, they lacked the required tools and animals; the local chiefs too blocked their efforts (Morgan, 1948).

Despite these difficulties, the plantation survived and 2 tons of sugar was harvested in its first year, rising to 4 tons in the second. However, all the inputs in the economy were owned by the King so industry remained small. Reform came with the 1839 Bill of Rights and the introduction of a Western-style constitution in 1840. With the Great Mahele of 1848, commoners gained not only voice but land rights. With private property rights established, capital could now be invested in with confidence.

With these political reforms, the sugar industry took off. Two other developments also spurred this growth. Firstly, California's gold rush led to a population explosion on the west coast of America. This created a whole new market for Hawaiian sugar. Production expanded rapidly, with output rising to 700 tons per year by the 1850s (Kuykendall, 1953). Secondly, the US civil war cut the North off from the South's sugar, providing a huge boost to the Hawaiian industry. Sugar could now be sold to an open, close-by, war-inflated economy. The price of sugar soared from 6c a pound to 17c a pound in 1864 and by 1870 the Kingdom was producing 9,000 tons annually (Kuykendall, 1953). Coffee, rice and salt output also grew, but sugar out-stripped them all. By the 1870s, the Hawaiian economy had become westernised and the sugar industry was now the driving force in society.

The Revolution of 1893

The Reciprocity Treaty of 1875 between Hawaii and the US greatly expanded Hawaiian exports of sugar and rice to the US as well as Hawaiian imports of US manufacturing. Sugar exports to the US rose from 21 million pounds in 1876 to 114 million pounds in 1883 and to 224.5 million pounds in 1890 (La Croix and Grady, 1997). This meant that the US was losing out on over five times the revenue from tariffs when it came to renegotiating the Treaty. By 1883, groups were already organising in the US to abrogate or modify it. Indeed, La Croix and Grady (1997) estimate the loss to the US in tariff income was $2.5 million in 1882.

The Treaty also generated problematic strategic dynamics, as the structural change it launched gradually worsened Hawaii's bargaining position with the US. The Hawaiian economy had become dependent on the US market, whereas the Hawaiian market represented such a negligible share of US exports that the incentives to renew the Treaty were heavily biased; indeed, when it expired in 1883, the US extracted better terms.

The Treaty also transformed Hawaii's internal politics by massively increasing the wealth of the Caucasian owners of sugar plantations (La Croix and Grady, 1997). The planters had clear incentives to support the treaty's successful renegotiation, whereas most native Hawaiians, as well as the King, were adamantly opposed to further economic concessions. Tensions came to a head in 1887 when the government's opponents (chiefly Caucasian industrialists) forced the King to accept a new cabinet and a constitution, stripping the monarchy of much of its authority. A special election in

September 1887 brought a pro-American government to power. Consequently, the US was given rights to Pearl Harbour and the Reciprocity Treaty was renewed in 1887.

Just when the plantation owners were confident that their sugar would enter the US market tariff-free, the McKinley tariff was introduced, erasing the advantages that reciprocity provided to Hawaiian sugar producers over other foreign competitors selling in the US market. This laid the incentive to eliminate the uncertainty of reciprocity through annexation. At the same time, the new Queen, Kamaka'eha Lili'uokalan, was trying to restore power to the monarchy.

U.S. opportunism in 1890 with respect to the trade treaty threatened the wealth of the planters and thereby played a critical role in the overthrow of the Hawaiian monarchy (Kuykendall, 1967). Silva (2004) presents the 'Petition Against Annexation' that was signed by 21,269 native Hawaiian people, or more than half the 39,000 native Hawaiians and mixed-blood persons reported by the Hawaiian Commission census for the same year. However, by this stage the sugar interest groups were too powerful. They solicited the US support, the Queen yielded and in 1898 the islands were formally annexed.

The Plantation Economy (1876 – 1941)

Initially, after the Reciprocity Treaty came into effect in 1876 the vast US market was opened for duty-free entry of sugar and other agricultural products. The American connection propelled Hawaii into a period of economic boom. By the time the US entered WWII in 1941 Hawaii had developed into one of the biggest plantation economies in the world, producing 4% of global sugar supply. Overall, Hawaii grew at three times the pace of the US, which is phenomenal given the speed of economic expansion on the mainland. By 1940, Hawaii's GDP per capita was approximately equal to that of the US. This extraordinary rise in income and productivity was almost entirely the result of the sugar and pineapple plantations.

Sugar was already the main crop in Hawaii in 1876, but the Treaty led to the sustained expansion of the product. In 1876, sugar was sold for 6.5 cents an ounce in the US, 3 cents of which was duty. The removal of this tariff almost doubled revenues. Margins of this magnitude stimulated a rapid expansion of sugar acreage which rose from 15,000 acres in 1876 to 238,000 acres in 1941 (Hitch, 1992). And with higher productivity, output of raw sugar rose from 13,000 tonnes to 947,000 tonnes, 73-fold, over the same period. By 1941, one in three employees in the workforce in Hawaii was employed in the sugar or pineapple industry. The combined revenues of the sugar and pineapple companies were $135 million, compared with combined revenues of $35 million for the rest of the agriculture and manufacturing industries (Schmitt, 1968). Only the expansion of the military in the late 1930s kept the economy from being absorbed entirely by sugar and pineapple.

Sugar Revisited

Sugar production requires a lot of labour (to plant and mill) and water (which required irrigation systems). Both these factors gave a competitive advantage to large plantations. Also, for mills to be cost efficient, they must be large. Economies of scale dictated that plantations become larger and larger. In 1880 the average sugar plantation was 413 acres. This grew by 1000 acres a decade so that by 1940, the average size of the 38 plantations was over 6000 acres. Improvements in transport, first by railroad and then by truck, greatly reduced transport costs, again lowering the costs of production for sugar plantations. Keeping average costs lower required a lot more work than merely expanding output and exogenous declines in transport costs.

The first major obstacle the plantations faced was the chronic shortage of labour supply. The total number of immigrant labourers recruited to Hawaii during this period has been estimated at around 300,000 (Platt, 1950). There were three distinct periods of immigration in Hawaii. The first stretched from 1851 to 1885 when some 30,000, mostly Chinese, labourers came to the islands. Between 1885 and 1908 most of the 140,000 immigrants that came were Japanese, though the Chinese still came in vast numbers until 1898, when the US' introduction of the Chinese Exclusion Act prohibited the Chinese from entering Hawaii. Here we can see the double-edged sword that was annexation for the sugar elite. While it reduced tariffs to the US, it cut off the main source of cheap labour to the industry.

The nature of the labour contract changed when Hawaii became a territory of the US. The 13th Amendment to the Constitution prohibited any form of servitude, thereby invalidating the standard Hawaiian labour contracts that required a multi-year period of service. To emphasis the ban, Congress passed in 1900 an act specifically prohibiting the importation into Hawaii of foreigners under contract agreement to perform labour. Almost half of the 350,000 workers that came in were imported under these free labour contracts. The Hawaii immigration board therefore tried to focus on Caucasian workers, but with only limited success. Because federal law barred the importation of Chinese and Japanese labour, that left the Philippines as 'the only source of a permanent labour supply' (Platt, 1950, p.11). Between 1909 and 1934, the Hawaiian Sugar Planters' Association (HSPA) brought 118,449 Filipino workers to Hawaii (the Philippines Independence Act in 1934 stopped further Filipino migration to the US). In addition, large numbers of workers came on their own initiative in search of unskilled jobs. Overall, plantation labour recruitment was a success. By 1928 there were more people on sugar plantations' payrolls than there had been residents 50 years earlier. However, by 1934, the cheap labour supply had been exhausted. Not surprisingly, as labour became scarcer, it became more expensive. The average daily wage for sugar workers in 1835 was 12.5 cents. By 1874 that had risen to 45 cent a day and on the eve of WWII in 1941 it had shot up to $2.18. Moreover, pay levels for sugar workers had largely converged to US levels prior to Hawaii becoming a territory in 1898. In 1866 pay was at 60% of the US wage, but by 1890, it had risen to 85% (Schmitt, 1968). To afford such a wage Hawaii had to have the most efficient sugar industry in the world.

Productivity in the Sugar Industry

When Ladd & Co. planted 25 acres in 1835, they produced about one-twelfth of a ton per acre, or about one ton per 100 workers. By the early 1990s, this had soared to 5.5 tons per acre and 143 tons per employee. The sugar industry certainly benefited from the advancement in technology spurred on by the Industrial Revolution. Irrigation, planting cycle developments and falling transport costs were particularly beneficial to the industry. This outstanding increase in productivity was the result of an innovative, scientific approach to farming that overcame the high costs of land, labour and transportation that prevailed in the industry. Most sugar cane at this time was heavily subsidised by governments to provide jobs and secure foreign exchange. However the US did not subsidise the sugar industry in this way, so its survival depended on high levels of productivity.

The industry became consolidated under the 'Big Five': Castle & Cook, Theo H. Davies & Co, C. Brewer & Co., American Factors (now Amfac) and Alexander & Baldwin. They eventually gained control over other aspects of the Hawaiian economy including banking, warehousing, shipping, and importing. The operators of the 'Big Five' held senior post in most other large businesses in Hawaii. Hitch (1992, p.90) describes this as 'a network of interlocking directorships made the whole business community … all members of one family.' The companies did not compete with each other but rather cooperated to keep the prices of their goods and services high.

Why Didn't Capital Flow to Hawaii?

Worth considering is why capital didn't flow to Hawaii after 1875. Most of the investment in the sugar plantations came from reinvested profits. Besides this, capital flows to the industry were negligible. As the theory goes, capital should flow from rich to poor nations due to higher marginal products of capital (MPK). But Lucas (1990) argues that this may not occur due to differences total factor productivity ('A'). However, the technological advancements in Hawaii during the late 19th century suggest that productivity levels were not all that different. Institutional quality however, was not equal across countries. As noted above, the financial system was remarkably undeveloped for much of the late nineteenth century. The problem with Lucas' (1990) argument is that 'A' could represent any number of factors. Some of the explanatory power behind 'A' may be applicable in this case but, given the rapid growth in productivity in the sugar industry over this period, other factors seem to have been at play.

Clark (1978) has argued that it is not discrepancies in 'A' between countries that explains this paradox, but rather labour productivity which is culturally determined. After all, technology is easily transmitted; the Hawaiian sugar industry was certainly operating one of the most efficient production technologies in the sugar industry at the time. Clark (1978) finds that, in 1910, one New England cotton textile operative performed as much work as 1.5 British, 2.3 German, and nearly 6 Greek,

Japanese, Indian, or Chinese workers. Given that the Hawaiian labour supply consisted largely of Chinese and Japanese workers in the late 19th century, there is certainly some merit to this argument. It might also have been the case that the MPK simply wasn't higher. Hawaii is land-scarce. Indeed, Lucas (1990) doesn't account for this variable. Therefore, labour productivity in Hawaii should decline much faster than in the US. By contrast, the US was rapidly expanding its land endowment at this time and experienced a huge population boom, mostly due to immigration. These factors, combined with rapid industrialisation, meant that the MPK was high and there were vast opportunities for capitalists. Moreover, as we will see, Hawaii only really came to the world's attention after WWII. This led to huge inflows of capital despite the fact that wages had converged to US levels. Therefore, investors may simply have been myopic given the developments in US economy.

Finally, Clemens and Williamson (2000) have argued that, had Lucas (1990) used a less contemporary example, such as the first great global market boom after 1870, he might have been less surprised by the results. This was a period characterised by huge global flows of capital, but it was between Britain and the new world, two developed nations. The authors reject any international market failure explanation of the Lucas Paradox. Rather, they argue that British capital heading abroad, especially that invested in private enterprise, went where it was most profitable—chasing natural resources, educated populations, migrants, and young populations. Assuming this conclusion holds true for the US, then it explains why capital did not flow to the sugar industries in Hawaii; it had few natural resources, not a particularly educated workforce, limited US migrants, and an average-age populations.

The Demise of Sugar

WWII had a big impact on the sugar industry. To continue operating, the plantations had to cope with a flood of economic controls; replenishing stocks as they became depleted, making a profit when the price was fixed but costs were variable and finding labour in a very tight market. Despite the military's belief in the importance of sugar, one in three plantation workers left to seek better paid work between 1939 and 1945. There was also a string of operational disruptions to production due to the realities of war. For example, harvesting at night was not permitted due to blackout rules. Remarkably, despite these obstacles, in 1945 sugar was only 14% below 1941 levels.

The real damage was not to output, but rather to the place of sugar in the Hawaiian economy. The most crucial development of this period economically was the total displacement of the 'Big Five' who now took their orders from the military leaders. The plantation oligarchy was instantly and totally eclipsed. Not only did they lose control of their companies, but their plantation economy was largely overshadowed by the immensity of the war economy that was added to it.

Of greater significance was the emergence of labour unions. They had gained momentum prior to 1941, but were restrained from striking because to do so would seem disloyal. However, workers were now dissatisfied: they were unable to change jobs even as others sought higher-paid

work; and they were paid only 41c per hour even though the military paid 62c an hour per worker to the owners, to compensate them for the use of plantation labour. Most of this difference was due to tax, overheads and compensation for the loss of sugar, but workers believed that plantation owners were profiting at their expense.

The power of the union stems from the overspecialised nature of the economy; shutting down warehouses and ports would bring the industry to a standstill, as the International Longshore and Warehouse Union (ILWU) demonstrated during the 1946 sugar contract negotiations, with a shutdown of 79 days. Unionisation brought rapid wage appreciation. Recall that, in 1941 the average daily wage of a sugar worker was $2.18. By 1986 it had risen to $68.72 and by the early 1990s this reached over $100 (Schmitt, 1968).

The final nail in the coffin of the sugar industry came in 1974 when congress abandoned the Sugar Act of 1934. At the time, 80% of the world's sugar was either consumed locally or committed to some other country at a fixed price; the remaining supply was sold at severely reduced rates on the world market. The Sugar Act prevented this excess supply from being dumped onto the US market by providing complicated government subsidies to domestic sugar industries. The result of its repeal was that cheap sugar flooded that US market, and the industry could no longer compete in any meaningful way.

Interestingly, although sugar accounted for a much smaller proportion of the economy in 1988, production levels were roughly the same. Since 1990, the importance of sugar has been diminishing primarily due to a decline in the output of this industry and not rising incomes. By 2009, Hawaiian sugar production was down to 127,000 tons and was produced by the last remaining sugar company in Hawaii, Alexander & Baldwin; this company is also contemplating switching to ethanol production which could mark the end of this industry in Hawaii. Today the sugar industry is paltry by comparison with the tourism sector: in 2007 it generated just over $40 million in sales.

Conclusion

For over 100 years, sugar plantations held a remarkable position with respect to the Hawaiian economy. Their power was such that they could dictate the policies of the Kingdom. Inevitably, such a privileged position brought about collusion and excess profits. It was truly a sweet deal. When the McKinley tariff threatened these profits, the industry responded by displacing the monarchy and seeking union with the US. However, once within the protective barriers of the US, the industry soon realised that annexation was a double-edged sword. The cheap and plentiful supply of Asian labour was cut off. Furthermore, the emergence of a new generation of educated, American-born, labourers who held strong beliefs about paternalism further damaged the supply of labour to the industry. Thus, the industry soon found itself unable to compete with the low-wage sugar economies of the world. Ultimately, the forces that made the deal so sweet for the sugar elite – US ideology – led to the industry's decline.

References

Clark, G., 1987. 'Why Isn't the Whole World Developed? Lessons from the Cotton Mills'. *Journal of Economic History*, 47(1), pp.141-173.

Clemens, M. A. and Williamson, J. G., 2000. 'Where did British Foreign Capital Go? Fundamentals, Failures and the Lucas Paradox: 1870-1913'. National Bureau of Economic Research, Cambridge, MA. *Working Paper no.8028*.

Cushing, R. L., 1985. 'The Beginnings of Sugar Production in Hawaii'. *Hawaiian journal of History*, 19, pp.17-34.

Hitch, T. K., 1992. *Islands in Transition: The Past, Present, and Future of Hawaii's Economy*. Honolulu: First Hawaiian Bank.

Kuykendall, R. S., 1953. *The Hawaiian Kingdom: Volume 2: 1854-1874: Twenty Critical Years*. Honolulu: University of Hawaii Press.

Kuykendall, R. S., 1967. *The Hawaiian Kingdom: Volume 3: 1874-1893: the Kalakaua Dynasty*. Honolulu: University of Hawaii Press.

Kuykendall, R. S. and Day, A. G., 1979. *Hawaii: A History from Polynesian Kingdom to American Statehood*. Englewood Cliffs, NJ: Prentice-Hall.

La Croix, S., 2002. 'Economic History of Hawai'i: A Short Introduction'. Dept. of Economics, University of Hawaii. *Working Paper no.02-3*.

La Croix, S. and Grandy, C., 1997. 'The Political Instability of Reciprocal Trade and the Overthrow of the Hawaiian Kingdom'. *The Journal of Economic History*, 57(1), pp.161-189.

Lucas, R. E., 1990. 'Why Doesn't Capital Flow from Rich to Poor Countries?' *American Economic Review*, 80, pp.92-96.

Morgan, T., 1948. *Hawaii: A Century of Economic Change, 1778 - 1876*. Cambridge, MA: Harvard University Press.

Platt, S. L., 1950. *Immigration and emigration in the Hawaiian sugar industry*. Honolulu: Hawaiian Sugar Planters' Association.

Silva, N. K., 2004. *Aloha Betrayed: Native Hawaiian Resistance to American Colonialism*. Durham: Duke University Press. pp.123-163.

Schmitt, R. C., 1968. *Demographic Statistics of Hawaii: 1778-1965*. Honolulu: University of Hawaii Press.

HOW THE FUNDAMENTALS OF AMERICAN AND JAPANESE CAPITALISM DIVERGE

Graeme O'Meara

Junior Sophister

Japan was once thought of as a considerable challenger to the United States' economic dominance. While it never lived up to those lofty expectations, its form of capitalism still proves fascinating. Graeme O'Meara gives an absorbing account of how Japan's economy works, comparing it directly to the free market capitalism of the United States.

Introduction

How can we define a 'capitalist' economy? Milton Friedman describes it as the 'organisation of the bulk of economic activity through private enterprise operating in a free market' (Friedman, 1962, p.56). J.M. Keynes offers a humorous thought on capitalism: 'Capitalism is the astounding belief that the most wicked of men will do the most wickedest of things for the greatest good of everyone' (cited in Albert, 2000, p.152). Specific histories and cultures lead to many varieties of capitalism; it can be thought of as a stage in social revolution (Groenewegen, 1997). Let us distinguish three prominent styles of capitalism which dominate the global economy: enterprise, social and collective capitalism. Enterprise or 'pure capitalism' is strongly associated with the United States. It is based on the ideas of Adam Smith (1723-90), David Ricardo (1772-1823) and updated by modern economists such as Friedman. 'Its central feature is faith in the untrammelled workings of market competition, born out of the belief that the market is a self-regulating mechanism [Smith's 'Invisible Hand']' (Heywood: 2002: 180). The open market is the epitome of freedom in a capitalist society. Friedman contends that:

> The great advantage of the market is that it permits wide diversity. It is in political terms, a system of proportional representation. Each man can vote, as it were, for the colour of tie he

> wants and get it; he does not have to see what colour the majority wants and then, if he is in the minority, submit (Friedman, 1962, p.15).

Enterprise capitalism is edged by desire for profit, high productivity, flexibility of labour and emphasis on expansion/growth. Trade unions are seen as an obstacle to productivity and are therefore minimised.

Social capitalism involves a significant role for governments in regulating markets and providing social support for the poor, and is mostly associated with continental Europe and is therefore not relevant to this particular argument. However, it isw collective capitalism with which we associate Japan. Its entity is co-operative long-term relationships, originating from the post World War II boom. 'Relational markets' consist of groups such as the *keiretsu*, the *kigyo shudan* and the *kankei-gaisha*, with whom major industries are members and aim to ensure each other's competition by providing finance and expertise. Trademarks of Japanese employment include seniority-based promotion, lifetime employment, pensions/social protection and a high emphasis on teamwork. This collective identity in the workplace is underpinned by narrow wage differentials between manager and worker (Heywood, 2002). Workers are seen as 'members' of the firm rather than employees; 'the firm is a surrogate family' (Eccleston, 1989, p.69).

In this essay, I intend to provide an in-depth analysis into the elements of these two capitalist models that I feel possess fundamental differences. I will use the American style as a base for contrasting the Japanese equivalent.

Liberal Market Economies: the US

In the US, capitalism is pursued by means of an unregulated 'free market'. Organisations act independently, yet interdependently in the sense that firms can pursue any strategy they desire, but such strategies are generally formulated based on what rival firms do. Financial systems or markets for corporate governance encourage firms to be attentive to earnings and shares on the stock market. Companies secure finance based on a valuation in equity markets, where investors depend on publicly available information (yearly financial reports/value of assets) to assess the firm. This also applies to bonds, the issuing of shares and bank lending. Compensation systems for top management who increase net earnings of the company are a trademark of liberal market economies. Trade Unions (seen as a barrier to productivity) have little power to sway management who are under no obligation to establish any representative bodies for employees. Typically, the free market economy relies on macroeconomic policy and market competition to control wages and inflation. A fluid labour market means that workers are in plentiful supply and at any firm's disposal. This makes it less attractive to pursue production strategies

that promise long term employment and since people drift between jobs, transferable skills rather than company specific knowledge are encouraged. The fluid labour market often facilitates transfer of technical information via the movement of scientists from company to company. In this respect, companies are loath to invest in apprenticeship schemes with industry specific skills for fear that competitors will snatch their apprentices (Hall and Soskice, 2001). Companies often provide 'in house' training in marketable skills that employees have incentives to learn. This produces a well equipped labour force suited to job growth in the service sector, but often leaves firms short of employees with highly specialised or company specific skills (Hall and Soskice, 2001).

The liberal market economy is also based on standard market relationships and formal contracts. In America, these relations are mediated by markets, doctrines of contract laws and rigorous anti-trust regulations that prevent collusion to control prices. In this respect, companies wishing to engage or co-operate with other firms are discouraged by the US legal system. The liberal market economy also depends heavily upon the licensing or sale of innovations to assist technology transfer, which is feasible in sectors where patenting is possible, such as biotechnology, micro-electronics and semiconductors. Creators of new inventions profit by licensing it to multiple users, which explains the presence of venture capital firms in the liberal market economy: one success at standard setting can pay for many failed investments. (Hall and Soskice, 2001) In conclusion, inter firm collaboration plays a far less important role in the process of technology transfer than in Japanese capitalism.

Characteristics of Japanese capitalism

Traditional economic theory advocates that free competition is a source of sound economic development, maximum social welfare, innovation and equitable prices. Industrialised nations are always in the position of trying to balance free competition and prevent monopolies. US industry is highly concentrated with numerous monopoly firms such as IBM, Xerox and Boeing. For example, when the airline industry in the US was deregulated in the early 1980's, concentration in the market was higher than ever by the end of the decade. (Adams, 1990) Monopolies tend not to form in Japan; in the US, Eastman Kodak is the only player in the photographic industry, while Japan (with a market half the size) has two only competitors. When a new market emerges in Japan, firms rush to get a foothold - Sega and Sony challenging Nintendo being a case in point.

This 'supermarket' strategy, in which each company has a hand in every area, worked well during Japan's economic boom between 1960 and 1990. 'Made in Japan' gadgets, once considered poorly manufactured, ended up as world leaders in quality, humiliating America's electronics along the way (The Economist, 2009).

And despite acquisitions, bankruptcy or voluntary exit, firms rarely leave an industry. This intense competition in Japan has resulted in lower profitability, but greater innovation and quality in production. 'This pattern of fierce competition without losers has underpinned Japanese economic and competitive success in the post war period' (Tezuka: 1997, p.84). Japanese capitalists suppress competition, augment innovation and boost growth in three fundamental ways: through favourable industrial policy, membership of the 'keiretsu' and a regime of guaranteed lifetime employment.

Japanese Industrial Policy

'Both academic analysts and policy makers in other countries too often overlook the role of Japan's industrial policy in sustaining strong domestic competition' (Tezuka, 1997, p.85). After World War II, the Japanese state acted as a 'capitalist developmental state' which undertook development functions to make up lost ground on the other capitalist powers. Industrial policy decided where and how much investment was made, what kinds of technology development and training took place, as well as the pace and direction of innovation and diffusion (Coates, 2000). The Ministry of International Trade and Industry (MITI) used its influence over the Japanese Development Bank to guide post war companies into industries and technology it thought desirable. It also steered private capital out of low wage textile production into heavy industry – steel, chemicals, shipbuilding and autos. David Coates (2000) describes the Japanese state as a 'gatekeeper' after the war; it controlled entry of capital, technology and manufactured goods into the economy, preventing the domestic market from being colonised by foreign companies bent on export penetration. These policies continued for decades afterwards; in the late 1970s the MITI sponsored a project to develop VLSI semiconductor technology. The government contributed ¥30 billion and industrialists provided ¥42 billion. This strengthened the semiconductor industry in Japan, generating over 1,000 patents on the VLSI process, which the government then licensed out to firms who couldn't afford the investment in the project. As a result, ten semiconductor companies emerged, which led to Japan becoming a world leader in semiconductor manufacturing.

Keiretsu

'Keiretsu' translates as 'order' or 'system,' and is seen as a continuation of the 'zaibatsu,' after the latter was closed down by the Allied forces in the late 1940's. It operates through cross shareholdings, acting as an efficient competitive mechanism. Two distinct types are evident: the horizontal keiretsu (six large groups across a wide range of industries) and the vertical keiretsu (supplier-assembler networks and relationships in major manufacturing industries). Each (horizontal) keiretsu has a member company in one of Japan's major sectors (steel, chemicals, shipbuilding). Groups are loosely co-ordinated by minority cross shareholdings, regular communication by top executives and co-operation for mutual benefit

(Tezuka, 1997). Members do not expect high dividends or returns on investment, they commit to prevent hostile takeovers of a fellow member, external pressures or high levels of autonomy in company management. The banking member acts as a financier and if one member is seen to be struggling, the other members will offer aid to ensure the group's prosperity and continued competition. In this respect, risks are eliminated when making investments, and because risk is shared, lower interest rates can be obtained. For example, when Mazda ran into trouble during the first oil crisis, the Sumitomo group bank provided financial assistance, installed a new CEO and encouraged other members of the keiretsu to purchase Mazda vehicles (Tezuka, 1997).

The vertical keiretsu is held together by a complex mix of suppliers (*kankei-gaisha*) and manufacturing firms – typically in auto and electronics industries. Firms tend to avoid reliance on one supplier; the rate of single source parts in the auto industry in Japan is 12.1%, well below the US rate of 69% (Womack, 1990). The lead firm in the keiretsu is 50% dependent on one supplier, 30% on a second, and 20% on a third. At least one supplier is not a member of the group, and the lead firm will encourage its second and third supplier to match the cost/quality of its first supplier – this way it pushes for improvement and keeps competitive pressure on its first supplier. If a supplier falls behind, the lead firm provides aid, and if there is no improvement over time they will be replaced by another supplier. Suppliers are also encouraged to sell outside the keiretsu to gain experience and inflate economies of scale. However, the lead firm is its main customer, as one keiretsu supplier noted: 'As soon as I succeeded in becoming such a supplier, I was considered part of its 'family'. I was expected to be loyal to that company whatever the sacrifice' (Sakai: 1990, p.32). The Experience Curve shows that as cumulative experience in an industry rises, quality of service rises and costs fall, and that each time cumulative experience doubles, costs per unit fall by 10-30% (Dyer and Ouchi, 1993). Japanese firms consolidate their business with a few high quality suppliers and create the appropriate conditions to permit suppliers to make the necessary investments to accelerate down the experience curve (Dyer and Ouchi, 1993). When a supplier wins a contract with a leading firm in the auto industry, for example Nissan or Toyota, they are guaranteed four years of business (the lifetime of the curve) and know that if they perform well, they can win a contract for the next model. In this way, firms encourage long term planning and investment. Conversely, American firms continually split business among many suppliers, destroying the experience curves, to the extent that no supplier can accelerate down its experience curve to accumulate cost advantages. Suppliers in the US stated that they stand a 69% chance of winning business again with a firm, while a survey estimated the average contract lasts 2.5 years. (Dyer and Ouchi, 1993) As a result, each model changes as the experience curve of the previous supplier is destroyed and the new supplier incurs start up costs.

Lifetime Employment

The Japanese Ministry of Labour (JMOL) defines lifetime employment as the practice of companies hiring their core employees primarily from among new graduates and other young persons, to plan their continual training and development, to continue their employment within the company group over a long period of time (usually until age 55 or 60) and not to discharge or lay off such employees except in very unusual circumstances (JMOL, 1995, p.2).

This applies to core employees (not seasonal or part time workers), and generally occurs in larger companies. A JMOL survey of manufacturing companies with a workforce of over 1,000, showed percentages of male white collar employees who remained with their first employer at various ages: 30 – 78.9%, 40 – 70%, 50 – 66.2%, 60 – 33.6% (JMOL, 1995, p.5). This survey has shown that, on average, two thirds remain until age fifty and by sixty most retire or move to subsidiaries. The general trend is to maintain this system of lifetime employment, as only 2% of firms in the above survey had laid off workers in the last year, while 56% intended to continue offering lifetime placements.

Newly employed graduates face seniority based promotion and salary increases: 'There is no fast track in the Japanese company, but there is a very long track', (Tezuka: 1997, p.88) and jobs tend to be vaguely defined, allowing for versatility. As previously mentioned, duties at the workplace are organised in groups, as opposed to individual tasks. Contrary to the fear of losing their job, employees face the panic of lagging behind. This internal competition could have created a hostile environment, were it not for group work, which means the best performers cannot succeed without the co-operation of their team: '[C]ollective action supersedes individual action' (Konzelmann, 2005, p.49). Similarly, imprecise description of tasks pushes workers to set goals for themselves, resulting in elevated productivity. Just like the keiretsu and industrial policy, lifetime employment obtains full commitment and co-operation from all players, while allowing weaker contenders to remain in the system. Whereas in the market orientated economy there are winners and losers (the unemployed), Japanese society faces lower social costs as lifetime employment becomes a substitute for government welfare programs. 'Because families or companies traditionally looked after people, the state did not have to. Moreover, there is a stigma in Japan if an unemployed person asks for help: 'If you don't work, you don't eat' as the saying goes' (The Economist, 2009). It is in light of these characteristics that I believe American and Japanese capitalism possess fundamental differences.

American and Japanese Models Compared

The Japanese and American systems of corporate governance and employment relations diverge mainly due to the differences assigned to market and production processes. The shareholder system of corporate governance in the US prioritises the providers of capital, and views labour as a factor of

production/cost to be minimised. Conversely, the stakeholder orientated form of corporate governance in Japan observes labour as a productive resource. Thurow (1992) claims that in Japan, employees are the number one stakeholders, customers follow second and shareholders third. In contrast, the USA places shareholders first, customers second and employees a 'distant' third. For Japan, markets serve as an arena in which efficiency is recognised (Konzelman, 2005). Japanese management is controlled by a variety of disciplinary mechanisms that ensure the interest of stakeholders is served. 'In Japan, producing goods and services that enhance the lives of others is good. Spending one's life in the speculative purchase and sale of financial claims is bad' (Dore, 1993, p.77).

An emphasis on the requirements of production is evident in the archetypal Japanese firm: rates of turnover are low, lifetime employment exists, substantial expenditure is made on firm specific training and even when staff is laid off, companies often provide retraining to boost their employability (Konzelmann, 2005). In this respect, employee welfare is seen as the responsibility of the firm and Human Resource management is stakeholder orientated. The Japanese system delivers high quality productivity in tandem with highly co-operative employment relations because the logic of production is not subordinated to the market, i.e. because the desire for profit is not overwhelming, efficiency tends to be higher. Contrastingly, Americans see markets as outlets for productive activity, where competitiveness depends on setting the pace in effective co-ordination of production and innovative activity and in creating as well as responding to consumer demand. Each model is not without its strengths and weaknesses: the free market permits dynamism, job rotation, radical innovation and perfect knowledge, while in the long run it tends to perform weakly when dynamic efficiency requires specific investment (Groenewegen, 2005). For Japan, its forte is investment in human capital, the availability of finance and the execution of team work both inside and outside the workplace. However, profits tend to be lower, and there are innumerable rigidities within the system that make managerial decisions time consuming.

Conclusion

If we juxtapose both capitalist models, fundamental differences can be observed from the outset. Lifetime employment is not a policy of US employers because the fluid labour market means that labour is at their disposal, collaboration of major industries is not encouraged by the US legal system, and the state plays a more subtle role in dictating where industry is to prosper. While America maintains a conscious welfare state and social security system that looks after the unemployed, the Japanese state has placed this responsibility in the hands of private enterprise. The drive to dominate financial markets causes American capitalists to use dynamism to cope with competition, while the Japanese employ and sustain competition as a way of boosting innovation. The overbearing profit motive in US capitalism often means that specific investment in human capital or intermediate goods is not made, and so

performance suffers. When the Japanese do make this essential investment, profit margins are diminished. Which model works best? The answer depends on the objective of the business: to make profit by whatever means possible, or to do so collectively.

References

Adams, W., 1990. *The Structure of American Industry*. 8th ed. New York: MacMillan Publishers.

Albert, M., 2000. *Moving Forward: Programme for a Participatory Economy.* Edinburgh: AK Press.

Coates, D., 2000. *Models of Capitalism: Growth and Stagnation in the Modern Era.* Massachusetts: Blackwell Publishers.

Dore, R., 1993. 'What Makes the Japanese Different?'. In: D. Coates ed., *Models of Capitalism: Growth and Stagnation in the Modern Era.* Massachusetts: Blackwell Publishers.

Dyer, J.H. and Ouchi, W.G., 1993. 'Japanese-Style Partnerships: Giving Companies a Competitive Edge' *Sloan Management Review.* 35(1), pp.51-63.

Eccleston, B., 1989. *State and Society in Post-War Japan*. Cambridge: Polity Press.

Friedman, M., 1962. *Capitalism and Freedom.* Chicago: University of Chicago Press.

Groenewegen, J., 1997. 'Institutions of Capitalisms: American, European, and Japanese Systems Compared'. *Journal of Economic Issues,* 31(2), pp.333-347.

Hall, P.A. and Soskice, D., eds. 2001. *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage.* New York: Oxford University Press.

Heywood, A., 2002. *Politics.* New York: Palgrave Macmillan

JMOL (Japanese Ministry of Labour) , 1995. 'The Present Situation and Future Prospects of the Japanese Style Employment System (Nihon teki toyo siedo no genjo to tenbo)'. In: G.K. Ornatowski, ed. 1998. 'The End of Japanese Style Human Resource Management'. *Sloan Management Review.* 39(3), pp.73-85.

Konzelmann, S.J., 2005. 'Varieties of Capitalism: Production and Market Relations in the USA and Japan'. *British Journal of Industrial Relations,* 43(4), pp.593-603.

Ornatowski, G.K., 1998. 'The End of Japanese Style Human Resource Management'. *Sloan Management Review,* 39(3), pp.73-85.

Sakai, K., 1990. 'The Feudal World of Japanese Manufacturing'. *Harvard Business Review*. 68, pp.38-49.

Tezuka, H., 1997. 'Success as the Source of Failure? Competition and Cooperation in the Japanese Economy'. *Sloan Management Review,* 38(2), pp.83-93.

The Economist, 2009. 'Non-Regular and Non Wanted'. *The Economist*, [Online] 05 February. Available at: http://www.economist.com/node/13059773

Thurow, L., 1992. *Head to Head: The Coming Economic Battle Among Japan, Europe and America.* New York: William Morrow & Co. Inc

Womack, J., Jones, D.T., and Roos, D., 1990. *The Machine That Changed the World*. New York, Rawson Associates, MacMillan Publishing Company.

THE ROLE OF BEHAVIOURAL ECONOMICS IN INCREASING SAVINGS IN DEVELOPING COUNTRIES

Clare Delargy

Senior Sophister

The choice between short-term spending and long-term saving is something that mankind is constantly faced with. People in the developing world in particular often lack the financial institutions that allow for saving and investment, on top of having to live on a very low income. Clare Delargy uses behavioural economics to examine people's propensity to save, and provides some perceptive insights into how savings decisions can be altered in order to stimulate development.

Introduction

One and a half billion people in the world live in conditions of extreme poverty, surviving on less than $1.25 a day (United Nations, 2010). Explanations as to why these people remain trapped in poverty range from the paternalistic, which contend that the poor are psychologically and attitudinally inferior to other members of society and as such require interventions from government, to the libertarian, which hold that the poor are rational individuals whose actions are informed by their own beliefs and values, rendering societal interventions unnecessary (Bertrand, Mullainathan and Shafir, 2004). Behavioural economists have countered these theories by suggesting that poverty is not a psychological weakness but rather the consequence of behaviour, influenced by the same behavioural biases and irrationalities that affect other members of the population. However, for the poor these biases have a much graver impact. Savings are an important aspect of development; they help to fund education, entrepreneurship and healthcare. It is vital that policies attempt to increase the number of people in the developing world who are saving their money in financial institutions. This essay will argue that behavioural economics can use insights from studies to influence rates of savings.

Saving in the Developing World

Saving money is crucial for development; it allows individuals to have enough capital to invest in education and pay for healthcare and it facilitates entrepreneurship. Individuals can use money they have put aside to invest in and expand their business (Dupas and Robinson, 2009). The

outcomes of a failure to save when on the poverty line can mean an inability to continue education or gain adequate healthcare. Access to banking is crucial in order to save. Without it individuals are left with few options; they can save at home, but this does not generate any interest; or they can use savings options, which may involve negative interest (Bertrand, Mullainathan and Shafir, 2004). Banerjee and Duflo (2006) surveyed individuals in 13 developing countries and discovered that, when money was borrowed, 6% came from a formal source. Only 14% of those surveyed had savings accounts. The number of people in the world without access to banking services may total 2-3 billion (Karlan and Morduch, 2009). Financial institutions have been reluctant to provide services in these areas; a disproportionate number of those without bank accounts are poor, and therefore offering such services may not be profitable (Karlan and Morduch, 2009).

Though people living in the developing world must survive on extremely little money, this does not automatically result in an inability to save. Instead, Banerjee and Duflo (2006) found that individuals living in poverty still spend money on extraneous items such as religious festivals, alcohol and tobacco and express wishes to curtail their spending on such products. There are indications that people in the developing world want to save and that they form strategies to combat their lack of access to traditional financial institutions. The presence of informal savings groups such as rotated savings and credit associations (ROSCAs) are a testament to this. Saving using these mechanisms can involve large risks, yet, as it may be the only opportunity these people have to save, they continue to use them (Wright and Mutesasira, 2001). These schemes allow a group of people to contribute their savings into one fund which is loaned out to different members in turn. Dupas and Robinson (2009) found that, though women in rural Kenya had large constraints on savings, the desire to save still remained. Forty percent of participants in an experiment accepted a savings plan even though it offered a negative real interest rate, as the bank in question had high withdrawal fees. In spite of these high transaction costs, take up of the plan led to savings.

The Behavioural Economics of Saving

There are a number of findings from the behavioural economics literature which shed light on human behaviour and can be used in an attempt to understand why people may not save money. Firstly, individuals tend to discount the future and overweigh the present when making decisions. Thaler (1981) found that this discounting does not occur linearly; instead it levels off as the time distance increases. Valuations, such as whether someone would prefer €10 now or €11 next week, decline sharply for decisions about the present; people disproportionately prefer to have the €10 immediately. However this difference in valuation is not observed when people are asked to make choices for the future. For instance, when offered €100 in one year's time versus €110 in one year and one week, these two options tend to hold the same value for people. They are more willing to wait an extra week for a greater sum of money in one year than those who have to resist €10 immediately. When making decisions in the present people can be impulsive and choose short term gains over long

term outcomes. Conversely, when they have to plan for the future, individuals often make choices that have greater long term benefits while correspondingly overestimating their ability to adhere to these plans. According to Mullainathan (2004) there is an inherent clash between what people intend to do in the future and what they actually do when it arrives. Individuals are often intuitively aware of this conflict and impose penalties on themselves to ensure that they carry out a particular task.

Trope and Fishbach (2000) offered a glucose test to participants, telling them that the results they received would inform them what glucose levels result in optimum cognitive functioning. They were told that they could not have sugary foods for either six hours or three days prior to the test and were asked to put forward a sum of money that they would agree to pay if they did not abstain. Those who only had to give up glucose for six hours set an average fine of $1.49, while those who were told to give up high glucose foods for three days set an average penalty of $3.86. The participants in the latter condition anticipated that they would find it difficult abstaining and so set a stricter commitment device for themselves. Bauer, Chytilová and Morduch (2008) found that women living in rural Indian villages, being more likely to discount the future, did not save their money at home. Instead, recognising the temptation that keeping savings at home would bring, these women were more inclined to be members of and borrow from local microcredit organisations, using this as a form of commitment device to generate savings. These studies suggest that individuals can be aware of their inability to stick to a set of behaviours they wish to complete and are willing to set in place devices to ensure that they follow through. Planning problems can explain why individuals in the developing world who, when asked, indicate that they value the education of their children highly, consequently fail to save enough to ensure that they can attend school (Mullainathan, 2004). In order to save enough money, individuals may have to make multiple sacrifices months prior to their children attending school. A determination to save may not be sufficient as individuals may be compelled to use money intended for education for health emergencies, another indication that the high risk environments that people in developing world live in increase the negative effects of behavioural biases.

Loss aversion may also play a role in the lack of savings behaviour in the developing world. Tversky and Kahneman (1992) explain that people react differently to losses and gains of the same size; losses are felt more keenly than gains of a similar magnitude leading to behaviour which is loss averse. Karlan and Morduch (2009) hold that putting aside money for savings requires an individual to forgo goods which they could obtain immediately in exchange for unclear future benefits. They are faced with the choice of losing instant benefits when the future gain is not even assured. Thaler and Benartzi (2004) suggest that when people become accustomed to having a particular income and the lifestyle that accompanies it, putting money into savings, resulting in less disposable income, may provoke a sense of loss.

Another facet of behaviour which influences savings is a tendency to procrastinate and a failure to change the status quo (Samuelson and Zeckhauser, 1988). Thus, if it is the default position

not to have a savings account, people may fail to start one even if they wish to do so. Thaler and Benartzi (2004) successfully increased savings by designing a plan in which, for individuals who opted into the scheme, the default option was to increase the amount saved with each pay increase. Participants were free to opt out of this part of the scheme, but the majority remained with the plan and tripled their savings rate over 28 months. A closer examination of defaults on savings plans may help to explain why people do not save more money. Defaults linked to rises in income, such as those in Save More Tomorrow (Thaler and Benartzi, 2004), could be incorporated into savings plans in the developing world to ensure that, if an individual's income increases, greater savings will result.

Changing People's Behaviour

The behavioural economics perspective allows policymakers to understand why education campaigns, though helpful, are insufficient to provoke total behaviour change. Taking a rational view of the mind suggests that, once people are given adequate information about an issue, they will act in their own best interests. Dinkleman, Levinsohn and Majelantle (2006) demonstrate that information about the transmission of the HIV virus and behaviours necessary to prevent it do not lead to an eradication of such behaviours. The research indicates that a campaign designed to increase savings behaviour by purely providing information may not lead to a lasting change in behaviour. Hyperbolic discounting, self control, status quo bias and loss aversion, among other heuristics, can result in individuals failing to do things that they want to do. These findings can be used in policies attempting to increase savings in developing countries.

Commitment devices allow individuals to meet their future goals or perform desired behaviour, which they recognise would be difficult to achieve due to personal conflict (Bryan, Karlan and Nelson, 2009). The presence of commitment devices in savings plans used by formal institutions would allow individuals to save money for their long term benefit and prevent them from spending it on short term gains, which research indicates they want to do. In a field experiment, Ashraf, Karlan and Yin (2006) examined the effect of offering savings policies with inbuilt mechanisms to ensure commitment. They assessed the discount rate of a group of people living in the Philippines and offered a savings policy to half the sample. Women traditionally manage the finances in Filipino culture and women who were less likely to discount the future were more inclined to choose the plan. The plan involved a voluntary commitment made by the individual to not withdraw any money until a specific monetary target or date was reached. Participants decided themselves to enter into the commitment, but once it was made, money could not be withdrawn until the target was reached. In total, 28% of individuals offered the plan decided to avail of the opportunity. Savings in this group increased by 81% over 12 months, indicating that not only is the programme effective at generating savings but that the savings commitment entered into is viewed as long term. The authors demonstrate that the type of savings programme offered by banks will not only influence the amount people save but also the sort of person who will save. In this case, the programme attracted those who had a lower

discount rate. Using savings accounts such as these not only produces commitment but also represents a safer way of saving money as it is less vulnerable to theft (Dupas and Robinson, 2009).

An increased use of commitment savings plans in developing countries, offered on a voluntary basis, could lead to an increase in savings rates. Individuals would be offered a savings contract with a local bank and the commitment required would be in the form of keeping the money in the account for either a specific amount of time or until a specific monetary target is reached. Karlan and Morduch (2009) recommend marketing and labelling savings plans in a positive manner; drawing particular attention to the gains involved may help to increase uptake and commitment. Policies regarding savings should be accompanied with financial education about the benefits of saving in institutions, which may prevent feelings of loss aversion from affecting take up. Utilizing savings plans could result in an increase in investment in education or other areas, thus promoting development. Policy makers should also consider including automatic increases in savings if an individual's income rises, drawing on the successes of Save More Tomorrow (Thaler and Benartzi, 2004).

Though Ashraf, Karlan and Yin's (2006) study is a significant in attempting to promote saving in the developing world, it cannot establish that these increased savings have generated a corresponding increase in participants' welfare. In fact, it may be that the monetary constraints participants imposed upon themselves resulted in a loss of welfare, such as an inability to pay for urgent medical care. Similarly, the low uptake rate of 28% indicates that such plans should not be made mandatory butoffered only on a voluntary basis. However, it may be that this low uptake rate is a reflection of the absence of interest rate increases in exchange for a reduction in access to money, which is something future policies could rectify. Many criticisms of policies derived from behavioural economics literature centre around questions regarding liberty; is it correct to use research findings on behavioural biases to elicit changes in behaviour? If all humans are susceptible to behavioural biases, surely this applies to those administering policies based on this research too. Can we trust these individuals to generate effective policies without allowing themselves to be influenced by such biases (Whitman, 2010)? However, the approach discussed is not considered to be overly paternalistic in nature, as it provides individuals with a choice to enter a savings plan to combat biases that, as the literature indicates, they are already aware of and attempt to combat themselves.

Conclusion

Using behavioural economics to help stimulate development is a relatively new approach and is by no means perfect; Duflo (2005) laments the fact that it has not to date produced an overarching theory explaining how its findings are relevant to development economics. It is also controversial as to whether it is possible to directly apply findings from western studies, such as behaviour regarding pension contributions (Benartzi and Thaler, 2004), to people living in developing countries. This western perspective, a characteristic of much psychological research (Henrich, Heine and

Norenzayan, 2010), needs to be counterbalanced by research examining the decision making in people living in poverty. The assumption that the findings in the behavioural economics literature can be universally applied must be verified experimentally; however, early studies such as those by Ashraf, Karlan and Yin (2004) suggest that these findings may be universal. Similarly, while behavioural economics has developed theories examining the importance of human behaviour on savings and finance among the poor, it has yet to do so in many other areas, such as education (Banerjee and Duflo, 2010).

Field research suggests that individuals in developing countries exhibit the same biases and behavioural anomalies as those which have already been established, through experiments, in western societies. Providing access to credit and financial institutions is not a panacea which can solve all the problems facing those living in the developing world; lack of education, corruption, inequality and disease continue to be huge constraints on development (Karlan and Morduch, 2009). However, Ashraf, Karlan and Yin's (2006) study demonstrates the potential that commitment savings plans have to increase savings in the developing world, which in turn could have a marked impact on the lives of those living in difficult economic conditions.

References

Ashraf, N., Karlan, D. and Yin, W., 2006. 'Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines'. *Quarterly Journal of Economics*, 121 (1), pp. 635–672.

Banerjee, A.V. and Duflo, E., 2010. 'Giving Credit Where it is Due'. *The Journal of Economic Perspectives*, 24 (3), pp.61-79.

Banerjee, A.V. and Duflo, E., 2006. 'The Economic Lives of the Poor'. *Journal of Economic Perspectives*, 21 (1), pp.141–167.

Bauer, M., Chytilová, J. and Morduch, J., 2008. 'Behavioural Foundations of Microcredit: Experimental and Survey Evidence From Rural India'. Institute of Economic Studies, Charles University, Prague. *Working Paper IES no.2008/28*.

Bertrand, M., Mullainathan, S. and Shafir, E., 2004. 'A Behavioural Economics View of Poverty'. *American Economic Review*, 94 (2), pp.419–423.

Bryan, G., Karlan, D. and Nelson, S., 2009. 'Commitment Contracts'. Economic Growth Centre, Yale University. *Centre Discussion Paper No.980*.

Dinkelman, T., Levinsohn, J. and Majelantle, R., 2006. 'When Knowledge is Not Enough: HIV/AIDS Information and Risky Behavior in Botswana'. National Bureau of Economic Research, Cambridge, MA. *Working Paper no.553*.

Duflo, E., 2005. 'Field Experiments in Development Economics'. Econometric Society. In: *World Congress of the Econometric Society*. London, UK, 19-24 August 2005.

Dupas, P. and Robinson, J., 2009. 'Savings Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya'. National Bureau of Economic Research, Cambridge, MA. *Working Paper no.14693.*

Henrich, J., Heine, S.J. and Norenzayan, A., 2010. 'The Weirdest People in the World'. *Behavioural and Brain Sciences*, 33 (2-3), pp.61-83.

Karlan, D. and Morduch, J., 2009. 'Access to Finance'. In: Rodrick, D. and Rosenzweig, M., eds. *Handbook of Development Economic.* Amsterdam: Elsevier. Ch.2.

Samuelson, W. and Zeckhauser, R.J., 1988. 'Status Quo Bias in Decision Making'. *Journal of Risk and Uncertainty*, 1, pp.7–59.

Thaler, R. and Benartzi, S., 2004. 'Save More Tomorrow: Using Behavioural Economics to Increase Employee Saving'. *Journal of Political Economy,* 112, pp.164-187.

Thaler, R., 1981. 'Some Empirical Evidence on Dynamic Inconsistency'. *Economic Letters*, 8 (3), pp.201–207.

Trope, Y. and Fishbach, A., 2000. 'Counteractive Self-Control in Overcoming Temptation'. *Journal of Personality and Social Psychology*, 79(4), pp.493–506.

Tversky, A. and Kahneman, D., 1992. 'Advances in Prospect Theory: Cumulative Representation of Uncertainty'. *Journal of Risk and Uncertainty,* 5, pp.297–323.

United Nations, 2010. 'The Millennium Development Goals Report'. United Nations. In: *UN Summit on the Millenium Development Goals*. United Nations, New York, 20-22 September 2010.

Whitman, G., 2010. 'The Rise of the New Paternalism'. *Cato Unbound*, [Online] 5 April. Available at: http://www.cato-unbound.org/2010/04/05/glen-whitman/the-rise-of-the-new-paternalism/

Wright, G.A.N. and Mutesasira, L., 2001. 'The Relative Risks to the Savings of Poor People'. *Microlinks*, [Online] January. Available at: http://www.microlinks.org/ev.php?ID=24393_201&ID2=DO_TOPIC [Accessed: November 20, 2010]

A Discussion of Foreign Currency Debt in Emerging Market Economies

Grace Walsh

Senior Sophister

At a time when Ireland faces sovereign debt issues of its own, the difficulties that emerging market economies face in financing their sovereign debt should not be dismissed. While we take it for granted that Ireland's sovereign debt is almost entirely in euro, developing countries are forced to issue much of their debt in foreign currencies. Grace Walsh provides an insightful analysis into the merits of foreign currency debt, ultimately concluding that the pros are outweighed by the cons.

Introduction

In the wake of the recent financial crisis emerging market economies have become the darlings of investors seeking above average returns when global yields are low. The 'North Atlantic Crisis' gravely affected vulnerable developing economies by reducing the developed world's demand for imports, supply of capital flows and need for foreign workers. Russian GDP fell by 9.5% in 2009 (The Economist, 2009) whilst Cambodia's foreign direct investment fell by 50% compared to 2009. These figures illustrate the extent to which the financial crisis affected emerging market economies.

The importance of emerging market economies as engines of global economic growth has become more important than ever. In 2000 emerging market economies contributed 37% of world output, and by 2008 their share amounted to 45%. The BRIC economies' (Brazil, Russia, India and China) share of world output rose from 16% to 22% during the same period. Of the increase in world output from the year 2000 to 2008, 60% took place in the BRICs alone (The Economist 2009). Thus the economic consequences generated by the activities of emerging markets will impact on all countries, developing and developed.

The purpose of this paper is to explore the pros and cons of foreign currency debt with regard to emerging market economies. Foreign currency debt usually takes the form of sovereign bonds or international bailout packages from institutions such as the International Monetary Fund.

Borrowing: A Necessary Evil?

Since the 18th century, debt issued on the international markets has been denominated in the currency of the market of issue and the currency of the issuing country (Bordo and Meissner, 2006). Developed countries such as the United States and United Kingdom issue debt in their domestic currency, as not only does demand exist to hold such liabilities but the stability of the domestic economy and exchange rate provide assurance to international investors that the value of their investment is not likely to experience a severe depreciation. Developed countries typically have a larger proportion of their sovereign debt denominated in the domestic economy than emerging market economies whilst emerging market economies in contrast hold a higher proportion of their sovereign debt in foreign currencies (Cottarelli et al., 2008).

Emerging markets traditionally experience less demand for their domestic currency denominated bonds. This is partially due to the instability associated with the macroeconomic policy of emerging market economies. Emerging markets typically suffer from high inflation, low levels of capital, low levels of labour productivity, low domestic savings rates and under-developed financial systems. These factors can act as a deterrent to foreign capital flows. Additionally, emerging market currencies are empirically less stable than those of developed economies. For example, Argentina suffered currency crises in June 1975, February 1981, July 1982, September 1986, April 1989, February 1990 and again in 2002. Mexico suffered a currency crisis twice in 1982 and again in December 1994 (Calvo and Reinhart, 1999). In 1982, five Latin American countries suffered a currency crisis including Chile, Bolivia and Uruguay. Such regular crises actively deter investors from investing in emerging market economies, as not only is exchange rate risk prohibitive but past performance also suggests poor economic governance is a future possibility, raising the risk profile of the sovereign's debt.

The Composition of Emerging Market Debt

Emerging market economies have a different debt structure to developed economies. Investment in the sovereign debt of emerging market economies is quantifiably riskier than that of developed economies. This reflects the risk premium investors require for holding such debt. The composition of emerging market economy debt may make such economies more prone to crisis in the event of a default, as a currency depreciation or banking crisis may soon follow. Cottarelli et. al. (2005) illustrate that developed economies, such as the US, have debt almost entirely denominated in the home currency, in contrast to emerging market economies whose foreign currency debt typically stands at 42% and amounts to 63%, on average, in the year prior to default. A study by Cottarelli et al. (2005) makes clear that the composition of the debt is an important factor in whether or not default occurs, as emerging market economies have a larger share of floating rate or indexed debt than

developed countries. Ceteris paribus, should interest rates increase average borrowing costs will also rise. This may increase pressure on the sovereign's ability to repay its debt and increase the temptation to default if the cost of repayment rises to a prohibitive level.

Secondly, emerging market economies borrow for shorter periods of time than developed countries, whether borrowing in their own currency or another. Jeanne (2003) argues that foreign currency debt, especially of the short-term variety, is the most dangerous form of sovereign debt. The study by Cottarelli et al. (2010) clearly indicates that difficulties in repaying sovereign debt arise from a combination of long maturities and the composition of debt, which for emerging market economies is mostly composed of foreign currency borrowing, floating rate and indexed debt. These characteristics of emerging market debt combined with historical evidence indicates that investing in emerging market economy debt is riskier than that of developed market economies who tend to borrower for longer periods in their own currency. The next two sections will go on to discuss the pros and cons of foreign currency debt.

Advantages of Foreign Currency Debt

Foreign currency debt has many advantages for the borrower. It provides access to financial capital to fund investment, increases financial globalization and promotes better macroeconomic policy and governance in the borrowing country.

The issuance of foreign currency debt is a method of financing domestic investment which otherwise may not be possible if there is a low level of domestic demand for bonds or if domestic saving is low. The benefits of investment are manifold. It creates employment and generates a stream of revenue, which benefits the state in terms of increased tax revenue and consumers who receive higher levels of income. Investment increases the national income of the country and subsequently GDP per capita indicating increased economic welfare, it increases the productivity of labour and facilitates convergence to the level of developed economies. Though the Lucas Paradox (Lucas, 1990) and findings by Prasad, Rajan and Subramanian (2007) effectively illustrate that capital does not flow to poor countries from rich countries in the volumes predicted by theory, for countries with low levels of domestic financing foreign currency debt is an essential source of finance.

Issuing foreign currency debt forces the government to adopt better policies to ensure that its level of sovereign risk does not deter debt issuance or equity flows in the country. Thus it is important for international markets and investors to observe prudent economic management in the borrowing country, Kose et al. (2009) argue that increased financial globalization 'unleash[es] forces that result in better public and corporate governance' in addition to imposing discipline on macroeconomic policy. Improved governance and disciplined monetary and fiscal policy enhance economic growth and stability.

Thirdly, Jeanne (2000) argues that borrowing in a foreign currency is a powerful commitment device. The borrower has more incentive to repay its loan and build a credible relationship with

international investors. Jeanne uses a model of an entrepreneur borrowing from abroad to illustrate the strong incentive for the sovereign to pay back its foreign debt. In this model the failure to achieve high returns early in the project is punished by termination of the loan. Thus the borrower has strong incentive to make his venture successful and deliver high returns to investors. Lenders are aware of this and subsequently charge a lower rate of interest than the entrepreneur could have attained in the domestic market.

Extrapolating this example to emerging market economies, we see that it may be more cost efficient to raise capital abroad rather than at home, especially if there is a low level of domestic demand for sovereign debt. It is evident from this model that there is strong incentive not to default on the loan and to deliver returns to investors as they fall due, to ensure continued access to the capital markets. The maturity mismatch of borrowing and lending debt is explored in the next section.

Thus there are numerous positive benefits of foreign currency debt for emerging market economies. We will now turn to analyze the disadvantages of this activity.

Disadvantages of Foreign Currency Debt

Borrowing in a foreign currency creates numerous potential difficulties. Jeanne (2000) argues that large quantities of short-term foreign currency liabilities have made crises in the borrowing country deeper, more contagious and generally more difficult to manage. Krugman (1999) argued that foreign currency debt amplifies financial crises to the point of making them self-fulfilling.

The primary issue with accumulating foreign currency debt is the obligation to repay investors in a currency which the government has no control over. A country whose liabilities are denominated in a foreign currency is unable to hedge exchange rate risk which leaves it greatly exposed to global financial forces outside of its control. The country in whose currency the debt is denominated may set interest rates and influence exchange rates according to the domestic situation, which may not suit the borrower and make repayment more difficult. Bordo and Meissner (2006) argue that countries whose foreign currency liabilities are not matched by foreign currency assets may find it more difficult than countries with comparably more foreign assets to repay their foreign currency debts in the event of depreciation. Emerging market economies, especially those with floating exchange rates, often experience currency depreciation caused by sudden stops in capital inflows, loose monetary policy or disadvantageous terms of trade (Bordo and Meissner, 2006). This explains why emerging markets often peg the exchange rate to that of larger developed economies such as the US, a phenomenon known as a ‘fear of floating’ (Calvo and Reinhart, 1999). These combined factors make the repayment of foreign debt more challenging, increasing the risk of default, reducing external funding and potentially causing economic collapse.

If the exchange rate floats and depreciation occurs, the currency mismatch will cause bankruptcies. This, however, provides greater incentive for investors to hedge their positions. A greater incentive and propensity to hedge by investors reduces exchange rate exposure meaning that

large exchange rate fluctuations will not cause financial crises (Eichengreen and Hausmann, 1999). However, floating exchange rates are just as likely to cause default as pegged exchange rates, if not more so, as the country is more vulnerable to global shocks originating in the developed world or other developing markets.

In an increasingly globalised world the effects of sovereign debt default or crisis are not limited to the borrowing country should default occur. Argentina's 2001 default on $81 billion illustrates the negative consequences of default externally and domestically. The default occurred following three years of severe recession, a dramatic reduction in foreign direct investment, high inflation, high interest rates and a large budget deficit. GDP declined by 10.9% (The Economist, 2010). When the Argentine government announced its default its sovereign debts amounted to $144.5 billion, equating to 53% of national GDP (Shapiro and Pham, 2006). One hundred and fifty-three bonds in seven currencies and eight national jurisdictions were defaulted on. 53% of the debt was denominated in US dollars and 33% in euros. In a study by Shapiro and Pham (2006), direct costs for worldwide investors amounted to $83.6 billion in capital losses, interest payments never received and foregone investment returns. The Argentine default caused the devaluation of the peso against the dollar, euro and other major currencies reducing the value of existing investments in Argentina by those investors. It resulted in the direct loss of $83.6 billion in taxes paid by Argentine bondholders to their national governments. Finally, the default required a substantial bailout package to be compiled by the IMF. This $23.2 billion package was funded by taxpayers across the world. Thus, in addition to the loss of funding experienced by the borrowing sovereign, a default on foreign currency debt also negatively impacts a number of other international parties.

No discussion of emerging market economies' foreign currency debt is complete without a discussion of original sin. The case of original sin emphasizes two negative facets of foreign currency debt – currency mismatch and maturity mismatch. Eichengreen (2001) defines original sin as when the home currency cannot be used to borrow abroad or to borrow long-term, even domestically. Latin American countries such as Mexico, Uruguay and Colombia are the most notorious victims of original sin.

Currency mismatch arises because projects that generate pesos or francs will be financed with dollars. Maturity mismatch arises because long-term projects will be financed by short-term investments. When original sin prevails, both fixed and floating exchange rates represent serious challenges. If the exchange rate floats, a large depreciation could cause bankruptcies and an eventual default on foreign currency debt. A fixed exchange rate defended by the sale of reserves and interest rate increases precipitates defaults on domestic debt (Eichengreen and Hausmann, 1999). Eichengreen and Hausmann (2005) showed that countries with higher levels of original sin have a greater exchange rate volatility and higher macroeconomic volatility than those with lower levels of original sin. The only viable solution is dollarisation (or its euro equivalent) as income flows are then in the same currency as liabilities.

The case of Uruguay illustrates that dollarisation can in fact be a temporary solution to original sin whilst the sovereign rebuilds its credibility, whilst the success of Panama illustrates how effective long-term dollarisation can be. Panama has followed a policy of dollarisation since 1904. The benefits of pursuing this policy include the lowest rate of inflation in Latin America, full integration into the US financial system, high economic growth rates and stable access to international capital markets.

It is possible for sovereigns to redeem themselves through good behavior, eradicating the blemish of original sin. This may entail reform of macroeconomic policies including the temporary adoption of dollarisation, improved political stability and the development of a deep, liquid financial system to restore international faith in the domestic currency so that the sovereign can use it to borrow abroad and domestically.

Conclusion: Is Foreign Currency Debt Unavoidable?

In light of the disadvantages of foreign currency debt, is it possible for emerging market economies to borrow in a foreign currency and not bear the risk associated with foreign currency borrowing? As discussed above, one such method of this is dollarisation. If the liabilities of the emerging markets are in the same currency as their assets then many of the negative consequences associated with foreign currency borrowing could be eradicated. Dollarisation is not a miracle cure nor does it ensure sound fiscal management, but it provides relief against many of the pressures of exchange rate risk. An alternative solution is the introduction of a common, global currency for all countries. However, the real-life difficulties of achieving this prevent it from being a realistic or credible option. Unless the borrower can borrow in its own domestic currency, foreign currency debt will always pose significant advantages and disadvantages for emerging markets. Thus they must find credible ways of coping with the risks and opportunities foreign currency borrowing brings, in order to continue to converge to the level of developed economies.

The purpose of this paper was to evaluate emerging market economies' foreign currency debt. For small emerging market economies, repayment of sovereign debt is challenging as they struggle with large government deficits, reduced capital flows and reduced western demand for exports (IMF, 2010). Throughout the course of this essay, having compared the advantages and disadvantages of foreign currency debt, it is apparent that the negative consequences of such debt for emerging markets outweigh the advantages. Exchange rate risk, liability mismatch and the international effects of default illustrated, using Argentina's 2001 expreience as a case study, the dangers of foreign currency debt, particularly when it is of the short-run variety. Despite this, Shapiro and Pham (2006) provide evidence that foreign lending by investors to emerging market economies has increased despite 59 sovereign debt defaults during the period 1976 to 2004, 16 of which have occurred since 1998. However, markets have short memories and investors will always seek high levels of returns and diversification of their portfolios, just as emerging market economies will continue to finance growth

by borrowing from abroad. Borrowing and investing in foreign currency debt will continue into the future, just as the benefits of borrowing in foreign currency for emerging market economies will most certainly be a mixture of the pros and cons discussed in this paper.

References

Bordo, M. and Meissner, C., 2006. 'The Role of Foreign Currency Debt in Financial Crises: 1880-1913 versus 1972-1997'. *Journal of Banking and Finance*, 30(12), pp.3299-3399.

Calvo, G. and Reinhart, C., 1999. 'When Capital Flows Come to Sudden Stop: Consequences and Policy Options'. *Working paper no.6982*, [Online].
Available at: http://ideas.repec.org/p/pra/mprapa/6982.html

Cottarelli, C., Forni, L., Gottschalk, J. and Mauro, P., 2010. 'Default in Today's Advanced Economies: Unnecessary, Undesirable and Unlikely'. *IMF Staff paper*, [Online].Available at: www.imf.org/external/pubs/ft/spn/2010/spn1012.pdf

Eichengreen, B. and Hausmann, R., 1999. 'Exchange Rates and Financial Fragility'. Federal Reserve Bank of Kansas City, *Proceedings*, [Online].
Available at: http://ideas.repec.org/p/nbr/nberwo/7418.html

Eichengreen, B., 2001. 'What Problems Can Dollarization Solve?' *Journal of Policy Modeling* 23(3), pp.266-267. [Online]. Available at: http://ideas.repec.org/a/eee/jpolmo/v23y2001i3p267-277.html

Eichengreen, B. and Hausmann, R. 2005. *Other people's money: debt denomination and financial instability in emerging market economies*. Chicago: University of Chicago Press.

Jeanne, Olivier., 2000. 'Foreign currency debt and the global financial architecture'. *European Economic Review*, 44, pp.719-727.

Jeanne, Olivier., 2003. 'Why do Emerging Market Economies Borrow in Foreign Currency?'. IMF, *Working Paper*. [Online]. Available at: www.imf.org/external/pubs/ft/wp/2003/wp03177.pdf

IMF (International Monetary Fund), 2010. 'World Economic Outlook Update'. *IMF*, [Online] 26 January. Available at: http://www.imf.org/external/pubs/ft/weo/2010/update/01/

Lucas, R., 1990. 'Why Doesn't Capital Flow from Rich to Poor Countries?'. *American Economic Review*, 80(2), pp.92-96. [Online]. Available at: http://ideas.repec.org/a/aea/aecrev/v80y1990i2p92-96.html

Krugman, P., 1999. 'Balance Sheets, the Transfer Problem and Financial Crisis'. In: P. Isard, A. Razin and A. Rose, eds. I*nternational Finance and Financial Crisis, Essays in Honor of Robert P. Flood. Norwell*. Massachusetts: Kluwer Academic Publishers.

Kose, M., Prasad, E., Rogoff, K. and Wei, S., 2009. 'Financial Globalization: A Reappraisal.' IMF, *Staff Papers*. [Online]. Available at: http://www.palgrave-journals.com/imfsp/journal/v56/n1/abs/imfsp200836a.html

Overseas Development Institute (ODI). 'The Global Financial Crisis and Developing Countries: What Happened, and What Have We Learned?'. *ODI*, [Online] 19 March. Available at: http://blogs.odi.org.uk/blogs/main/archive/2010/03/19/52457.aspx

Prasad, E., Rajan, R. and Subramanian, A., 2007. 'The Paradox of Capital'. *Finance and Development*, 44(1). [Online]. Available at: www.imf.org/external/pubs/ft/fandd/2007/03/prasad.htm

Shapiro, R. and Pham, N., 2006. 'Discredited – The Impact of Argentina's Sovereign Debt Default and Debt Restructuring on U.S. Taxpayers and Investors'. *American Task Force Argentina*, [Online] October. Available at: www.sonecon.com/docs/studies/argentina_1006.pdf

The Economist, 2009. 'Not Just Straw Men'. *The Economist*, [Online] 18 June. Available at: http://www.economist.com/node/13871969

The Economist, 2010. 'Default Settings'. *The Economist*, [Online] 31 March. Available at: http://www.economist.com/node/15814868?story_id=15814868

SHOULD THE RENMINBI FLOAT?

Jason Somerville

Senior Sophister

China has been widely accused of being a 'currency manipulator' in recent years, leading to concerns about a potential 'currency war'. Jason Somerville provides a balanced study of the issue, and ultimately argues that while China's currency appears to be undervalued, an incremental move towards floatation is in the best interests of both China and the rest of the world.

Introduction

> 'The argument for a flexible exchange rate is, strange to say, very nearly identical with the argument for daylight savings time. Isn't it absurd to change the clock in summer when exactly the same result could be achieved by having each individual change his habits?' (Friedman, 1953, p.173).

The valuation of the Renminbi (RMB) has been a source of lively debate for over two decades. After China de-pegged its currency from the US dollar (USD) in 2005 the rhetoric dried up, but the onset of the financial crisis and China's renewed intervention in the currency market has reignited the debate and become a source of political tension between China and the rest of the world – in particular the United States. There have been vociferous complaints by the US government that, by keeping its currency weak, Chinese exports are artificially cheap, causing job losses in America, Japan and other Asian economies. Stepping back from the political dimensions of this debate, this paper explores first whether the RMB is undervalued using a variety of economic measures. Second, the normative question 'should the RMB float?' will be explored. Finally, the 'fear of floating' argument will be assessed in the context of the Chinese economy in an attempt to shed some light on China's reluctance to let the RMB float.

Historical Background

In 1993, following the unification of the Chinese exchange rate regime, the yuan fluctuated within a relatively narrow range against the USD before being officially pegged at 8.28 yuan per dollar in 1996. This peg was rigidly upheld until July 2005. For most of its early period, this pegged exchange rate regime attracted little criticism. Indeed, it was widely seen as contributing to internal and external stability. However, from 2002 onward, there was mounting pressure, both external and internal, on China to de-peg the RMB and allow it to appreciate against the USD. In July 2005, the People's Bank of China (PBoC) announced that it would allow the RMB greater flexibility.

The cumulative appreciation against the dollar reached 17.5% by July 2008. However, the onset of the financial crisis prompted the Chinese government to temporarily change its policy due to concerns regarding economic stability. The RMB was unofficially re-pegged to the USD at a rate of 6.83 yuan per dollar. In June 2010, China again lifted the peg and the RMB resumed its appreciation against the dollar. As of November 22, the RMB had appreciated 19.7% against the dollar since the peg was lifted in 2005. The question is, has this appreciation caused the RMB to converge on its equilibrium value, or is it over/under valued?

Is the RMB Undervalued?

'All models are wrong, but some are useful' (Box and Draper, 1987, p. 424). As this quote indicates, there are limitations to every model and so several will be explored in an attempt to determine whether or not the RMB is undervalued. Since the RMB was only recently de-pegged from the USD, no comprehensive empirical analysis has been published. However, the majority of the adjustment occurred in the 3 year period between July 2005 and July 2008. Investigations that have covered this period will therefore provide an accurate yardstick for answering the question: 'is the RMB under or overvalued?'

Goldstein (2004) used the 'underlying balance approach' and 'adjustment of global payment imbalances approach' to measure the equilibrium exchange rate of RMB/USD. He found that the RMB was undervalued by about 15% to 25% against the USD prior to July 2005. Another approach to estimating the equilibrium value of the RMB has been to use behavioural equilibrium exchange rate (BEER) specifications. These models incorporate a variety of channels through which the real exchange rate is affected (e.g. Funke and Rahn, 2005). However, these approaches failed to identify if a currency is misaligned relative to another country's currency for the same reason that relative PPP fails to do so – because they typically rely upon price indices and not actual prices.

An alternative framework has been presented by Prasad and Wei (2005) who use a flow model to determine the relative value of the RMB. By examining the composition of capital inflows into and out of China, they argue that much of the reserve accumulation that has occurred in recent years is due to speculative inflows; hence, the actual degree of misalignment is small. Conversely, Frankel (2005) has argued that the RMB was 35% undervalued from a long-run perspective.

An interesting argument has been put forward by Cheung, Chinn and Fujii (2007) in their paper entitled 'The Overvaluation of Renminbi Undervaluation', in which they highlight a number of problems with these approaches. In attempting to determine whether the RMB is misaligned, they find that, once sampling uncertainty and serial correlation are accounted for, there is little statistical evidence that the RMB is undervalued, even though the point estimates usually indicate economically significant misalignment.

So prior to de-pegging in July 2005, while there were some statistical shortcomings, the overall consensus emerging from this line of research was that the RMB was undervalued relative to its long-run equilibrium. This is not surprising given that the Chinese economy averaged a growth rate of 9.25% since the currency peg was introduced in 1996.

Since 2005's de-pegging, the RMB has appreciated nominally by 20% against the USD. However, the debate continues to rage. Cheung, Chinn and Fujii (2009) estimated the deviation of the RMB from its expected value at different stages over the de-pegging period. They found that the RMB was substantially below the value predicted by estimates based upon a cross-country sample, when using the 2006 vintage of the World Development Indicators. The economic magnitude of the misalignment is substantial – on the order of 50 per cent in log terms. Again, these authors add the caveat that the misalignment is typically not statistically significant, in the sense of being more than two standard errors away from the conditional mean. Interestingly, when 2008 data is used, the estimated undervaluation is on the order of only 10 per cent. Of note is their conclusion that 'these results are not informative with regard to the question of how a change in the RMB/USD exchange rate would affect the overall US trade deficit' (Cheung, Chinn and Fujii, 2009, p. 2). This highlights the uncertainty associated with what is the 'best' exchange rate policy for China, and indeed the world.

Nair and Sinnakkannu (2010) examined the evolution of the RMB between 2005 and 2009. They found that, on average, the RMB was undervalued in its real terms against the USD by 20.6% and against the euro (EUR) by 15.5% between July 2005 and June 2009.

Table 1: The overvaluation (+) and undervaluation (−) of the RMB against the USD, Japanese Yen (JPY) and EUR in %

Year	RMB/USD	Stdev	RMB/JPY	Stdev	RMB/EUR	Stdev
H2/2005	-19.40%	0.85	+2.202	0.48	-15.095%	0.53
2006	-21.06%	0.81	+2.079	0.64	-16.26%	0.61
2007	-20.82%	0.75	+5.171	1.37	-15.44%	0.81
2008	-20.25%	0.78	+8.501	1.08	-14.49%	0.89
H1/2009	-20.88%	0.80	+7.748	1.23	-16.17%	0.81

Overall	-20.57	0.93	+5.182	2.87	-15.46%	1.01

(Nair and Sinnakkannu, 2010)

Though the nominal exchange rate of RMB/USD has slowly increased since its de-pegging, the real exchange rate with the USD was found to be consistently undervalued.

Having established that, despite empirical limitations, the RMB appears to be undervalued by most statistical measures, it is imperative to explore the normative question: '*should* the RMB float?' Due to the nature of this analysis, there is no definitive answer. Value judgments ultimately come into play and so it is not surprising that this issue is a source of political tension. Here the main economic arguments, both for and against floating the RMB, will be presented.

The Argument for a Floating Exchange Rate

> 'The biggest problem with China's economy is that the growth is unstable, unbalanced, uncoordinated, and unsustainable' (Premier Wen Jiabao, 2007).

Given that the RMB appears to be undervalued, the argument for a flexible exchange rate is effectively an argument for currency appreciation. The rationale for a flexible exchange rate regime was put forward in Milton Friedman's (1953) famous book '*Essays in Positive Economics'*. He argues that 'it is far simpler to allow one price to change, namely, the price of foreign exchange, than to rely upon changes in the multitude of prices that together constitute the internal price structure' (Friedman, 1953, p.173). By keeping the exchange rate fixed, incentives have become distorted and agents responding in an optimal way may generate a sub-optimal outcome.

Funke and Rahn (2005) have argued that China has achieved spectacular growth by selling deliberately undervalued exports and transforming itself into the 'workshop of the world'. This growth has been underpinned by a policy of depressing its exchange rate in order to improve its international competitiveness. Indeed, Makin (2009) has argued that an inflexible yuan has yielded higher short-run output gains for China at trading partners' expense through a form of 'exchange rate protection'. The author goes onto assert that the RMB's misalignment has contributed to the global imbalances that drove the financial crisis.

Obstfeld and Rogoff (2009) have expanded upon this, arguing that such imbalances are intimately connected with the recent financial crisis. An artificially low exchange rate drove export led growth, culminating in large surpluses. They go on to suggest that these surpluses contributed to the United States' ability to borrow cheaply abroad and thereby finance its unsustainable housing bubble. Worryingly, this trend has continued. In the first 10 months of 2010, China's trade surplus

totalled $147.77 billion. This balance of payments surplus implies that the reserve component of the monetary base is increasing. Table 2 shows that China's foreign reserve accumulation has been growing steadily since 2001 and amounted to almost half of GDP in 2009. The fear is that these surpluses and the on-going reserve accumulation will provide the fuel for the next bubble.

Table 2 China's foreign reserve accumulation from 2001 to October 2010

Year	USD Billion	Percentage of China's GDP
Dec 01	215.60	16.3
Dec 02	291.10	20.0
Dec 03	403.25	24.6
Dec 04	609.93	31.6
Dec 05	818.89	36.5
Dec 06	1,068.60	40.2
Dec 07	1,528.25	45.2
Dec 08	1,946.03	45.0
Dec 09	2,399.15	48.1
Jun 10	2,454.28	N/A

(Chinese State Administration of Foreign Exchange)

While the US has not officially labelled China a currency manipulator, as the *Financial Times* columnist Martin Wolf (2010) put it recently, 'If a decision to invest half a country's gross domestic product in currency reserves is not exchange rate manipulation, what is?'.

Despite this expansion in the monetary base, inflation has been remarkably low over the past decade. High inflation would at some point render China uncompetitive at the prevailing exchange rate and lead to pressure for an appreciation of the exchange rate (Greenwood, 2008). Therefore, in order to offset the increase in the monetary base, the PBoC has issued 'Sterilisation bonds'. These are intended to soak up the increase in the monetary supply and therefore dampen inflationary pressures. However, as Mussa (2007) points out, sterilisation blocks the monetary, price and interest rate mechanisms that come into play to equilibrate imbalances. Furthermore, it becomes increasingly difficult to sterilise the inflow over time (Frankel, 2005).

Figure 1: Exchange rate and domestic inflation in China 1987 -2010

(Economist Online)

It could be argued that the appreciation of the RMB is inevitable given the Balassa-Samuelson effect. According to this argument, as productivity increases in rapidly developing economies like China, wages will increase causing the price level to rise. Eventually, this will lead to a nominal appreciation of the exchange rate due to excessive inflation (Xiao, 2008). Frankel (2005) has argued that, from a longer-run perspective, the prices of goods and services in China are low – not just low relative to the United States (at .23), but also low by the standards of a Balassa-Samuelson relationship estimated across countries (which predicts .36).

Fundamentally, an appreciation in the value of the Renminbi is necessary to boost domestic consumption. Increasing domestic consumption is essential to support the local economy, reduce China's dependence on export-led growth, reduce the risk of inflation and to allow Chinese workers to consume imported goods and services which are currently comparatively expensive to local goods. Guo and N'Diaye (2009) have explored in some depth the sustainability of China's export-oriented growth over the medium to longer term. They find that, if exports are to continue to drive growth it would require significant gains in market share through lower prices in a range of industries. They suggest this could be achieved through a combination of increases in productivity, lower profits, and higher implicit or explicit subsidies to industry. However, the evidence suggests that it will prove difficult to accommodate such price reductions within existing profit margins or through productivity gains. Therefore they conclude that 'rebalancing growth toward private consumption would provide a large impetus to output growth' (Guo and N'Diaye, 2009, p.2)

The Argument Against Floating

> ‘Continued pressures to revalue the RMB run the risk of disrupting the internal balance of saving and investment and pushing China on to a path of slower economic growth’ (Bosworth, 2004, p.1).

Those who argue against floating the RMB point to the parallels between Japan in the early 1980s and China today. In the third quarter of 2009 nominal GDP in Japan – though still vast by global standards – sank below its level in 1992. In the three decades between 1960 and 1990, Japan experienced rapid economic expansion, making it the world’s second largest economy. However, such a sudden expansion led to overheating, as excesses accumulated and bubbles formed. Japan's mistake was that it bowed to the demands of the US in the 1980s and allowed its currency (the Japanese Yen) to appreciate. It was the appreciation of the Japanese Yen in the 1980s that led to the bursting of a housing bubble and the stagnation of the Japanese economy for the last two decades.

Similarly, China has experience phenomenal growth over the past two decades and recently surpassed Japan as the world’s second largest economy. This had led market commentators to suggest that a huge property market bubble has formed, with some suggesting it is the biggest in financial history (Xie, 2010). The fear is that, by allowing the RMB to appreciate, the bubble could burst and devastate the Chinese economy. Rather, a policy of gradually deflating the bubble appears to be the best way forward for the economy. Floating the RMB and leaving its value to the whim of market forces could lead to a swift appreciation of the currency, causing the bubble to burst. The knock-on effects for economic activity could be disastrous.

However, the counter-argument is that by allowing the RMB to float, the PBoC will be free to deploy monetary policy to deflate the housing bubble. This is the classic policy trilemma. A country must ‘trade-off’ the following three goals: a fixed exchange rate, an independent monetary policy and capital mobility. According to the Mundell-Fleming model, any two policies imply that the third is not possible (Feenstra and Taylor, 2008). Which two policies are the best for China, and indeed the world, is unclear.

There are less ambiguous arguments to be made against letting the RMB float. There are structural differences between US and Chinese economy which have been overlooked. The US economy is characterised by large national and multinational corporations which are well equipped to hedge currency risk. Therefore a floating exchange rate does not pose much of a threat to US industry. The Chinese economy, on the other hand, is more dependent on smaller household businesses that are reliant on a weak RMB. Any appreciation of the RMB could erode the small profit margins that traditionally characterise these industries.

Indeed, the infant industry argument has long been touted as justification for a fixed exchange rate in China. According to this theory, some newly established activities are initially high cost relative to the established foreign enterprises and it requires time for them to become competitive

(Krueger and Tuncer, 1982). Therefore protection, such as 'exchange rate protection', might be justifiable.

One of the real points of contention concerns China's hoard of foreign assets – in particular dollar denominated debt (Table 2). Any appreciation of the RMB could lead to a huge capital loss for China. As Paul Krugman (2009) puts it, China has gotten itself into a 'dollar trap'. He argues that 'China acquired its $2 trillion stash – turning the People's Republic into the T-bills Republic – the same way Britain acquired its empire: in a fit of absence of mind.' A further appreciation of the RMB could continue to erode China's external wealth. That said, from a global perspective, this would be a zero-sum game, with countries like the US benefiting the most. So it is clear that what is best for China is not necessarily what is best for the world. This is again a normative issue.

Fear of Floating?

In their influential paper, Calvo and Reinhart (2002) put forward the theory that nations have a 'fear of floating'. Countries that say they allow their exchange rate to float mostly do not. While the authors pinpoint a lack of credibility on the part of the monetary authority as the main motivation behind this fear, this appears less applicable in the case of China. However, they do point to other causes. They find that exchange rate volatility appears to be more damaging to trade in emerging markets; perhaps because trade is predominantly invoiced in US dollars and hedging opportunities are more limited. Indeed, they found that, despite the rationale for accommodating real terms of trade shocks, countries often choose not to. Fear of an abrupt exchange rate swing in these cases may dominate the perceived need to allow for the nominal and real exchange rate to adjust. If it is the case China too has a 'fear of floating' then, from a purely economic perspective, there is a case for allowing the RMB to float. To recall Friedman's (1953) argument, it is more efficient to allow one price, that is, the value of RMB, to change than to adjust a plethora of individual prices that make up the internal price structure of the Chinese economy.

Conclusion

The argument for floating the RMB is steeped in a myriad of political and economic issues. Leaving aside the political arguments, the empirical literature on the topic is also far from unanimous. While there are a number of empirical shortcomings, the RMB does appear to be somewhat undervalued relative to the USD and EUR. However, that doesn't necessarily imply that a free-floating currency would be the best policy for China, or indeed the world.

If one solid conclusion can be drawn from this analysis, it is that if China's peg is motivated purely by a 'fear of floating,' then there may be a clear economic rationale for moving to a more flexible exchange rate. Realistically, an incremental move towards a floating regime is in China and world's best interest so as to avoid the volatility associated with large swings in the exchange rate.

The Chinese authorities appear to hold the view that the benefits of floating outweigh the costs. The above comments from the Chinese Premier reinforce the understanding that the nation needs to rebalance growth. Central to achieving that is a more flexible exchange rate and an independent monetary policy. Despite temporarily re-pegging the RMB to the USD during the financial crisis, the RMB has been moving towards a float regime since mid-2005. Since the unofficial peg was lifted in June 2010, the RMB has appreciated 2.6%. Therefore, whether the debate has been settled in economic literature appears less relevant; China is moving, albeit slowly, toward a more flexible exchange rate regime.

There are certainly benefits of this movement for the world economy. In an economic downturn, when central bank's find themselves in a liquidity trap they often expand the money supply in order to devalue their currency. This has taken the shape of quantitative easing recently. While this advisable for asymmetric shocks to an individual nation, a shock as big and universal as the financial crisis makes such a policy a game of 'beggar-thy-neighbour', and ultimately amounts to a zero-sum exercise. Such policies have the potential to become a race to the bottom and mark an age of global 'exchange rate protection'. Given the economic consequences of protectionism is the 1930s, such an outcome could have a detrimental outcome for the world economy. Worryingly, talk of 'currency wars' has been making headlines in recent months and the political back-and-forth between the US and China has been intensifying that risk. While the political arguments have been omitted from this analysis, these developments have the greatest potential to dictate the flexibility of the RMB over the coming years.

References

Bosworth, B., 2004. *Valuing The Renminbi.* Washington, DC: The Brooking Institution.

Box, G. E. P. and Draper, N. R., 1987. *Empirical Model-Building and Response Surfaces.* Wiley.

Calvo, G. A. and Reinhart, C. M., 2002. 'Fear Of Floating'. *The Quarterly Journal of Economics*, MIT Press, 117(2), pp.379-408.

Cheung, Y., Chinn, M. and Fujii E., 2007. 'The Overvaluation of Renminbi Undervaluation', *Journal of International Money and Finance.* 26(5), pp.762-785.

Cheung, Y., Chinn M. and Fujii, E., 2009. 'China's Current Account and Exchange Rate'. In: R. Feenstra and S. Wei, eds. *China's Growing Role in World Trade.* Chicago University Press.

Feenstra, R. and Taylor, A. M., 2008. *International Economics.* Worth Publishers.

Frankel, J., 2005. 'On the Renminbi: The Choice Between Adjustment Under a Fixed Exchange Rate and Adjustment Under a Flexible Rate'. National Bureau of Economic Research. *Working paper no.11274*.

Funke, M. and Rahn, J., 2005. 'Just how undervalued in the Chinese renminbi?' *The World Economy*, 8(4), pp.465–489.

Goldstein, M., 2004. 'China and the Renminbi Exchange Rate'. In: C. Bergsten and J. Williamson, eds. *Dollar adjustment: how far? Against what?*. Washington, DC: Institute for International Economics.

Greenwood, J., 2008. 'The Costs and implications of PBC Sterilization'. *Cato Journal*, 28(2), pp.205-217.

Guo, K. and N'Diaye, P., 2009. 'Employment Effects of Growth Rebalancing in China'. IMF. *Working Paper no.09/169*.

Krueger, A. O. and Tuncer, B., 1982. 'An Empirical Test of the Infant Industry Argument. *American Economic Review*, 72, pp.1142-1152.

Krugman, P., 2009. 'China's Dollar Trap'. *The New York Times*, 02 April.

Makin, A., 2009. 'Is China's Exchange Rate Policy a Form of Trade Protection?'. *Business Economics*, 44(2), pp.80–86.

Friedman, M., 1953. *Essays in Positive Economic*s. Chicago: Chicago Press.

Mussa, M., 2007. 'IMF Surveillance over China's Exchange Rate Policy'. In: Peterson Institute of International Economics, *Conference on China's Exchange Rate Policy*. Peterson Institute, 19 Ocotber 2007.

Nair, M. K. V. and Sinnakkannu, J., 2010. 'De-pegging of the Chinese renminbi against the US dollar: analysis of its effects on China's international trade competitiveness'. *Journal of the Asia Pacific Economy,* 15(4), pp.470-489.

Obstfeld, M. and Rogoff, K., 2009. 'Global Imbalances and the Financial Crisis: Products of Common Causes'. In: Federal Reserve Bank of San Francisco, Asia Economic Policy Conference. Santa Barbara, CA, 18-20 October 2009.

Prasad, E. and Wei, S.-J., 2005. 'The Chinese Approach to Capital Inflows: Patterns and Possible Explanations'. IMF, *Working Paper no.05/79*.

Wolf, M., 2010. 'How to fight the currency wars with stubborn China'. *Financial Times*, 05 Ocotober.

Xiao, G., 2008. 'China's Exchange Rate and Monetary Policies: Structural and Institutional Constraints and Reform Options'. *Asian Economic papers*, 7(3), pp.31-49.

Xie, A., 2010. 'I'll Tell You When Chinese Bubble Is About to Burst'. *Bloomberg*, [Online] 25 April. Available at: http://www.businessweek.com/news/2010-04-25/i-ll-tell-you-when-chinese-bubble-is-about-to-burst-andy-xie.html.

WOULD EUROPE BE BETTER OFF WITHOUT THE EURO?

Mairead Gallagher

Senior Sophister

Mairead Gallagher addresses the question that is on many people's minds at the moment, but more importantly tackles the fundamental structural problems of the European Monetary Union. The essay serves to remind us that there are benefits to the single currency, and argues that addressing its flaws is a more constructive course to take than questioning its initial wisdom.

Introduction

In light of the current financial crisis, much speculation has surrounded the issue of whether 'Europe would be better off without the euro'? This paper seeks to address the issue but also proposes that the real question should in fact be whether collectively, individual euro area countries are prepared to take on board the lessons learned from the current situation and implement necessary institutional reforms (counter-cyclical fiscal policy, improved regulation and increased flexibility in labour markets) in order to realise the benefits of the European Monetary Union (EMU) rather than further handicap the area.

Before considering the implications of the currency union for the current crisis, I will review the pros and cons of the EMU in relation to 'optimum currency area' (OCA) criteria (Mundell, 1961). I will then follow with a review of the impact of the euro on Europe in terms of: improving the credit environment, sources of divergences in inflation across the euro area, effects on competitiveness and the role of fiscal policy as a stabilisation tool.

OCA Analysis of EMU

OCA theory focuses on integration and asymmetries. It attempts to weigh the relative costs and benefits of monetary union against one another to ascertain whether the region in question is an OCA. Essentially, increased integration raises the efficiency benefits while asymmetries increase the stability costs. Difficulty in quantifying the respective pros and cons hinders this task; yet the general consensus is that Europe is not an OCA. Nonetheless, one must acknowledge that in deciding whether

Europe would be better off without the euro, there are factors beyond the economic considerations mentioned in this paper which warrant attention – after all the euro is a political construct.

The main criteria cited for successful monetary union are (Frankel and Rose, 1998): trade integration, similarity of shocks and positive correlation between business cycles, labour mobility and fiscal transfers.

Loss of Monetary Policy Autonomy

The primary cost of monetary union arises from the impossible trinity – a fundamental idea in international economics which states that the following three policy objectives cannot be simultaneously achieved: monetary policy autonomy, capital mobility and a fixed exchange rate.

By definition a currency union implies fixed exchange rates and the Single European Act disallows capital controls; consequently, euro area members renege on the possibility of exercising independent monetary policy as an adjustment mechanism to shocks. The magnitude of costs resulting from monetary union depends on the degree of asymmetric shocks and the availability of adjustment mechanisms to adapt appropriately to such asymmetric shocks (discussed below). With independent monetary policy redundant, fiscal policy becomes increasingly important to euro area members, as does the need for labour mobility and the existence of fiscal transfers.

Trade

One of the initial arguments for EMU was the expected increase in trade resulting from the lower transaction costs associated with the euro. Frankel and Rose (1998) propose that the optimum currency area criteria are endogenous, such that EMU would promote intra-European trade which would in turn lead to more similar business cycles among members – increasing the benefits of EMU (as well as reducing asymmetric shocks). The counter-argument, as proposed by Krugman (1993) and others, is that as trade becomes more integrated, specialisation will occur which will cause more asymmetric shocks within the euro area.

Evidence to date does not support the specialisation argument; however, given that the pro-trade effect from the euro is estimated to be in the range of 5-15%, with 9% being the best estimate (Baldwin, 2006) it is difficult to support the assertion that business cycles will become more closely aligned due to the trade integration mechanism. Despite a near 10% increase in trade flows for euro area members, it does little to alleviate the pain of forgoing independent monetary policy, especially since non-member countries experienced about a 7% increase in trade due to the euro (Baldwin, 2006).

Common Shocks

It is widely accepted that in responding to common shocks the ECB delivers superior solutions than would be attained if monetary policy was conducted at individual country level. Eichengreen (2010) remarks that 'the last decade would have been very difficult for Europe without the euro. There would have been chaos in foreign exchange markets after 9/11, after the Madrid train bombings, and after Lehman Brothers.' It is conceivable that uncoordinated monetary policy would inadequately account for spillover effects arising from individual country level actions, possibly leading to excessive or insufficient policy responses.

Furthermore, monetary union has facilitated global cooperation in relation to common shocks. The ECB and Federal Reserve participated in coordinated policy interventions to stabilise the euro-dollar exchange rate in 2002 (Lane, 2009). Such cooperation would be too cumbersome to be timely and effective if decision making had to take place between the Federal Reserve and each individual euro area country. Yet, despite the success of the ECB in dealing with common shocks, the euro area is also subject to considerable asymmetric shocks which cannot be stabilised using monetary policy. The implications of these asymmetric shocks will be considered later.

Labour Mobility

Labour mobility provides a mechanism for adjusting to asymmetric shocks. If a country or region experiences a negative shock, unemployment will rise in response to a decline in output. With a mobile labour force, workers who become unemployed can migrate to other regions in search of work. This mechanism reduces the need for stabilisation policy, thus reducing the costs of monetary union.

However, according to Wyplosz (2006, p.215) 'the labour mobility criterion is not even remotely satisfied'. Evidence from 2003 shows that only 1.5% of people were born in a different EU country from the one in which they live, compared with over 30% of residents in the US living in a different state to which they were born (Feenstra and Taylor, 2008). This trend is not expected to have changed significantly as euro area labour mobility continues to be hindered by language and cultural barriers as well as further impediments associated with differing social welfare systems across the euro area. Without labour mobility, 'competitive disinflation' is necessary, which involves periods of sustained high unemployment and wage disinflation until competitiveness is regained and the current account deficit and unemployment are reduced (Blanchard, 2007). The process is prolonged if wages are subject to nominal rigidities; in this case re-establishing competitiveness requires either productivity growth above that in other regions and/or inflation to erode the real wage.

Competitiveness and the ability to stabilise shocks are seriously compromised by the preference for social protection in continental Europe. Without institutional reform, many euro area countries will endure the painful process of 'competitive disinflation' currently faced by Portugal, Ireland, Greece and Spain. The detrimental effects are evident in the case of Germany, whose growth was consistently lower than that of the euro area from 1995-2006 as it underwent this process (Blanchard, 2007).

Fiscal Transfers

The US is commonly used as a benchmark for EMU. One stark contrast between the euro area and the US is that the euro area does not provide insurance against asymmetric shocks. While fiscal federalism is a prominent feature of the US economic system no such feature exists in the euro area; rather fiscal policy is conducted independently at national level. Furthermore, the EU budget is only about 1% of EU GDP compared with about 33% for the US, which indicates the limited capacity of the euro area to finance fiscal transfers. Fiscal federalism acts to cushion asymmetric shocks by providing fiscal transfers from non-affected countries to affected countries, facilitating fiscal stimulus. Therefore, some centralisation of fiscal policy within the euro area could serve to alleviate the problems with asymmetric shocks.

However, there are several drawbacks from such an insurance system, including political resistance and issues of moral hazard. To circumvent moral hazard, the Maastricht Treaty explicitly included a 'no bailout clause'. However, that clause has been reneged upon in response to the current financial crisis, as the European Financial Stability Facility (EFSF) has been set up to provide funds to troubled euro area economies.

Improvement in Credit Environment

A significant benefit of adopting the euro, particularly for peripheral countries, was gaining access to the integrated euro area bond market; a much larger pool of funds from which to obtain finance than would have been available in their previously underdeveloped domestic bond markets. By entering the EMU and subsequently designating monetary policy to the ECB, euro area countries were able to import low inflation and therefore credibility. With monetary policy conducted by the ECB, whose commitment is price stability, inflation risk premiums fell significantly. Prior to joining the euro, countries such as Italy had a poor reputation with regard to devaluation and consequently faced large risk premiums on domestic denominated bonds.

Wyplosz (2006) recognises that even with a tight peg in a fixed exchange rate system, instances occur where the temptation to run inflationary monetary policy becomes too high to resist. In a currency union this possibility is eradicated, bringing clear benefits to the euro area by increasing credibility and thus eliminating speculative attacks on the currency (attacks are still possible on sovereign debt however).

The euro led to a substantial convergence in yield spreads across the euro area. The impact was largest in peripheral countries where spreads above the DM were significant prior to monetary union. These countries 'faced a permanent reduction in the cost of capital' (Lane, 2009, p.8). The sharp rise in the availability of long-term credit coupled with expectations that entering monetary

union would lead to growth and convergence with the 'core' economies translated into an increase in aggregate demand in the periphery – particularly in Greece, Ireland, Spain and Portugal.

Insulation From Potential Foreign Debt Crisis

The global savings glut in the early 2000s certainly contributed to the increase in the supply of credit. However, without the integrated euro markets the impact on the periphery would have been limited because pre-euro domestic equity and bond markets were too small and illiquid to attract substantial investors.

Furthermore, for countries with less developed securities markets the savings glut would have led to a rise in the accumulation of foreign currency debt. With respect to the current financial crisis, if national central banks had insufficient foreign reserves to provide liquidity to domestic banks a foreign debt problem, as well as an amplified banking crisis, may have ensued (Lane, 2009). The euro has to some extent protected Europe from a worse crisis, in that less financially developed European countries would have accumulated considerably more foreign debt if the euro and its integrated markets did not exist.

Diverging Inflation Rates and Implications for Competitiveness

Asset bubbles appeared in some national housing markets and were further inflated by expansions in bank credit, predominantly in Spain and Ireland. Sustained increases in demand caused inflationary pressure, which led to both wage and price increases in the booming euro area economies and in turn contributed to an eventual decline in competitiveness. It is clear that a 'one size fits all' approach to monetary policy is not appropriate in the euro area. With some economies booming and others almost stagnant, inflation differentials began to emerge between countries. Inflation across the euro area in 2000 ranged from 1.5% in Germany to 5.6% in Ireland, with a range of 2.4 percentage points over the period 1999-2004 (Lane, 2006).

In accordance with the Fisher Equation (Fisher, 1977), with all euro area members facing the same nominal interest rate (by definition of a fixed exchange rate), differences in inflation across countries leads to differences in the real interest rate each economy faces. Rather perversely, those who are experiencing higher than average inflation face a lower real interest rate thus stimulating the economy even further, while those with lower than average inflation face a higher real interest rate thus dampening growth in their economy. In this respect, the euro serves to amplify the differences in economic fundamentals between euro area economies. A self-correcting mechanism exists to offset divergences in growth and inflation, however, this process is more harmful to growth than taking preventative measures to maintain sustainability before becoming uncompetitive. The self-correcting mechanism occurs where prices and costs of factors of production rise to make the economy less competitive. This eventually leads to a slowdown and accompanying unemployment.

Traditionally an economy would engineer a devaluation of its currency to regain competitiveness. However, with the devaluation option no longer available and with many economies experiencing low productivity growth, the only alternative is to decrease nominal wages and prices of non-tradables (Blanchard, 2007). This is a painful process as previously noted; Germany pegged to the euro at an overvalued rate and thus suffered from the inability to compete with exporting countries. The German economy underwent a process of competitive disinflation which resulted in their economy growing at a consistently lower rate than that of the euro area. If euro area economies aren't prepared to implement necessary reforms to labour markets and fiscal systems then the costs of monetary union may overshadow the benefits.

Heterogeneity between euro area members is further augmented by the single currency in that different countries have different sensitivities to fluctuations in the euro exchange rate, as determined by their trade linkages. For example, Ireland trades much more extensively with the US and the UK (as a % of GDP) than other members of the euro do. Therefore, fluctuations in the euro exchange rate impact Ireland's ability to export, which feeds back into home demand through the impact on real wages in Ireland and in turn has implications for inflation.

If euro members continue to harbour such differences in fundamentals it is possible that Europe would be better off without the euro unless adequate stabilisation tools can be employed. Accordingly, fiscal policy has become increasingly important for moderating the effects of business cycles.

Fiscal Policy as a Stabilisation Tool

The availability of cheap credit enabled countries to run current account deficits and perhaps even encouraged pro-cyclical fiscal policy. Many countries justified their current account deficit because of expectations of higher growth and a lower real rate of interest (Blanchard, 2007, p.4). However, hindsight tells us that the level of aggregate demand in these economies was unsustainable, particularly in view of the housing bubbles.

Furthermore, markets did not sufficiently penalise countries running deficits. Market failure occurred because investors believed that euro area securities were equally risky; there was a general failure to properly account for default risk from individual countries. 'An increase in the deficit by 1 per cent of GDP raises the spread by 4 basis points for a non-euro area country but only by 1.5 per cent for a euro area member' (Eichengreen, 2007, p.22). These market failures could be attributed to a lack of belief in the 'no bailout clause' in the Maastricht Treaty. Irrespective of market failures, peripheral countries should have paid heed to the large volume of literature which reiterates the importance of running 'sufficiently large surpluses during boom periods in order to finance the loss of revenue and increased spending commitments during downturns' (Lane, 2006, p.16). Overheating could have been abated to some extent using counter-cyclical fiscal policies, which would also have enabled national authorities to form a 'rainy day fund.'

While in some respects monetary union may have aggravated the current situation for Europe, forewarnings regarding the need for counter-cyclical fiscal policies were abundant and ultimately, fiscal policy responsibility lies in the hands of national policymakers. Given the current economic situation, the question to be asked is not so much whether Europe would be better off without the euro but whether euro area countries would be better off designating control of fiscal policy to some independent institution.

The constraints imposed by the Stability and Growth Pact may have encouraged pro-cyclical fiscal policy by inhibiting the degree of expansionary fiscal policy possible during downturns, yet not imposing rules regarding the level of surpluses required during upswings. Many euro area countries now find themselves locked out of the bond markets – a consequence of imprudence in the good years combined with recent contagion effects. As a result, despite the fact that countries such as Germany, Netherlands and Austria proceeded more cautiously, they now have to contribute to the bailout packages for weaker euro area economies.

Exiting the Euro

Euro area members are currently stuck between 'a rock and a hard place' (The Economist, 2010). The weaker economies may find the option to leave the euro and devalue their currency appealing, particularly as they are faced with years of austerity to complete the disinflation necessary to become competitive again. On the flip side, stronger creditor countries may wish to exit the euro to avoid subsidising countries with excessive deficits and unsustainable debt to GDP ratios.

Effects of exiting differ between economies. In the case of weaker economies, investors would anticipate that they will run inflationary monetary policy which would result in credit-rating downgrades, higher risk premia and larger sovereign spreads; access to bond financing is unlikely to improve by leaving the euro (Eichengreen, 2007). As for stronger economies, the costs of leaving are more political in nature, in that domestic exporters would be severely harmed by currency appreciation, perhaps causing political backlash. Regardless of which type of economy exited, capital flight would ensue, further jeopardising the future of the euro.

Perhaps the most important reason to remain within the currency union is that, in abandoning the euro, a country may be denied the privileges of the single market (Eichengreen, 2007). While it is not clear that this would happen, it is a very big risk to take – particularly for exporting nations.

Conclusion

From the initial review of the euro area in relation to OCA criteria it is clear that there is insufficient flexibility within the euro area to absorb asymmetric shocks. The costs of these shocks are heightened by a lack of labour mobility and an absence of adequate insurance transfers. Furthermore, there is a lack of empirical evidence to support the view that trade integration results in greater

homogeneity between euro area business cycles. With frequent or sustained asymmetric shocks, all else equal, Europe would be better off without the euro. Quite simply, the costs of monetary union might be too high given the limited flexibility.

However, all else is not equal. Monetary union has brought great gains to Europe in terms of improving the long-term credit environment and attracting foreign direct investment (see Lane, 2006), which have facilitated growth and a significant degree of convergence across countries. Moreover, the level of financial integration achieved through monetary union has insulated many European countries from possibly worse crises resulting from foreign debt problems and following speculative attacks. Without the euro it is conceivable that there would be chaos in the foreign exchange markets due to defaults and speculative attacks, which would have a huge impact on trade (with the euro, attacks are limited to sovereign debt).

It is evident from the persistent divergences in inflation that a 'one size fits all' monetary policy is inappropriate. However, the effects of these divergences could be ameliorated through labour mobility, national authorities running more counter-cyclical fiscal policies and by the development of a fiscal transfer system.

Whether Europe would be better off without the euro remains indeterminate. For some countries it is very costly to remain in the currency union, particularly in view of the austerity measures necessary to regain competitiveness (or on the other hand the cost of funding bailouts). Future euro area prosperity is conditional on institutional reform in financial regulation, fiscal policy and labour markets.

References

Baldwin, R. E., 2006. *In or Out: Does It Matter? An Evidence-Based Analysis of the Euro's Trade Effects*. London: Centre for Economic Policy Research.

Blanchard, O., 2007. 'Adjustment within the euro. The Difficult Case of Portugal'. *Portuguese Economic Journal,* 6(1), pp. 1-22.

Eichengreen, B., 2007. 'The Breakup of the Euro Area.' National Bureau of Economic Research, Cambridge, MA. *Working Paper NO. 13393.*

Eichengreen, B., 2010. 'The Breakup of the Euro Area'. In: *TheBrowser*, [Online] 22 November. Available at: http://fivebooks.com/interviews/barry-eichengreen-on-euro [Accessed: November 29, 2010]

Feenstra, R. C. and Taylor, A. M., 2008. *International Economics*. New York: Worth Publishers.

Fisher, I., 1977. *The Theory of interest*. Philadelphia: Porcupine Press.

Frankel, J. and Rose, A. K., 1998. 'The Endogeneity of the Optimum Currency Area Criteria'. *The Economic Journal.* 108(449), pp. 1009–1025.

Lane, P. R., 2006. 'The Real Effects of European Monetary Union'. *The Journal of Economic Perspectives,* 20(4), pp. 47–66.

Lane, P. R., 2009. 'European Monetary Union and Macroeconomic Stabilisation Policies in Ireland'. Trinity College Dublin and CEPR, *IIIS Discussion Paper*. London: Mimeo.

Mundell, R., 1961. 'A Theory of Optimum Currency Area'. *The American Economic Review*, 51(4), pp.657-665.

The Economist, 2010. 'The Future of the Euro'. *The Economist*, [Online] 2 December. Available at: http://www.economist.com/node/17629661 [Accessed: December 4, 2010]

Wyplosz, C., 2006. 'European Monetary Union: The Dark Sides of a Major Success'. *Economic Policy,* 21(46), pp.207–261.

THE GOLD STANDARD AND THE GREAT MONETARY DEPRESSION: HOW THE MONETARY SYSTEM CONTRIBUTED TO TURNING AN 'ORDINARY SLUMP' INTO THE GREATEST ECONOMIC CONTRACTION OF THE 20TH CENTURY

Amandine Lobelle

Senior Sophister

Recent calls by Robert Zoellick, the President of the World Bank, to reintroduce the gold standard as part of the global currency reserve system have brought the almost forgotten currency regime back into conversation. Amandine Lobelle examines the detrimental effects that being pegged to the gold standard had on the Great Depression, arguing that it made what would have simply been an ordinary downturn a colossal one.

Introduction

There has been a central, recurring theme in economics, from Cantillon to modern day writers, that gold is an ideal monetary standard, domestically and internationally, due to its qualities as a standard of value and a medium of exchange. The renowned classical gold standard (1870s-1914), seen as a 'seal of approval' (Bordo and Rockoff, 1996) amongst the international community, was built on this premise, whereby a certain mass of gold defined the monetary unit and served as the 'ultimate medium of redemption'[1] (White, 2008, p.2).

Similarly, there has been a fixation amongst academics to understand the Great Depression, the ability of which has been hailed as the 'Holy Grail' of macroeconomics (Bernanke, 1995). Many academics spend their entire professional careers on determining the roots of the crisis and the long-lasting impacts it had on 20th century monetary thought and policy. This essay will analyse recent theories attributing the interwar gold standard, which existed from the mid-1920s to the mid-1930s, as a watershed that promulgated an ordinary slump into the Great Depression.

The interwar gold standard was reinstated following the massive exogenous shock of World War I (WWI). However, it was not to operate in the same environment. The repercussions of the

[1] In the United States, one ounce of pure gold was equivalent to US$20.67 (White, 2008, p.2).

Great War meant that the 1920s were characterised by domestic political constraints in the following forms: unions and interest groups; international political disputes over war debts and reparations; hyperinflation; unemployment; and incompatible conceptual frameworks in different countries which prevented policymakers from developing a common understanding of the economic problem. Eichengreen (1995) argues that these factors were obstacles to the crucial international cooperation required after WWI to rebuild a suitable gold standard. Furthermore, as Ahamed (2010) points out, during the transitory 1920s international communication was not easy and central banks were still privately owned; their key objective was to preserve the value of the currency rather than stabilise the economy.

The Wall Street Crash and post-1929 recession that ensued was one of the greatest economic shocks of the 20th century; industrial production fell by 30% in the U.S., 25% in Germany and 20% in Britain (Ahamed, 2010, p.374) and between 1929 and 1933, unemployment rose from about 3% to nearly 25% in the US (Bernanke, 2004). Eichengreen (1995, p.4), however, maintains that had a *stable* gold standard been maintained, the post-1929 recession would have been just another economic contraction. In reality, the badly constructed system and its consequent collapse was another source of financial instability during the early 1930s. Not only did it restrict the monetary policy that governments and central banks were able to enact, but it allowed for volatile capital and gold flows to occur, which widened the gap between surplus and deficit countries and undermined the solvency of financial institutions. This uncertainty, combined with the fact that central banks were not cooperating to improve the system and had begun breaking the 'rules of the game', precipitated speculative attacks on the major reserve currencies which had a deflationary feedback effect on other gold standard countries. Indeed, no country could begin recovering from the recession whilst still being tied to the gold standard.

This essay will expand on the themes outlined above. Essentially, domestic and international finance was intimately linked through the gold standard which required high degrees of credibility and international cooperation to survive. Unfortunately, these requisites were not to hold for the early 1930s which would have dire repercussions on the world economy.

The Gold Standard Restricted Policymakers' Abilities to Combat Deflation and Economic Contractions

Liaquat Ahamed's recent book *The Lords of Finance* (2010) highlights how the decisions of the leaders of the four most important central banks of the time (Britain, France, United States and Germany) and the decisions that they took while in office were key to the Great Depression. Although human error and a lack of communication and cooperation were big factors in translating the downturn into a depression, it should be noted that countries which were bound to the fixed exchange rate monetary system could not depreciate their currency; a tool which Eichengreen (1995, p.21)

believes was crucial to macroeconomic growth, through increases in output, employment, investment and exports**.**

The physical aspect of this restriction, as posited by Bernanke and James (1991), is that the majority of European central banks had insufficient powers and were limited in terms of the open market operations they could use as a result of stabilisation programmes in the 1920s; they could only rely on discount policies (i.e. interbank lending rates) to affect the money supply instead of, for example, the buying and selling of government securities. Since major commercial banks borrowed quite infrequently from their central banks, their control over the money supply was quite weak and unable to deal with severe macroeconomic contractions (Bernanke and James, 1991, p.10). Furthermore, being on the managed gold standard forced countries to adopt similar discount rate policies: if one country decided to increase their interest rate, other countries would have no option but to retaliate and increase their rate too. Failure to do so could risk the loss of gold reserves as financial investors transferred their funds to countries where returns were higher (Bernanke, 2004). Eichengreen (1995) even attributes the worldwide downturn of 1929-1930 to the degree of monetary policy interconnection between countries on the gold standard.

In addition to physical restraints and obligations, the mental paradigm of the time also restricted policymakers. The consensus ideology was that the gold standard should be saved at all costs. Indeed, policymakers perversely believed that macroeconomic impairments such as unemployment would stabilize and output would resume if the gold standard was maintained, while any attempts to increase employment directly would fail (Eichengreen and Temin, 2000, p.195). This is possibly why, despite dealing with grueling deflation, output losses and mass unemployment, it took most of the Western powers (who seemed to idly sit by as their economies plummeted) several years to abandon the gold standard, with Switzerland being the last to leave in 1936.

Hamilton (1988) argues that the speculative attack on the dollar in 1931 was caused by the restrictive monetary policy of the Federal Reserve in the second half of 1931. In fact, 'on October 9 1931, the Reserve Bank of New York raised its rediscount rate to 2.5% and on October 16 to 3.5% - the sharpest rise within so brief a period in the whole history of the system, before or since' (Friedman and Schwartz, 1963, p.317). This contractionary policy in the midst of rapid economic decline was characteristic of the gold standard mentality.

The Gold Standard Broke Down the Price-Specie Flow Mechanism and Encouraged Volatile Capital Flight

Given that the gold standard was a managed system, there was no automatic check to redistribute gold reserves between member countries; hence Hume's automatically adjusting price-specie flow mechanism did not hold. Theoretically, under a gold standard with fixed exchange rates, a contractionary monetary policy in one country, such as the United States, should be matched by gold

inflows to neutralize the effect of the Federal Reserve's actions on the money supply. However, Hamilton (1987) specifies three reasons why the price-specie flow mechanism did not hold from 1929-1931. Firstly, the money supply was falling not just in the United States but all over the world, partially due to a shortage of gold. Secondly, even though the discount rate fell in 1929-1930, bills discounted fell faster following a failure to increase un-borrowed reserves: despite the inflow of gold, high powered money fell by 5% to the effect that the Federal Reserve was in fact sterilizing gold inflows and further contracting high powered money. Thirdly, the collapse of world trade following the rise of trade barriers such as tariffs, quotas, and domestic content laws was also a disruption to the price-specie flow mechanism (Hamilton, 1987, pp. 159-160).

This failure in the price-specie flow mechanism led to the build-up of gold reserves by certain countries and furthered the already-existing balance of payments disequilibrium and financial instability. In fact, France and the United States, two surplus countries, were particularly guilty of augmenting this asymmetry. By 1932, the two countries owned 70% of the world's stock between them (Cesarano, 2006, p.55). Nurkse (1985, p.213) even attributes the buildup of French gold imports to the breakdown of the gold standard. An unavoidable consequence of this was that other (deficit) countries were forced to adopt restrictive measures to defend their gold reserves which in turn widened the asymmetry, accentuated a deflationary bias in the system, worsened the economic climate and hastened the collapse of the gold standard; surplus countries essentially shifted the burden of adjustment to deficit countries by forcing them to deflate.

Furthermore, uncertainty about the soundness of banks following the worldwide banking panics of 1929 and 1930 reinforced fears of exchange-rate devaluation and was a source of mounting financial instability and a fall in confidence in domestic banking systems during the early 1930s. Both of these factors precipitated volatile capital flight; expectations of devaluations triggered outflows of 'hot-money'[2] deposits (as well as those held by domestic depositors) and a fall in confidence in the banking system often led to a flight of short-term capital from the country, draining international reserves and threatening convertibility (Bernanke, 1995, p.7). These large and sudden capital flows were exacerbated by the gold standard because 'speculators necessarily will continuously reappraise the profit opportunities arising from any potential change in gold parity' (Hamilton, 1988, p.68); the gold standard essentially further undermined the solvency of financial institutions. By 1931, it became evident that the Federal Reserve and the Bank of England had 'succumbed to the lure of managed money', had begun breaking the 'rules of the game' and had committed abuses of credit by sterilising international gold flows[3] (Eichengreen and Temin, 2000, p.195). This further prevented the capital flows from exerting their normal stabilising influence on credit conditions and respectively prevented costs and prices from adjusting. Ferderer and Zalewski (1994, p.836) argue that this loosening of

[2] Short-term deposits held by foreigners in domestic banks (Bernanke, 1995, p.7).
[3] An example of such sterilisation of gold flows could be selling government securities on the open market.

credit conditions and the withdrawal of highly mobile foreign deposits led to further bank failures and attacked the basic premise of the gold standard, the ability of a country to convert its currency to gold.

Capital Flight Under the Gold Standard Resulted in Speculative Attacks on Currencies

Following the emergence of volatile capital and gold flows outlined above, in June 1931, the government commissioned Macmillan Committee called a 10% devaluation of the gold parity of the sterling and laid out a detailed plan for Britain to do so within the confines of the gold standard (Macmillan Committee on Financy and Industry, 1931, p.195-196). Unfortunately, such plans were not to bear fruit. Following the financial upheaval and bank runs in Continental Europe (notably that of Credit Anstalt in Austria), the soundness of Britain's international long-term loans had been called into doubt, which led speculators to attack the British pound (Bernanke, 2004). However, this may not have been the only incentive for a speculative attack. As Hamilton posits, it may be that the most persuasive argument simply lies in the fact that, ex-post, we can see that anyone who bet against the pound during the summer of 1931 profited very well (Hamilton, 1988, p.74). The nature of speculative attacks is such that as rumours of the devaluation spread, the damage was self-fulfilling; over the summer of 1931 capital began to flow out unchecked by the Bank of England who hesitated to raise the Bank rate out of fear of the damage this could cause to an already depressed economy (Ferderer and Zalewski, 1994, p.836). Essentially her unwillingness to raise interest rates left Britain with no other choice but to abandon the gold standard in September 1931, allowing the pound to float freely and have its value determined by market forces.

It is difficult to imagine the paradigm-shift that must have occurred following Britain's departure from gold. Although smaller countries had left the gold standard previously, these had been peripheral economies that were not at the heart of international trade and finance. For a core country and leading (albeit already severely depleted) economic power to actually abandon the mechanism that had governed policymaking throughout the downturn threw an already unstable international financial system into turmoil. As expected, this financial uncertainty and speculative activity would spread to other countries and eventually force other countries to abandon gold convertibility. In fact, although the United States remained on the gold standard until 1933, 'between September 16 and October 28, 1931, Federal Reserve holdings of gold fell from $4.729 billion to $4.002 billion' (Hamilton, 1988, p.74). This meant that the Federal Reserve had to pursue extreme measures in the discount rate policy, to reinstate that its commitment to gold was credible. It also meant that the United States was again susceptible to large and sudden capital flows which again increased uncertainty and deflationary pressures in other countries. In summary, as Hamilton argues, 'private speculators' potential profit opportunities should be regarded as an irresistible force in international finance, and any central bank policy that leaves open such opportunities is doomed to failure' (Hamilton, 1988, p.69). Failure, for the Bank of England, was being forced out of the gold standard.

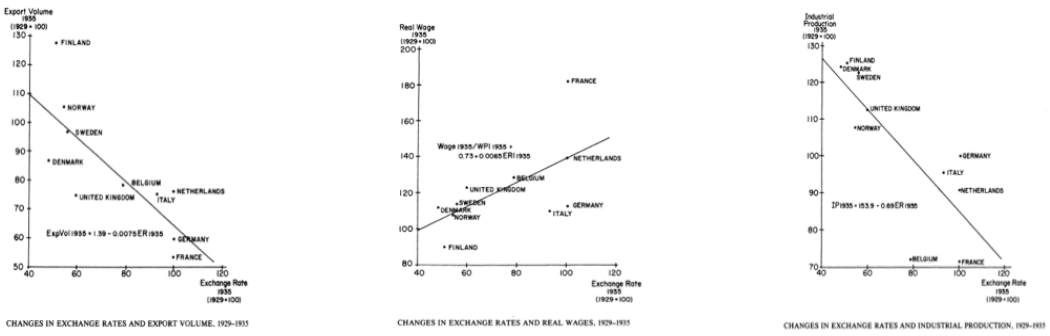
The Gold Standard Impeded Grounds for Economic Recovery

Perhaps the most potent argument for how the gold standard affected the onset of the Great Depression is that it impeded grounds for recovery. Temin (1993) goes as far as to argue that the gold standard was *the* primary transmission mechanism of the Great Depression: 'the single best predictor of how severe the Depression was in different countries is how long they stayed on gold. The gold standard was a Midas touch that paralyzed the world economy.' (Temin, 1993, p.92) Indeed, the gold bloc member countries were to endure severe contractions that lasted into 1935 and 1936.

Bernanke (1995, p.4) states that no country exhibited significant recovery whilst remaining on the gold standard, and there have been various econometric analyses to support this hypothesis. He found that subsequent to 1931 or 1932 there was a sharp divergence between countries which remained on the gold standard and those who left it; the difference in real wage growth was equivalent to about 6 percentage points per year (Bernanke, 1995, p.21). This discrepancy was attributed to the fact that countries which had left the system had greater freedom to initiate expansionary monetary policies, even if there was a 6 month time lag before expansionary policies were applied (a necessary interlude to convince the public and policymakers that abandoning gold was not going to cause inflation) (Eichengreen, 1995, p.393). Regaining monetary autonomy severed the link between the balance of payments and the price level to allow countries to lower interest rates or expand production without the onslaught of a currency crisis.

Eichengreen and Sachs (1985) found that countries were only able to begin recovering after leaving the gold standard with their study on the variance in industrial production, real wage and export volume. The data collected is from 10 European countries, consisting of the gold bloc countries still on the gold standard and those that had left and devalued by 1935. They find that gold bloc countries systematically had high wages and lower levels of industrial output than those who abandoned the gold standard, relative to their respective 1929 levels, as shown in the graphs below.

Figure 1: Changes in export volume, real wages and industrial production vis-à-vis exchange rates for 10 European countries 1929-1935

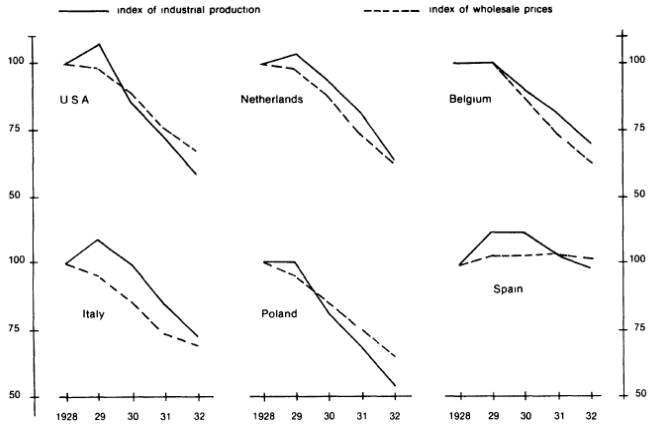


(Eichengreen and Sachs, 1985)

This analysis essentially shows the macroeconomic effect of depreciation once countries had left the gold standard (Eichengreen and Sachs, 1985). For all three macroeconomic variables in the graphs above, the gold bloc countries had not regained their 1929 levels in 1935. France in 1935, for example, relative to 1929 levels, had an export volume of c. 54%, real wages were c. 185% and industrial production was c. 72%. Conversely, Britain[4], who had left four years prior, was experiencing an export volume of c. 75%, a wage rate of c. 123% and industrial production of c. 113% relative to her 1929 levels. The graphs also show that countries which devalued early (Britain, Denmark, Sweden, Norway and Finland) grew more rapidly. There also seems to be a positive relationship between the magnitude of depreciation and the rate of growth, possibly due to the first mover advantage of regaining autonomy and depreciating their currencies. This final thought can be extended to the fact that Argentina and Uruguay, who suspended gold convertibility in 1929, largely avoided the financial crisis (Temin, 1993, p.96).

The final study by Choudhri and Kochin (1980) posits that flexible exchange rates insulate domestic output and prices from the influence of foreign output and price disturbances. To prove this, they perform a counterfactual evaluation by comparing gold bloc countries to Spain, who never joined the international monetary system. The graph of their results (below) shows that, despite substantial differences in industrialisation and trade, the impact of the severe contraction in the United States and other gold standard countries was not felt by Spain, who actually maintained a pretty consistent wholesale price index and whose industrial production output in 1932 was similar to in 1928. Additionally, this suggests that a contagion effect was in play; the gold standard countries' output and prices mirrored that of the United States suggesting that the gold standard was the binding mechanism

Figure 2: The Behaviour of Industrial Output and Wholesale Prices During 1928-1932: The Gold Bloc Countries and Spain Compare to the U.S.



[4] Britain's exchange rate of 59 indicates a 41% depreciation compared to 1929 levels.

(Choudhri & Kochin, 1980, p.569)

Conclusion

The purpose of this essay has been to show that the gold standard was a crucial instrument in propelling the ordinary downturn of 1929 into the Great Depression, a colossal shock in economic history and one which gave birth to a new field of macroeconomics, still revered and being used today. The fact that the gold standard was conceived in economically and politically turbulent times, and with very little international cooperation, meant that there were inherent structural weaknesses which added to the mounting financial instability of the early 1930s. In particular, the contradiction of having a *managed* system of gold exchange convertibility allowed certain surplus countries to take advantage of their balance of payments positions by hoarding gold reserves at the expense of other member countries. The distortion this created in the international economy, and the fact that deficit countries such as Britain were left without enough gold reserves to defend themselves against speculative attacks, undermined the credibility of the gold standard which proved to be detrimental for all in the long run. As Hamilton (1988) puts it, the interwar gold standard was not 'as good as gold' but just as good as the *credibility* of the government's promise to maintain gold convertibility.

Furthermore, not only were policymakers from gold standard countries restricted to the 'rules of the game', but the semi-religious strength of the mental paradigm was such that their lacklustre reactionary policies seem almost laughable today. By 1931, when Britain let go of her commitment to gold, the system had lost the remains of its international reputation which resulted in further deflation and macroeconomic losses for the gold bloc countries over the following four years. In hindsight, we can see that the gold standard was the ballast to economic recovery in the early 1930s; no country could begin to recover while still tied to gold, and the last to leave the monetary system were the worst hit. In fact, those who never joined the monetary system to begin with were largely exempt from the depression.

The early 1930s experience with the gold standard is exemplary of the macroeconomic 'Impossible Trinity' whereby only two of capital mobility, independent monetary policy and fixed exchange rates can be met. With fixed exchange rates and free capital flows, central banks could not exert autonomous monetary policy; when they tried to and began breaking the 'rules of the game', the system collapsed. Today, the key objectives of central banks are generally the price level, output and sometimes unemployment rather than the value of the currency. Nevertheless, the current economic crisis has led to calls, notably by Robert Zoellick, President of the World Bank (Zoellick, 2010, cited in Financial Times, 2010), for a return to some form of gold-backed monetary system in the attempt of regaining economic stability. While some degree of international reform is definitely in order, history shows us that the limitations and contradictions of classical gold standard variants are not to be underestimated.

References

Ahamed, L., 2010. *Lords of Finance.* London: Random House.

Bernanke, B. S., 2004. 'Money, Gold, and the Great Depression'. Washington and Lee University, Lexington, Virginia, *H. Parker Willis Lecture in Economic Policy.*

Bernanke, B. S., 1995. 'The Macroeconomics of the Great Depression: A Comparative Approach'. *Journal of Money, Credit and Banking ,* 27(1), pp.1-28.

Bernanke, B. S., and James, H., 1991. 'The Gold Standard, Deflation, and Financial Crisis in the Great Depression: An International Comparison'. In: R. G. Hubbard, ed., *Financial markets and financial crises*. Chicago: University of Chicago Press & National Bureau of Economic Research. pp. 33-68.

Bordo, M. D. and Rockoff, H., 1996. 'The Gold Standard as a "Good Housekeeping Seal of Approval"'. *The Journal of Economic History ,* 56(2), pp.389-428.

Cesarano, F., 2006. *Monetary Theory and Bretton Woods: The Construction of an International Monetary Order.* New York: Cambridge University Press.

Choudhri, E. U. and Kochin, L. A., 1980. 'The Exchange Rate and the International Transmission of Business Cycle Disturbances: Some Evidence from the Great Depression'. *Journal of Money, Credit and Banking ,* 12(4), pp.565-574.

Eichengreen, B., 1995. *Golden Fetters: The Gold Standard and the Great Depression, 1919-1939.* New York: Oxford University Press.

Eichengreen, B. and Sachs, J., 1985. 'Exchange Rates and Economic Recovery in the 1930s'. *The Journal of Economic History* , 45(4), pp.925-946.

Eichengreen, B. and Temin, P., 2000. 'The Gold Standard and the Great Depression'. *Contemporary European History ,* 9(2), pp.183-207.

Ferderer, J. P. and Zalewski, D. A., 1994. 'Uncertainty as a Propagating Force in the Great Depression'. *The Journal of Economic History* , 54(4), pp.825-849.

Financial Times, 2010. 'Zoellick Seeks Gold Standard Debate'. *Financial Times*, [Online] 7 November. Available at: http://www.ft.com/cms/s/0/eda8f512-eaae-11df-b28d-00144feab49a.html#axzz1FlrSFAqe

Friedman, M. and Schwartz, A. J., 1963. *A Monetary History of the United States, 1867-1960.* Princeton: Princeton University Press.

Hamilton, J. D., 1987. 'Monetary Factors in the Great Depression'. *Journal of Monetary Economics ,* 19, pp.145-169.

Hamilton, J. D., 1988. 'Role of the International Gold Standard in Propagating the Great Depression'. *Contemporary Policy Issues ,* 6(2), pp.67-89.

Macmillan Committee on Financy and Industry, 1931. 'International Gold Standard Report'. In: B. Eichengreen and M. Flandreau, eds., 1985. *The Gold Standard in Theory and History*. New York & London: Methuen. pp. 185-199.

Nurkse, R., 1985. 'The Gold Exchange Standard'. In: B. Eichengreen and M. Flandreau, eds., 1997. *The Gold Standard in Theory and History*. New York & London: Methuen. pp. 201-227.

Temin, P., 1993. 'Transmission of the Great Depression'. *The Journal of Economic Perspectives* , 7(2), pp.87-102.

White, L. H., 2008. 'Is the Gold Standard Still the Gold Standard among Monetary Systems?' Cato Institute, *Briefing Papers,* 100, pp.1-8.

STANDING AT THE ABYSS: MONETARY POLICY AT THE ZERO LOWER BOUND

Barra Roantree

Senior Sophister

The current economic crisis has seen a variety of tools used to avert catastrophe, monetary policy being one of them. However, central banks are limited in how much they can increase the supply of money by the zero-lower-bound. Barra Roantree discusses this limitation and outlines the possible alternative policy solutions that can overcome it, including such drastic measures as placing a negative interest rate on physical currency, or even abolishing it altogether.

Introduction

The high levels of inflation experienced in the post-war era resulted in the demotion of the zero-lower-bound (ZLB) from standard economic textbooks and the concept merely being treated by 'a few generations of economists ... [as] a relic of the Depression era' (Bernanke, Sack and Reinhart, 2004, p.5). Following the success of monetary policy in bringing inflation under control, from a worldwide average of 14% in the early 1980s to 4% within twenty years (Goodfriend, 2007), concerns that Japan's 'Lost Decade' could prove to be a more widespread phenomenon lead to the resurgence of interest in the ZLB among professional economists and policy makers alike. With advanced economies suffering the biggest collapse since the Great Depression, nominal interest rates at historic lows and the prospect of sustained deflation, the ZLB has once again become a very real constraint on monetary policy.

This essay explores the operation of monetary policy at the ZLB and the adequacy of the existing monetary policy tool-kit. The theory behind the ZLB and why it poses a problem for the conduct of monetary policy is outlined first. These problems are then illustrated in the context of the ongoing 'Great Recession'. The bulk of the essay is given to consideration of alternative policy responses to the ZLB, before concluding.

The Zero Lower Bound

Over the latter half of the 20th century, a remarkable consensus has built up regarding the

perceived optimal conduct of monetary policy. Independent central banks should target a low but positive level of inflation, operating in a transparent manner with frequent communication to maximise credibility (Goodfriend, 2007). This should be achieved by controlling interest rates, which influence consumption and investment decisions, across the term structure utilising the monetary authority's position as the monopoly supplier of narrow money. The monetary authority should stabilise macroeconomic fluctuations, lowering the real interest rate during times of recession to stimulate demand and investment by increasing the quantity of reserves it lends to commercial banks (and vice-versa when the economy is overheating).

However, the ability of central banks to engage in macroeconomic stabilisation is constrained by the nominal rate of return on currency: zero. This is because lenders and bank depositors will prefer to hold currency if nominal rates of return on non-currency assets are below zero. Strictly speaking, the validity of a lower bound at zero 'depends upon the assumption that it is costless at the margin to store money (the economy's medium of exchange)'. While this assumption is incorrect, the costs of storage amount to a few basis points at most and so have a negligible quantitative effect (McCallum, 2000, p.901). The ZLB is pulled below zero by the carry costs of money or any non-pecuniary benefits of non-money assets, and pushed above zero by non-pecuniary benefits of money, such as the desire to engage in anonymous exchange (Yates, 2004). The implication for monetary policy remains: there is a lower bound close to zero below which the key interest channel of monetary policy is ineffective.

Even with nominal interest rates floored at zero, the monetary authority can still exercise some control over real interest rates to the extent that they can increase expectations of inflation (following from the ex-ante Fischer equation). Krugman (1998) outlines an optimal response for a central bank constrained by the ZLB where it 'credibly promises to be irresponsible'; that is to seek a higher future price level by continuing to hold nominal interest rates at zero beyond the duration of the recession. The higher expected inflation lowers ex-ante real interest rates, providing a boost to consumption and investment.

Such commitments lack credibility however, in a reversal of the classic time-inconsistency problem outlined by Kydland and Prescott (1977) where central banks are unable to credibly commit to low inflation given policy discretion. Rather, in a 'liquidity trap, the problem is that the markets believe that the central bank will target price stability, given the chance, and hence that any current monetary expansion is merely transitory' and ineffective (Krugman, 1998, p.139). The consequence of this is that real interest rates remain too high to stabilise economic output. More seriously, if the economy enters deflation, the real interest rate will increase, further reducing economic activity and leaving the monetary authority with no ability to reduce rates. There is the the potential that this 'liquidity trap', where rates across the yield curve are at zero, could lead to a deflationary spiral. This is a vicious circle where a negative economic shock causes deflation, which raises real interest rates, leading to a contraction in economic activity, prolonging expectations of deflation.

In short, the ZLB arises from the negligible carry costs of currency and has serious consequences for macroeconomic stabilisation. The ZLB implies an inefficient floor on the critical interest-rate channel of monetary policy, while time-inconsistency hampers central banks' efforts to lower real interest rates through the 'Fischer Channel'.

Standing at the Abyss: Responses to the Great Recession

The policy response to the recession following the 2008 financial crisis was unprecedented. Monetary policy rates across the developed world hit record lows, while the balance sheets of central banks expanded to previously unimaginable levels. The balance sheet of the ECB increased from 10% to 19% of Eurozone GDP between 2008 and 2009, while the corresponding increase was 6% to 15% for the Federal Reserve and 6% to 16% for the Bank of England (BoE). The dramatic nature of the response aimed to address the equally dramatic scale of the crisis. The collapse of Lehman Brothers in September 2008 compounded the financial turbulence that had been ongoing since late 2007, with heightened risk premia leading to the interbank market seizing up. Spreads between the EURIBOR and the overnight swap rate - a measure of risk premia - rose from near 0 to over 180 basis points (ECB, 2010, p.61). This resulted in funding problems for banks, who had become dependent on the interbank market for short term funding.

In response, central banks were quick to cut their policy rates, acting with unparalleled cooperation, and by spring of 2009 the Fed, ECB and BoE had cut rates close to 0%. With limited scope to further pursue traditional monetary policy, central banks began to implement a series of unconventional measures. While such measures differed in detail between various central banks, they can be broadly characterised as quantitative easing (QE) and qualitative, or credit easing (CE). QE involves the expansion of the Central Bank's balance sheet without change in the composition, while CE is a change in the composition, but not the size of the balance sheet (Lenza, Pill and Reichlin, 2010). While such a characterisation can be useful in classifying individual measures, Bernanke (2009) notes that the increase in central bank balance sheets was initially driven by the need to accommodate the large scale of the credit easing programme, and as such the measures should be seen as complimentary rather than substitutes.

The CE programmes of central banks were primarily directed at minimising the disruption to the traditional transmission mechanisms of monetary policy, given the breakdown in the interbank market. Measures included an increase in the eligible counterparties for central bank liquidity programmes[1]; swap agreements with other central banks; an expansion in the types and maturity length of securities bought by central bank open market operations; and the direct purchase of both state agency and private securities (ECB, 2010; Lenza, Pill and Reichlin, 2010). In addition, governments implemented large fiscal stimulus programmes along with bank guarantees and

[1] The number of eligible counterparties for ECB refinancing operations was increased from 140 prior to the crisis to over 2,000 (ECB, 2010, p.66).

recapitalisations (Stolz and Wedow, 2010).

Effectiveness

Given the unprecedented and extreme nature of the financial crisis, any evaluation of the effectiveness of the measures taken contains a high degree of uncertainty. However, that hasn't prevented the emergence of a burgeoning literature seeking to provide some estimate of the effectiveness of public interventions. Chung et al. (2011) find that 'alternative monetary policy instruments, such as asset purchases, have been effective at mitigating the adverse macroeconomic effects of the ZLB'. They estimate that the Fed asset purchase programmes are 'roughly equivalent to a 300 basis point reduction in the short-term interest rate' but that despite this 'the ZLB has importantly constrained the ability of conventional monetary policy to limit the depth and duration of the current slump'. Del Negro et al. (2010) estimate that the Fed's unconventional policies prevented an output gap twice the size of that actually experienced by the US, and had a significant effect on interest rate spreads.

Research on the impact of measures in Europe has focused on the more clearly observable impacts of unconventional measures on the financial market and the transmission mechanism of monetary policy. Joyce et al. (2010, p.38) find that British 'guilt yields were about 100 basis points lower than they would otherwise have been as a result of QE'. A directly observable measure of the effectiveness of the European policy response is the reduction in the EURIBOR-EONIA spread from over 180 basis points following the collapse of Lehman Brothers to less than 5 after intervention. A more sophisticated analysis carried out by Lenza et al. (2010) finds that 'the effect of the compression of the spreads stemming from the implementation of the ECB's so-called 'enhanced credit support' has been sizeable on both loans and interest rates, very modest on broad money, and has acted on the real economy with a delay.' In addition, Lenza et al. (2010, pp.34-35) find that measures 'have played a quantitatively significant role in stabilising the financial sector and economy after the collapse of Lehman Bros., even if insufficient to avoid a significant fall in economic and financial activity'.

The lesson to be drawn from the Great Recession is that central bank intervention was successful to the extent it minimised the disruption of the transmission mechanism caused by the financial crisis. However, numerous studies using heuristics like the Taylor rule have shown that monetary policy remains considerably too tight. One such study by Williams (2009) finds that 'an additional 2 to 4 percentage points of rate cuts would help bring unemployment and inflation rates more quickly to longer run values, but the ZLB precludes these actions'. While extraordinary measures in extraordinary times may have preserved the transmission mechanism of the interest rate channel of monetary policy, this is of little consolation given the inability of central banks to lower interest rates to the degree needed. With the economic recovery remaining sluggish at best and unemployment stubbornly high, alternative policy solutions that overcome the ZLB surely demand to be considered.

Higher Inflation Targets in Normal Times

Blanchard, Dell'Arricia and Mauro (2010) point to the large fiscal deficits incurred, as national governments tried to compensate for the zero lower bound, and their consequences for sovereign creditworthiness. They suggest that a higher targeted level of inflation would give policy makers a greater policy buffer as nominal interest rates would be higher to start with. Williams (2009, p.26) argues that if 'the current global recession represents the death knell of the era of the Great Moderation and the equilibrium real interest rate remains low, then the ZLB may regularly interfere with the ability of central banks to achieve macroeconomic stabilisation goals.'

Blanchard, Dell'Arricia and Mauro (2010, p.11) argue that 'many of the distortions from inflation come from a tax system that is not inflation neutral, for example, from nominal tax brackets or from the deductibility of nominal interest payments' and that fixing such distortions could tip the balance in favour of a higher inflation target. Some critics of the proposal have countered that raising the target risks unanchoring inflation expectations and would result in credibility problems for monetary policy: why would markets believe the goalposts won't be moved again? Ultimately the appropriateness of a higher inflation target rests on whether the net costs associated with a higher inflation target outweigh the benefits from a great policy buffer. The costs of higher inflation, and uncertainty as to whether even the increased policy buffer would provide sufficient room, means other alternatives still should be considered.

Gessel's Solution/Carry Tax

Keynes (1936) and Irving Fisher (1933) revived the innovative monetary theories of German socialist economist Silvio Gesell. Gesell (1958) was motivated by a belief that the growth in real capital was held back by too high a rate of interest; a result of the low carry costs of money. For Gessel (1958, Part IV-1), only 'money that goes out of date like a newspaper, rots like potatoes, rusts like iron, evaporates like ether, is capable of standing the test as an instrument for the exchange of potatoes, newspapers, iron and ether.' His solution was to replace metallic and paper money with stamped money, whereby, to maintain the face value of a note, stamps purchased at a post office would have to be applied to the currency weekly. This implied a depreciation, or negative interest rate on currency of 5.2 percent a year.

Gesell's solution lay in obscurity until resurrected by Goodfriend (2000) and Buiter and Panigirtzoglou (2003). The latter authors demonstrate in a rigorous manner how the Gesellian solution allows central banks to overcome the lower bound, while the former provides a more accessible modern reinterpretation. Goodfriend (2000, p.1015) proposes a 'per period, per dollar carry tax on electronic bank reserves when the interbank interest rate is pressed to zero'. Competition between banks seeking to avoid the tax would have the effect of pushing the interbank rate, upon which most

bank contracts are based, below zero by a magnitude determined by the carry tax. The hoarding of currency would be overcome by the introduction of a fixed carry tax on currency and vault cash when the interbank rate hit zero. This carry tax on currency would be imposed through the modern equivalent of Gesell's stamped currency; a magnetic strip embedded in each bill which could record the date last withdrawn from the banking system and so the tax due.

Regardless of what method is used to impose a carry tax on currency, doing so would give central banks scope to reduce the policy rate significantly below zero, breaking the constraint of the ZLB. In the aftermath of the financial crisis, the Swedish Riksbank imposed a negative interest rate on reserves held at the Bank of minus 0.25 percent, a first for a modern central bank (Ward, 2009). Though not large enough to encourage the hoarding of currency, as the cost was absorbed by the banks, the move served to reduce what the Deputy governor of the Riksbank, Lars Svensson, described as 'zero interest rate mystique [that had] exaggerated the problems' associated with sub-zero rates (cited in Goodfriend, 2000, p.1015). Given the significant costs associated with the ZLB and the technological advances made in the past century, the prospect of Gesell's solution seeing implementation is perhaps not as far fetched as once thought. The inconvenience of checking the magnetic bills on currency to see whether they were valid would in and of itself increase the carry costs of currency, and discourage the use of physical currency. In this case, taking the bold step of abolishing physical currency, as discussed next, would a priori merit consideration.

Abolish Physical Currency

Buiter (2009) forcefully advocates the abolition of physical currency as a solution to the zero lower bound.[2] All money would then be held in accounts at banks subject to interest, positive or negative, determined by the carry tax on reserves. Not only does the technology exist for the abolition of paper and metal currency, but it is already installed in the majority of retail premises in the advanced world. Moving to a completely electronic system of payments would have additional benefits for taxpayers and governments, given that the main beneficiaries from the anonymity provided by currency are criminality, tax evaders and the black economy. The blight of so-called 'tiger robberies' and armed raids on cash-transit vans would also be solved by the move to an electronic system.

Those who worry about an overbearing state infringing individual rights by monitoring purchases of citizens can be placated by the provision of pre-paid cash cards, like those currently offered by the major credit card companies or the Transport for London 'Oyster card'. A limited amount could be transferred to such cards with a carry tax charged on the average balance, thus preserving an individuals ability to engage in legitimate anonymous activity. An additional cost to consider from the abolition of currency is that the central bank would lose revenues from seigniorage.

---

[2] Currency is used interchangeably with physical currency in what follows.

A less radical version of this solution could be introduced initially. The abolition of currency in denominations greater than €2 would significantly reduce the lower-bound on interest rates, given the large costs associated with hoarding significant quantities of currency in such small denominations without detection. This is similar to the Gesellian solution discussed above, acting as a permanent fixed carry tax on currency providing the central bank with the scope to reduce interest rates to significantly below zero.

Fiscal Policy

The aftermath of the 2008 financial crisis has seen a remarkable rise in public deficits as governments borrowed huge amounts to recapitalise the fragile banking system and provide a fiscal stimulus to recession hit economies. Despite the demise in popularity of Keynesian style fiscal policy over the latter part of the 20th century, the constraint of the ZLB convinced governments that such policies were necessary.

Woodford (2010b, p.40) shows that in Great Depression-like circumstances - that is when 'a disturbance to the financial sector results in insufficient aggregate demand even with the central bank's policy rate at the lower bound of zero, and when there is feared to be a substantial probability of the constraint continuing to bind for years to come, standard models of the kind widely used in analyses of monetary stabilisation policy imply that the government expenditure multiplier should be larger than one, and may be well above one'. Indeed, Kuttner and Posen's (2001, p.157) analysis of Japan's experience 'shows that fiscal policy works pretty much the way Keynes suggested it does: contractions are contractionary and expansions expansionary, and even wasteful public spending has a clear multiplier effect (although it is disadvantageous in other ways)'.

However, as the ongoing sovereign debt crisis is demonstrating, expansionary fiscal policy requires that a country has sufficient creditworthiness to borrow in often risk adverse markets. Political factors, such as ideologically motivated opposition to government intervention (e.g. Republican control of the Houses of Congress in the United States) may prevent an economically justified fiscal stimulus. In addition, the beneficial effects of fiscal stimulus are reduced the more open, the smaller and the more indebted a country is (Woodford, 2010b).

Commitment Devices

While promises to target a higher future inflation rate suffer from time-inconsistency problems (as discussed above), Bernanke and Reinhart (2004) argue that open market operations in longer term bonds can act as a commitment device. As a large participant in the Treasuries market across the maturity spectrum, the Treasury can push longer term yields on Treasury bonds down, signalling that it is committed to keeping interest rates low. Clouse et al. (2003) suggest that central banks could write options on future Treasury bills as a stronger way of committing themselves to

lower rates. Those sceptical of the credibility of such commitments could then insure themselves against a policy reversal at the cost of the central bank.

Another suggestion is that the primary target of monetary policy itself should be revisited. Woodford (2010a) argues that rather than concerning themselves with an inflation target, central banks should instead aim for a price level target. By targeting the price level, the central bank can more credibly commit to a higher future inflation rate in deflationary times, as the price level target implicitly promises to make up for any undershooting of the target.

Conclusion

This essay has explored the constraints on monetary policy imposed by the zero-lower-bound and what alternative policy solutions exist. It first discussed how the negligible carry costs of money imply a lower bound close to zero for nominal interest rates, given lenders will prefer to hold money at sub-zero rates. Time-inconsistency hampers the credibility of central bank commitments to target higher levels of inflation. The combined result of this is that central banks cannot stabilise economic fluctuations by reducing real interest rates. The essay then illustrated the consequences of the binding zero-lower-bound in the context of the ongoing Great Recession. Policy alternatives for overcoming the zero-lower-bound were then considered. These included: the abolition of physical currency, Gesellian stamped currency and the use of fiscal policy.

What this essay has demonstrated is that while the zero-lower-bound poses a serious challenge for monetary policy, alternative solutions do exist. Much like John Law's advocacy of paper currency was decried as lunacy, so too will suggestions that physical currency be done away with entirely, or replaced by the script stamp of Gesell. However, the economic and human costs, in terms of lost output and increased unemployment, of central banks' inability to symmetrically respond to severe economic shocks surely warrants that radical alternatives be thoroughly explored.

References

Bernanke, B., 2009. 'The Crisis and the Policy Response'. In: *Stamp Lecture*. London School of Economics, 13 January 2009. [Online]Available at: http://www.federalreserve.gov/newsevents/speech/bernanke20090113a.htm

Bernanke, B. Reinhart, V., 2004. 'Conducting Monetary Policy at Very Low Short-Term Interest Rates'. *American Economic Review,* 94(2), pp.85–90.

Bernanke, B., Reinhary, V. and Sack, B., 2004. 'Monetary Policy Alternatives at the Zero Bound: An Empirical Assessment'. *Brookings Papers on Economic Activity* , 2004 (2), pp.1–100.

Blanchard, O., Dell'Ariccia, G. Mauro, P., 2010. 'Rethinking Macroeconomic Policy'. International Monetary Fund, Washington, DC. *IMF Staff Position Note SPN/10/03*.

Buiter, W. H., 2009. 'Negative Interest Rates: When Are They Coming to a Central Bank Near You?'. *Financial Times* 7 March. [Online] Available at: http://blogs.ft.com/maverecon/2009/05/negative-interest-rates-when-are-they-coming-to-a-central-bank-near-you/

Buiter, W. H. and Panigirtzoglou, N., 2003. 'Overcoming the Zero Bound on Nominal Interest Rates with negative interest on currency: Gesell's solution'. *The Economic Journal,* 113(490), pp. 723–746.

Chung, H., Laforte, J. P., Reifscneider, D. and Williams, J. C., 2011. 'Have We Underestimated the Likelihood and Severity of Zero Lower Bound Events?'. Federal Reserve Bank of San Francisco *Working Paper Series, 01, pp.1–58*.

Clouse, J., Henderson, D., Orphanides, A., Small, D. H. and Tinsley, P., 2003. 'Monetary Policy When the Nominal Short-Term Interest Rate is Zero'. *Topics in Macroeconomics,* 3(1), Article 12.

Del Negro, M., Eggertsson, G., Ferrero, A. and Nobuhiro, K., 2010. 'The Great Escape? A Quantitative Evaluation of the Fed's Non-Standard Policies'. Federal Reserve Bank of New York*, Working Paper.*

ECB, 2010. 'The ECB's Response to the Financial Crisis'. *ECB Monthly Bulletin,* October, pp. 59–75. [Online]Available at: http://www.ecb.int/pub/pdf/mobu/mb201010en.pdf

Fisher, I., 1933. *Stamp Scrip.* New York: Adelphi Company.

Gesell, S., 1958. *The Natural Economic Order*. London: Peter Owen Ltd.

Goodfriend, M., 2000. 'Overcoming the Zero Bound on Interest Rate Policy'. *Journal of Money, Credit and Banking*, 32(4), pp.1007–1035.

Goodfriend, M., 2007. 'How the World Achieved Consensus on Monetary Policy'. *Journal of Economic Perspectives*, 21(4), 47–68.

Joyce, M., Lasaosa, A., Stevens, I. and Tong, M., 2010. 'The Financial Market Impact of Quantative Easing'. Bank of England, London. *Working Paper no.393*, pp.1–44.

Keynes, J. M., 1936. *General Theory of Employment, Interest and Money*. London, Macmillan Cambridge University Press.

Krugman, P., 1998. 'It's Baaack: Japan's Slump and the Return of the Liquidity Trap'. *Brookings Papers on Economic Activity,* 2, pp.137–205.

Kuttner, K. N. and Posen, A. S., 2001. 'The Great Recession: Lessons for Macroeconomic Policy from Japan'. *Brookings Papers on Economic Activity,* 2, pp.93–198.

Kydland, F. E. and Prescott, E. C., 1977. 'Rules Rather Than Discretion: The Inconsistency of Optimal Plans'. *Journal of Political Economy,* 85(3), pp.473–491.

Lenza, M., Pill, H. and Reichlin, L., 2010. 'Monetary Policy in Exceptional Times'. ECB, *Working Paper Series no. 1253*, pp.1–40.

Mankiw, G., 2009. 'Reloading the Weapons of Monetary Policy'. *Greg Mankiw's Blog,* [Online] Available at: http://gregmankiw.blogspot.com/2009/03/reloading-weapons-of-monetary-policy.html

McCallum, B. T., 2000. 'Theoretical Analysis Regarding a Zero Lower Bound on Nominal Interest Rates', *Journal of Money, Credit and Banking,* 32(4), pp. 870–904.

Stolz, S. M. and Wedow, M., 2010. 'Extraordinary Measures in Extraordinary Times'. *ECB Occasional Paper Series,* 117, pp.1–70.

Svensson, L. E. O., 2001. 'The Zero Bound in an Open Economy: A Foolproof Way of Escaping a Liquidity Trap'. *Monetary and Economic Studies,* February. pp.1–46.

Svensson, L. E. O., 2003. 'Escaping From a Liquidity Trap and Deflation: The Foolproof Way and Others'. *Journal of Economic Perspectives,* 17(4), pp.145–166.

Ward, A., 2009. 'Bankers Watch as Sweden Goes Negative'. *Financial Times* 27 August. [Online] Available at: http://www.ft.com/cms/s/0/5d3f0692-9334-11de-b14600144feabdc0.html#axzz1FZDG1nMN

Williams, J. C., 2009. 'Heeding Daedalus: Optimal Inflation and the Zero Lower Bound'. *Brookings Papers on Economic Activity*, September, pp.1–37.

Woodford, M., 2010a. ‘Bernanke Needs Inflation for QE2 to Set Sail’. *Financial Times* 11 October. [Online] Available at: http://www.ft.com/cms/s/0/4d54e574-d57a-11df-8e86-00144feabdc0.html

Woodford, M., 2010b. ‘Simple Analytics of the Government Expenditure Multiplier’. National Bureau of Economic Research, Cambridge, MA. *Working Paper no. 15714*. [Online] Available at: http://ideas.repec.org/p/nbr/nberwo/15714.html

Yates, T., 2004. ‘Monetary Policy and the Zero Bound to Interest Rates: a Review’. *Journal of Economic Surveys,* 18(3), pp.427–481.

REGULATORY CAPTURE IN THE CONTEXT OF THE IRISH REGULATORY FRAMEWORK

Niamh Callaghan

Senior Sophister

Niamh Callaghan looks at the overall regulatory framework of the transport sector in Ireland, and reveals how prone it is to regulatory 'capture' by interest groups who often compete with each other to do so. She also demonstrates the benefits that have accrued from deregulation, for example in the Irish taxi industry, despite the resistance to embrace competition on the part of Irish regulatory authorities.

Introduction

Economic regulations intervene directly in market decisions such as pricing, competition, market entry and exit. Reform aims to increase economic efficiency by reducing barriers to competition and innovation, often through deregulation and the use of efficiency-promoting regulation (OECD, 2011). The quality of regulatory decisions impacts upon Ireland's competitiveness, which is so crucial to our economic survival at present and the welfare of Irish consumers (Gorecki, 2010). The transport sector has a particularly important role in our economy and has traditionally been highly regulated. Thus an investigation of the way such regulatory actions are implemented is a necessary task.

Regulatory capture is a real threat due to the number of different parties in this industry with competing motivations. Regulatory capture is defined as a situation in which one operator (or group of operators) in the market uses its influence or resources to extract a regulatory decision, or lack of decision, for their own benefit rather than the benefit of society as a whole (Department of the Taoiseach, 2004). In this essay, an outline of the theories of regulatory capture will be provided to create a platform on which to discuss the functioning of the Irish regulatory framework. Using examples such as CIE, the existence of this problem in Ireland as well as some success stories will be highlighted.

There are many arguments for regulation, both economic and social. The main economic reasons for regulation are to guarantee the provision of a service, to act as a counterweight to market power, to tackle externalities and to correct information asymmetries (Turnbull, 1999). Sometimes these arguments can be conflicting, reflecting the ongoing trade-off in economics between efficiency and equity. Regulation in the transport sector has traditionally been justified by the belief that in its absence the market would under invest in transport services, particularly in peripheral areas. The evidence does not support this argument as the bulk of investment in transport is undertaken by the private sector (Barrett, 1990).

Capture Theories

The 'capture' or 'interest group' theory concentrates on the role interest groups play in the formation of public policy. Information asymmetries must exist between businesses and the state to create an incentive to influence the behaviour of public institutions (e.g. regulators). Interest groups have a stake in the behaviour of state agencies which regulate them therefore independent oversight is crucial to responding to potential collusion between the two parties (Harrington, Vernon and Viscusi, 2005).

Stigler (1971) posits that conflicts of interest occur in all aspects of life but the fight is fairer when played out in the market place as opposed to the political arena. Market decisions are instantaneous and carried out by all participants as demand reacts to supply. In politics, decisions are made infrequently and by elected representatives to reflect the global good but in reality this system is open to manipulation. This analysis goes to the very core of the decision to place regulation in the hands of a few rather than leaving the decision with each individual in the marketplace.

Pioneering work by Stigler (1971) and later expanded by Peltzman (1976) helped to develop a theory to explain the occurrence of regulatory capture. The three main points outlined by the two authors were that: regulatory legislation is first and foremost a redistribution of wealth; the behaviour of legislators is motivated by their desire to remain in office implying that legislation is designed to maximise political support; and finally, interest groups compete by offering this political support[1]. The conclusions drawn are that better organised and funded interest groups will exert the most influence on regulatory bodies. Furthermore, they will benefit more from the gains from favourable legislation and will thus be willing to invest more resources into making capture possible. In particular, regulation will benefit small interest groups with strongly held preferences to the detriment of large interest groups with weaker preferences. This insightful argument put forward by Peltzman (1976) reinforces the empirical evidence that regulation tends to be pro-producer. Firms, due to their smaller size and the importance of profits,

[1] The political element referred to above does not merely include support from political parties or institutions but perhaps funding and support from private sectors.

will be more likely to 'capture' the regulator rather than the very large group of consumers with a much smaller interest in the regulation of one particular product or service which they consume[2].

Another model, proposed by Becker (1983), focuses on the competition between interest groups to become the 'capturer'. This model dismisses the role of the regulator and assumes that regulation is used to increase the welfare of the most influential interest group. The political equilibrium resulting from this model is Pareto inefficient[3] and as with all cases of capture, results in a large waste of resources.

Another reason a regulatory agency may be guilty of capture is due to the effect public and private interests have on the survival of regulatory officials and the agency's budget. A regulatory agency can have an interest in the success of the industry they are regulating and may thus be tempted to act more in the interests of the industry rather than the public[4]. A further issue is the funding of regulatory agencies. Agencies are often funded by a levy on the regulated firms which may be passed onto the consumer (Gorecki, 2010).

Costs of Inefficient Regulation

The stakes involved in regulation are very high. There is a trade-off between having regulation, which can lead to gross inefficiencies if not correctly implemented, or no regulation which can also lead to inefficiencies and a waste of resources in the form of monopoly profits and a deadweight loss for society. For instance, Winston (1998) estimated that the gains to passengers and companies from deregulation in the transportation field are $50 billion per year. This was a sizeable cost levied on consumers prior to deregulation.

The late Alfred Kahn (1971) notes that regulation contains no built-in mechanism to ensure efficiency, despite its purpose as a device to improve market conditions and remove inefficiencies associated with the conduct of monopolies[5]. Regulation restrains companies from fully exploiting their potential monopoly power. Thus, their incentive to operate efficiently is dramatically reduced as the supernormal profits which previously motivated them have been eliminated. Regulation can cause the alteration of the behaviour of the regulated firms in order to lessen the impact of the regulation on their profits incurring a further waste of resources.

---

[2] An example of this is illustrated later in the essay.
[3] Both groups could invest fewer resources and still achieve the same level of relative influence but, similarly to Nash equilibrium in the prisoners' dilemma game, no party will wish to deviate from this outcome even though it is inefficient.
[4] Such an interest may include a requirement in its brief that it should promote the industry, e.g. this occurred in the Irish financial industry much to the detriment of the banking system and the economy of this country.
[5] A higher price is charged to consumers and a quantity below equilibrium is supplied.

The costs of regulation are numerous and may not appear obvious at first glance. Regulation imposes a burden on businesses[6], directly and indirectly. Often firms are charged a levy to fund the activities of the regulator[7]. Other direct costs include compliance costs, encompassing employee training and efforts to understand and implement the rules, claims on management attention, discouraged investment and lost producer and consumer surplus from reduced output (Harrington, 2006). The initial costs of compliance are by no means the end of the story; there are many indirect effects on innovation and productivity which can be very damaging to the effected industry and the economy as a whole, as consumers are denied new and improved products or levies imposed are passed onto them in the form of increased prices (Hopkins, 1995). Furthermore, the costs of rent seeking[8], including the unproductive activities undertaken by firms and individuals when trying to influence regulatory decisions, should also be considered. Thus the costs associated with regulation are much larger than first thought. Posner's (1975) paper shows an economic model of the social costs of monopoly and he concludes that the costs are most likely higher in the regulated, as opposed to the unregulated, sector of the economy.

Regulatory Capture in the Irish Transport Sector

The Organisation for Economic Co-operation and Development (OECD) published a report in 2001 detailing the history of regulation in Ireland and the reforms which should be implemented in the future (OECD, 2001). The report highlights the existence of regulatory capture in Ireland and its effect on impeding economic progress. It points to a lack of clear priorities or strong incentives to transform deep-rooted regulatory traditions and practices as the reasons for lack of reform. The report acknowledges that significant progress has been made in some areas but further action must be taken to prevent regulatory capture, including reducing the influence of informal processes that provide opportunities for interests to monopolise information and thus favour 'insiders' (OECD, 2001).

Improving transparency is proving especially difficult, due to rent-seeking attitudes and the close relationships between elected representatives and producer interests, which generally act against free market principles and the interests of consumers. Consultation processes on reforming local transportation systems in particular have garnered much criticism for being greatly influenced by interest groups.

---

[6] This burden has been the subject of much debate in the US following President Obama's introduction of many new rules and regulations. Some have argued that the increase in red tape has caused the President to be labeled as 'anti-business' by some agents in the economy (The Economist, 2011).

[7] The National Transport Authority is funded by the state which also provides the subsidies to CIE.

[8] The opportunity to capture monopoly rents provides firms with an incentive to use scarce resources to secure the right to become a monopolist. Such activity is referred to as rent-seeking. Rent-seeking is normally associated with expenditures designed to persuade governments to impose regulations which create monopolies (OECD, 2011).

The Commission for Aviation Regulation (CAR)[9] has been successfully captured, as a result of Ministerial Intervention, by the Dublin Airport Authority (DAA), whom it was meant to regulate. CAR imposed price caps which were ignored by the DAA, resulting in the earning of supernormal profits (Commission for Aviation Regulation, 2008). The breach of the price caps incurred no rebuke for the DAA, thus allowing the behaviour to continue and other parties, namely customers, to suffer.

The government considered allowing private contractors to build and manage an independent Terminal 2 at Dublin airport, a premise strongly supported by Ryanair. However, this move was resisted by the DAA due to their desire to retain their monopoly position at the airport. The government then decided that the DAA would build the terminal but the contract for its operation would be put out to tender. There was yet another u-turn in March 2010 when the DAA were announced as the operators amid claims that other applicants did not meet the requirements for operating the terminal (Ryanair, 2005). These events clearly demonstrate the power wielded by the DAA over the government. The DAA now operates both Dublin airport terminals eliminating the possibility of competition between the two and potentially keeping charges artificially high.

Alternatives to Regulation

There are many economic benefits associated with the private ownership of real assets. With privatisation, a market of corporate control is created. The threat of bankruptcy, which is absent from nationalised firms, is a strong motivating factor for ensuring efficiency. The market, through Adam Smith's (1776) famous 'invisible hand', will evaluate the performance of management. In contrast, ministers and government departments cannot objectively assess the efficiency of a nationalised firm under the bind of regulatory capture. The remedy of market failures can rarely be undertaken without the creation of gross inefficiencies (Barrett, 1990).

Furthermore, not all markets with a dominant firm may be inefficient and require regulation. The theory of contestable markets (Baumol, 1982) states that the threat of potential entrants can be sufficient to ensure a firm behaves in a competitive manner[10]. The extent of freedom of entry should be considered in this assessment. Also, when competition within a market is not possible there may be scope to instigate competition for the right to operate in the market (Demsetz, 1973). This method may be useful in the case of natural monopolies, where due to economies of scale, it is irrational to have more than one firm. Competitive tendering is used successfully in the air traffic control industry to reduce costs (Barrett, 1990).

[9] CAR now falls under the remit of the National Transport Authority (NTA)
[10] The firm does not earn economic profit in the long run and that the long run price is equal to marginal cost.

It is very rare that a nationalised firm would pay dividends to the Exchequer indicating that inefficiencies prevail, thus providing evidence of the spiralling costs associated with state-owned companies. For example, CIE made a loss of €77.7 million euro in 2009 (CIE, 2010).

Success of Deregulation

Airline deregulation occurred on all Ireland and UK routes in 1986. It resulted in the reduction of fares by up to 54% on the Dublin to London route and an almost doubling of capacity. Furthermore, the number of tourists visiting Ireland increased three-fold from 2 million to 6 million (Barrett, 2003b). This is an example of a very successful liberalisation of the market resulting in gains for the industry in terms of increased efficiency, the consumer in the form of lower prices and the economy as a whole.

Deregulation in the Irish taxi industry occurred in 2000 despite constant pressure from incumbent taxi license holders to maintain the old regime. In the period 1978-1991, the number of taxis in Dublin remained relatively constant at approximately 1,800 (Barrett, 2003a). Considering the large economic growth in this period, these figures are startling and tell the story of a stagnant industry devoid of reform. The incumbents managed to successfully delay deregulation and to frustrate the Irish consumers for over thirty years due to their claims that safety levels would drop, their incomes would be depressed and a devaluation of licenses would occur. These arguments eventually proved unfounded and the results of deregulation have been overwhelmingly positive for consumers and new market entrants. The number of taxis soared following the judgment in 2000 proving that demand was not being met. However, incumbents remain unhappy with the decision and have once again attempted to capture the regulatory authorities by staging a series of strikes, poster campaigns and lobbying trade unions. As a result of these actions a Taxi Hardship Panel was set up to analyse the effect of deregulation on the livelihood of those taxi drivers whose licenses considerably diminished in value following deregulation (Barrett, 2003a). It is open to debate whether the formation of this panel constitutes regulatory capture but it appears that the price for a license was artificially high before deregulation and the current price reflects the equilibrium market price formed by the interaction of supply and demand. Barrett (2003a) considers this change in value analogous to an investment which is open to market fluctuations. It must be emphasised that some of the advantages of deregulation will accrue in the long-term and thus the success (or failure in some groups' eyes) should be analysed in a long-run context (Winston, 1998).

Regulatory Reform in Ireland

Reform of regulation in the transport sector has to be addressed immediately due to its extreme importance. Due to our location on the periphery of Europe, transport is vital to our export dependent economy. Low transport costs are vital to improving our competitiveness. The OECD (2001)

recommends the removal of licensing restraints throughout the economy, particularly in the bus industry and at ports. Further potential changes include the increased representation of consumer as opposed to producer interests during policy discussions.

A Government Statement on Economic Regulation (Department of the Taoiseach, 2009) announced the creation of a single transport regulator, the National Transport Authority. It encompasses the Commission for Taxi Regulation, the Commission for Aviation Regulation and the Irish Aviation Authority. This statement also outlines increased powers for ministerial oversight in respect of regulatory agency. Although an increased surveillance is on the surface a positive action, ministers are more prone to being 'captured'[11]. Furthermore, the statement acknowledges the positive impact deregulation has had on the transport sector. Despite this admission, the government has been unwilling to commit to deregulation of the bus industry. The framework proposes the introduction of a franchising system with both public and private contributions. However, licenses will only be granted to private operators for routes which the state body, Bus Átha Cliath, does not already operate on.

The regulatory decision-making of a minister is carried out with political considerations in mind. Many politically motivated decisions are short sighted and often not the best solution in the long term due to electoral pressures. In the case of the regulation of state owned firms[12], another consideration is the fact that the entity is owned by the state and thus the minister has an interest in the maximisation of its value, resulting in the potential dampening of competitors. Furthermore, government departments may not have adequate expertise to best evaluate regulatory decisions, leaving room for regulatory capture to occur (Gorecki, 2010).

Conclusion

Airline deregulation was considered ground breaking in the 1980s and taxi deregulation was considered a landmark for 'Irish law dealing with property rights and market access' (Barrett, 2003b, p.40). However, many problems still exist within the Irish regulatory framework. Despite the obvious gains from deregulation, as outlined above, there remains a reluctance to embrace open competition. In many cases, this unwillingness can be ascribed to regulatory capture.

References

[11] This is due to the nature of political decision-making as outlined above in Stigler's theory.
[12] CIE being one of the largest semi-state bodies in Ireland

Barrett, S., 1990. 'Transport Deregulation and Privatisation'. In: F. Convery and M. McDowell, eds. Privatisation. Dublin: Gill and Macmillan.

Barrett, S., 2003a. 'Regulatory Capture, Property Rights and Taxi Deregulation: A Case Study'. Economic Affairs, 23(9).

Barrett, S., 2003b. 'Privatisation in Ireland'. CESifo Conference Centre, Munich. Privatisation Experiences in the EU, [Online] 1-2 November. Available at: http://www.cesifo-group.de/portal/page/portal/ifoContent/N/neucesifo/CONFERENCES/SC_CONF_1999-2006/PRI03/PAPERS/PRI03-BARRETT.PDF [Accessed: February 02, 2011]

Barrett, S., 2004. 'Bus Competition in Ireland – The Case for Market Forces'. Economic and Social Research Institute*, Quarterly Economic Commentary: Special Articles*, 3, pp.1-20.

Baumol, W., 1982. 'Contestable Markets: An Uprising in the Theory of Industry Structure'. *American Economic Review*, 72(1), pp.1-15.

Becker, G., 1983. 'A Theory of Competition among Pressure Groups for Political Influence'. *The Quarterly Journal of Economics*, 98, pp.371-400.

CIE, 2010. 'Summary of Performance'. *CIE*, [Online]. Available at: http://www.cie.ie/about_us/summary_of_performance.asp

Commission for Aviation Regulation, 2008. 'Dublin Airport Authority plc Extract from Regulated Entity Accounts. Year Ended 31 December 2007'. Dublin: Commission for Aviation Regulation.

Dementz, H., 1973. 'Industry Structure, Market Rivalry, and Public Policy'. *Journal of Law & Economics*, 16(1), pp.1-9.

Department of the Taoiseach, 2004. 'Regulating Better'. Department of the Taoiseach, [Online] January. Available at: http://www.betterregulation.ie/eng/Government_White_Paper_'Regulating_Better'/ [Accessed: January 8, 2011]

Department of the Taoiseach, 2009. 'Government Statement on Economic Regulation'. Department of the Taoiseach, [Online] October. Available at: http://www.betterregulation.ie/eng/Government_ Statement_on_Economic_Regulation.pdf [Accessed: January 15, 2011]

Gorecki, P,. 2010. 'Regulation in Ireland'. In: The Irish Economic Association, *Economic Policy Conference 2010*. Kenmare.

Harrington, W., 2006. 'Grading Estimates of the Benefits and Costs of Federal Regulation. Resources for the Future Discussion Paper'. In: RFF, *Discussion Paper no.06-39*.

Hopkins, T. D., 1995. 'Profiles of Regulatory Cost'. Rochester Institute of Technology, *Report to the U.S. Small Business Administration*. [Online] November. Available at: http://archive.sba.gov/advo/ research/rs1995hoptot.pdf [Accessed: January 15, 2011]

Kahn, A., 1971. *The Economics of Regulation: Principles and Institutions*. Cambridge, MA: MIT Press.

Massey, P., 2007. 'DELAYED INDEFINITELY: Regulatory Reform of the Irish Bus Industry'. *Quarterly Economic Commentary*, 1, pp.38-61.

Massey, P., 2002. 'Competition and Deregulation: A Review of Irish Experience'. *The Antitrust Bulletin*, Winter.

OECD, 2001. 'Regulatory Reform in Ireland'. *OECD*, [Online]. Available at: http://www.taoiseach.gov.ie/eng/Publications/Publications_Archive/Publications_2006/Publications_for_ 2001/OECD_Report_on_Regulatory_Reform_in_Ireland_-_Executive_Summary.html [Accessed: December 14, 2010]

OECD, 2011. 'Glossary of Statistical Terms'. *OECD*, [Online]. Available at: http://stats.oecd.org/glossary/detail.asp?ID=4639 [Accessed: January 31, 2011]

Peltzman, S., 1976.'Toward a More General Theory of Regulation'. *Journal of Law & Economics*, 19(2), pp.211-40.

Posner, R., 1975. 'The Social Costs of Monopoly and Regulation'. *The Journal of Political Economy*, 83(4), pp.807-828.

Ryanair, 2005. 'Minister Cullen Breaks His Own Terminal 2 Promises and Safeguards'. *Ryanair,* [Online]. Available at: http://www.ryanair.com/en/news/gen-en-300905 [Accessed: January 31, 2011]

Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. London: W. Strahan and T. Cadell.

Stigler, G., 1971. 'The Theory of Economic Regulation'. *Bell Journal of Economics*, 2(1), pp.3-21.

The Economist, 2011. 'Red Tape Rising'. *The Economist,* [Online] 20 January. Available at: http://www.economist.com/node/17961890

Turnbull, P., 1999. 'Regulation, Deregulation or Re-regulation of Transport?. Symposium on the Social and Labour Consequences of Technological Developments'. *Deregulation and Privatization of Transport Discussion Paper No. 4*.

Viscusi, W., Harrington, J. and Vernon, J., 2005. *Economics of Regulation and Antitrust*, 4th ed. Cambridge, MA: MIT Press.

IRISH BUS COMPETITION AND THE SWORDS EXPRESS CASE

EOIN MARSH

Senior Sophister

Eoin Marsh explores the issue of Irish bus competition, a sector which is heavily regulated and dominated by state-funded players. He demonstrates the inefficiencies of Dublin Bus and how state authorities have prevented competition from entering the market, making a convincing argument for a radical overhaul of the sector.

Introduction

The issue of bus competition is one which has attracted growing attention in Ireland in recent years. The dominant state-funded player has had difficulties in providing cost-effective services while some private sector operators have been able to run services competitively. This paper will examine the role of the Department of Transport in favouring state-sponsored provision of bus services and the rationale presented in support of this. The competition on existing routes in Ireland, particularly on intercity services, will be scrutinised. Furthermore, the degree of bus competition in Dublin will be compared and contrasted with the scenarios in other cities such as London, Denver and San Diego where liberalisation has been introduced. The issue of regulatory capture will also be examined.

In addition, the recent case of Swords Express will be discussed and the extent to which state authorities have prevented competition from taking place on the market will be explored. This essay will provide an analysis of costs, with respect to both the public and the private supplier, and will also highlight the trends regarding subsidies within CIÉ.

The Context in Ireland

Strict regulation of the bus sector has prevailed in Ireland for almost eighty years; since 1932 Córas Iompair Éireann (CIÉ), which comprises of Bus Éireann and Bus Átha Cliath (Dublin Bus), has effectively run a monopoly. In contrast to transport sectors in which deregulation has been successfully introduced, such as aviation and taxis, the bus sector still experiences many obstacles that prevent competition and contestability. It is argued that the regulatory capture of the Department of

Transport by CIÉ is emphasised by the fact they are the sole recipient of public transport subsidies (Barrett, 2004).

The rationale for regulating the bus sector and preventing extensive competition from occurring is presented in a report which was undertaken by Steer Davies Gleave (SDG) in 2002. Though their paper does outline some potential advantages of bus competition, such as lower fares overall and increased frequency, these are outweighed by the problems that competition would bring. The report makes a strong argument against competition insofar as service levels may be reduced from intermediate towns and remote areas which have low demand and are loss-making. In addition, the point is made that fares might increase on routes where there is little or no competition, especially in the case of tendered routes (Steer Davies Gleave, 2002).

The report can be questioned in some respects however; it finds that the integration of services might be affected by new competitors. The market success of low cost point-to-point routes somewhat undermines the argument that the integration of routes should be prioritised. Also, requiring competing transport operators to integrate services runs the risk of encouraging collusion. The SDG paper also makes the case that, with the onset of competition, bus standards would become variable and that competitors would focus on high demand routes. These findings go against the experience of airline deregulation; standards such as punctuality, ease of booking and frequency have been improved with deregulation and services have extended beyond hub routes to more provincial destinations (Barrett, 2004). Overall, an examination of the SDG report indicates that there are at least as many reasons to favour competition as there are to deter it.

Evidence shows that bus competition in Ireland has been quite successful. The deregulation of certain intercity routes in recent years has exemplified this:

> 'Competition on Dublin-Belfast increased frequency from 7 to 39 services per day in each direction and reduced the fare from €19 to €12. Competition between Dublin and Cork reduced the fare from €20.50 to €7 and frequency increased from 6 to 14' (Barrett, 2008, p.595).

Similar findings are in evidence on other contested routes such as those between Dublin and both Galway and Waterford. Frequency of service per day is considerably higher and fare per mile much lower on these routes than on comparable uncontested routes (Barrett, 2004). In addition, data also suggests that despite restrictions on the ability of the private sector to operate services which compete with state-sponsored companies, this sector has outperformed each of the CIÉ operators. In 2003 it was found that the private sector bus fleet has passenger receipts which were 77% more than those of Dublin Bus, 41% more than Bus Éireann and 124% more than the state railways (Barrett, 2008).

For private sector operators in Dublin, contestability has been difficult to bring about. The notable exception in recent years has been Aircoach. Established in 1999, this company has expanded to operate five routes linking Dublin Airport with the south part of the city and runs a twenty-four hour service (Aircoach, 2011). The company must compete with Dublin Bus, however, which receives state subsidies. Aircoach has also had its licences restricted. Barriers to entry such as these hinder the ability of potential entrants to penetrate the market in Dublin.

In contrast to the case in Dublin, cities in a number of countries have introduced bus liberalisation. In the US for example, cities such as Denver and San Diego have introduced competitive tendering on urban bus services. The dominance of ideology in motivating policy-making in the US, in this instance with the Republican Party espousing privatisation, can be problematic and can bring about the shortcomings of market forces (as mentioned above), such as uncontested fares being increased and low demand routes being neglected. The benefits of liberalisation can be identified to a considerable extent however; for example, in Denver, competition has meant that the overall bus service has expanded and managerial and general in-house efficiency has been greatly improved. Similarly, San Diego also initiated tendering for routes; overall costs have been greatly reduced and the level of bus services increased (Richmond, 2000).

Additionally, when one examines the situation in the UK, it is evident that bus competition has existed in London since the mid-1980s. Bus services which were previously operated by London Transport were put out to competitive tender, as this body was broken up into a number of companies that were subsequently privatised. Though the formation of a large bureaucracy (Transport for London) and a recent rise in government subsidies can be seen as deficiencies in this model, important improvements in the bus system are in evidence; passenger numbers in London have grown since liberalisation. In addition, bus operating costs have fallen and bus services have increased (Compecon, 2003).

Notwithstanding the experience of cities abroad, the case for avoiding contestability in the bus sector seems to be prevailing to an increasing extent in Dublin. The establishment of the National Transport Authority (NTA) in December 2009 has allowed for the continuance of central service provision by the CIÉ operator Dublin Bus. Operating subsidies and capital grants for Dublin Bus have increased further as the company's exclusive right to provide its services is enshrined in the *Dublin Transport Authority Act 2008* (DTAA, 2008), which sets out the provisions for the NTA's functioning (NTA, 2011). This legislation does require that all new licences for state-sponsored routes are put out to tender, which means private operators will for the first time be able to apply to win the right for these services. This provision has yet to be put into practice however as demonstrated when Dublin Bus was recently unchallenged in changing its No. 63 service to operate a new route between Dún Laoghaire and Kilternan.

Lessons from Swords Express

Evidence presented from both the domestic market and from overseas emphasizes that efficiencies attained by private sector bus operators can often be significant and generally eclipse the deficiencies of market forces. This paper will now examine further the potential for lower cost operations, as well as the regulatory capture of policy makers and regulators with reference to one such private bus company, namely Swords Express. This company is an example of a competitive bus operator which has been affected by obstacles to contestability in the Irish bus sector. Since the company first applied for a licence to operate in April 2005, it has had to initiate court proceedings in order to facilitate its expansion, through the attainment of a further licence, and also to prevent unfair competition from being allowed in the market, in the form of illegitimate state subsidies being granted to its competitor, Dublin Bus.

Swords Express was awarded a licence in October 2007 and commenced running services a short time later. At the start of 2008, the company made an application to the Department of Transport for a second passenger licence, encouraged by its success to date. Subsequently, Dublin Bus made contact with the then Minister for Transport, Noel Dempsey, urging a re-routing of their No. 41X service, which served Swords, to a course via the Dublin Port Tunnel which mirrored that of the new operator. Under Irish Law, the *Transport Act 1958* (TA, 1958) requires that ministerial consent is obtained when a state-sponsored service provider wishes to initiate or alter an existing passenger road service which competes with another licensed service. Effectively, the granting of state subsidies on a route which has competition is not permitted because of its anti-competitiveness. However, the Department of Transport by-passed the Minister, having found that no competition existed on the route in question, despite the fact the two companies operated an almost identical route. As a consequence the 41X received subsidies, thereby allowing Dublin Bus to charge a fare of €2.30 on the route, whereas Swords Express, in receipt of no subsidy, was forced to charge customers between €3 and €4 (The Courts Service of Ireland, 2011).

As a result, judicial review proceedings were initiated later in 2008 with Swords Express contesting that competition did exist on the route which it operated and that Ministerial consent would be required. Consultants were engaged to adjudicate on the matter and it was found that competition was indeed existent and that the Department had erred. The Minister for Transport soon confirmed the continuance of the 41X but the service would no longer be allocated state funding. In the meantime, the Minister had yet to decide on the second licensing application of Swords Express, despite this having being agreed during judicial proceedings (*Irish Times*, 2010).

However, the victory for Swords Express in face of the anti-competitive behaviour of the State authorities proved to be short-lived. In 2009, another consultants' report was carried out on behalf of the Department of Transport. The findings indicated that the removal of one bus stop from the 41X's current route would eliminate any competitive advantage it had. Dublin Bus and the Department agreed that the revised route involved no competition and altered its service accordingly. Consequently, Swords Express began further court proceedings against the Minister for Transport in

December 2009. The company questioned this blur between competitive advantage and competition itself. McMahon J. found in their favour, stating that competitive advantage was not of statutory concern and that the Minister's decision to permit the altered route was made in error. It was also declared that the applicant's submission for a second licence had been unlawfully delayed, having been put on hold for almost thirty months. The extent to which a state-sponsored company was favoured illegitimately over a competitive private operator which was obstructed is evidenced by Mr. Justice McMahon's criticism of the approach of the Minister and that of his Department. The judge reprimanded them for their inconsistency in considering applications and for ignoring their own guidelines (Courts Service, 2011).

Data suggests that bus operators in the private sector can provide services more cost-efficiently than companies which are publically funded. A pre-budget submission to Minister for Finance Brian Lenihan in October 2010 by Antoin O'Lachtnain, Director of Swords Express, underlines the scope for greater cost effectiveness in the bus sector by allowing greater service provision in the private domain. The paper presents figures showing that the cost per bus per year of Dublin Bus is €316,000, an amount over twice that of Swords Express whose cost is €153,000. In addition, Dublin Bus has a total of 936 buses in its fleet compared with just 6 in use by Swords Express. Therefore, the costs of the publically subsidised company are quite striking given the potential for economies of scale when running such a large operation. It also points out that Swords Express is operating a more expensive and higher quality fleet of buses which would work against the company in an analysis of costs (O'Lachtnain, 2010).

O'Lachtnain's pre-budget submission paper also indicates CIÉ's growing level of inefficiency with regard to allocation of capital. It stated that, despite overall passenger numbers falling by roughly twenty million in the period between 2002 and 2009, CIÉ has increased its capital employed by over two-fold, with figures adjusted for inflation. This translates into a rise from €4.78 in capital required per passenger journey in 2002 to €11.12 per passenger in 2009. Moreover, subsidies have grown significantly, which suggests that there has been no improvement in operating efficiency in spite of the considerable capital investment. The inflation adjusted figures show an increase in subsidies from €338 million in 2002 to €451 million in 2009, an approximate one-third increase. This represents an extreme upsurge in subsidy per journey, from €1.25 to €1.79 in the seven year period. These increases in subsidies come at a time when CIÉ passenger services are worsening with significant cutbacks in operations being enforced (O'Lachtnain, 2010).

Conclusion

In conclusion, this paper firstly examines the dominance of CIÉ in the Irish bus sector and the regulatory capture the company has had over the Department of Transport. The cases for and against competition are presented. These are juxtaposed with evidence of predominantly successful competition on several bus routes across Ireland. In addition, the resistance to establishing

liberalisation on routes in Dublin is identified and compared with the largely positive experiences of foreign cities which have introduced competitive tendering for services. Furthermore, the likely continuance of protectionist policies is predicted with the establishment of the National Transport Authority, though the future practice of tendering for new routes is now written in law.

The implications of the recent Swords Express case are significant, in that the anti-competitive behaviour of state authorities with regard to bus competition is now in the public domain, after court proceedings in 2010. Furthermore, the criticism of the presiding judge directed at the Minister for Transport underlines their improper practices. Finally, an analysis of the costs of both Dublin Bus and Swords Express indicates the efficiencies that could be gained as a result of increased private sector provision of bus routes. Moreover, the evidence presented above reinforces the fact that, despite considerable growth in subsidies, CIÉ has become increasingly costly to run. Ultimately, while the reconfiguration of services in Dublin ought to allow new entrants to compete on new routes, it appears that obstacles to contestability will prevail for some time.

References

Aircoach, 2011. 'Our Service'. *Aircoach*, [Online] Available at: http://www.aircoach.ie/coach.bus.transport.dublin.airport.service.php [Accessed: February 2, 2011]

Barrett, S.D., 2004. 'Bus Competition in Ireland - The Case for Market Forces'. Economic and Social Research Institute, *Quarterly Economic Commentary*, Autumn 2004.

Barrett, S.D., 2008. 'Exporting Deregulation – Alfred Kahn and the Celtic Tiger'. *Review of Network Economics,* 7(4), Article 7.

Compecon, 2003. 'The Case for Bus Competition'. *Compecon*, [Online]. Available at: http://www.compecon.ie/buses.htm [Accessed: February 1, 2011]

Irish Times, 2010. 'Minister Erred in Bus Dispute, Rules Judge'. *Irish Times*, [Online] 3 July. Available at: http://www.irishtimes.com/newspaper/ireland/2010/0731/1224275911179.html [Accessed: February 2, 2011]

NTA, 2010. 'Bus Licensing'. *NTA*, [Online]. Available at: http://www.nationaltransport.ie/bus_licensing.html [Accessed: February 2, 2011]

O'Lachtnain, A., 2010. *Pre-budget submission to Minister for Finance*.

Richmond, J., 2000. *The Private Provision of Public Transport.* Harvard University, Kennedy School of Government: Taubman Center.

Steer Davies Gleave, 2002. 'Regulation of Bus Services Outside the Greater Dublin Area'. *Transport.ie*, [Online] 9 September. Available at: http://www.transport.ie/upload/general/2799-0.pdf

The Courts Service of Ireland, 2010. 'Digital Messenger Ltd trading as Swords Express -v- Minister for Transport'. *Courts.ie* , [Online]. Available at: http://www.courts.ie/__80256F2 B00356A6B.nsf/0/8F959BC0E75602C88025778B003DE28D?Open [Accessed: February 4, 2011]

THE BRITISH AIRPORT AUTHORITY MONOPOLY IN LONDON AIRPORTS: COMMON OWNERSHIP AND COMPETITION

Holly Bott

Senior Sophister

While discussions on transport competition often revolve around airlines themselves, in this essay Holly Bott explores competition between airports. She specifically examines the inefficiencies of the three main London airports – who until recently were owned and operated by the same company – and makes a compelling case that ensuring competition between airports is just as important as airlines.

Introduction

The British Airport Authority (BAA), the airport operator owned by Ferrovial, a Spanish construction group, has been the subject of scrutiny by the Office of Fair Trading (OFT) and the Competition Commission (CC) since early 2007. Before the sale of Gatwick in October 2009, BAA owned 7 airports in England and Scotland, including, 'the big three' in the London area, being Heathrow, Gatwick and Stansted. These airports alone accounted for more than 90% of air passenger travel in and around London, with all the BAA airports accounting for over 60% of total air passengers in the UK. With 'evidence of poor quality and high charges' (Fingleton, 2006 cited in The Independent, 2006), the 2006 OFT recommendation of a market inquiry into the BAA-owned airports was undertaken by the CC in 2007. The outcome of the detailed inquiry was the required divestment of Gatwick, Stansted and one of either Glasgow or Aberdeen airports by BAA in that order. This paper focuses solely on the London airports, since, arguably, this is where BAA wields its greatest monopoly power, and will explore the adverse effects on competition of the common ownership of the three major London airports. By way of comparison, the performance of Gatwick Airport, due to the recent transfer of ownership, is evaluated to highlight the more positive effects of increased competition between airports and a reduction in monopoly power.

Competition vs. Monopoly

Baumol's (1982) theory of contestable markets hinges on the presence of competition; markets rely on potential entrants, and more specifically on the knowledge that these potential entrants *may* enter the market, to ensure that prices tend towards long run marginal costs and move the market back towards perfect competition. The knowledge that new entrants are possible incentivises incumbents to be more efficient; banning these potential entrants only serves to make the market less contestable, encouraging laziness, complacency and inefficiency on the part of the incumbents. Protecting poor performers and restricting the exit of 'dying' firms further compounds this problem of moral hazard. How then does this relate to airports?

The Competition Commission (2009) report concluded what many have known for a long time; BAA have held a substantial monopoly over air passenger travel in the London area since the privatisation of the original, state-owned BAA in 1987. Smith (1776) notes that 'the monopolists, by keeping the market constantly understocked, by never fully supplying the effectual demand, sell their commodities much above the natural price, and raise their emoluments, whether they consist in wages or profit, greatly above their natural rate'. The common ownership of the 'big three' in London serves to effectively block out any potential entrants from the market and allows BAA to keep capacity low while charging continuously higher prices to airlines. A recommendation by the Competition Commission that the regulator of the airport industry, the Civil Aviation Authority (CAA) place an 'upper threshold on the airport charges at London Heathrow and Gatwick' (Fiddian, 2007) went largely ignored by the regulator, who allowed BAA to significantly increase its landing fees at the London airports. For example, for the period 2008-2009, the CC suggested a £10.96 cap on the charge for each passenger on board to land at Heathrow, and a £5.48 cap per passenger for Gatwick. In its 2007 proposal, the CAA allowed the charges imposed on airlines using these airports to increase to £11.97 and £6.07 respectively. The reasoning for these higher charges – increased investment for improved capacity and security at these airports – are yet to be seen in reality, with the delay of the additional planned runway at Stansted, now planned for 2019.

The Regulatory Problem

The regulator, then, can be the 'builder' of barriers, rather than the 'remover' that it should be. The common ownership of the three major airports in the London area allowed BAA to charge high prices across the board, with no significant increase in capacity or customer facilities, since it faces little competition from Luton and London City airports. The regulator, in hiking up the charges, only furthered the market from a competitive position and served, principally, the interests of the BAA monopolist. The virtues of contestability and competition are much extoled; why, then, does the regulator adopt barriers and restrict competition?

When a regulator is seen to primarily serve the interests of the incumbent(s) that it is designed to regulate, it is termed as 'regulatory capture'. Instead of allowing new entrants to enter the market and enabling competition to flourish, regulators who are 'captured' by incumbent firms, such as BAA,

protect the existing firms in the market from external competition. The 2008-2009 landing fee increase is only one example where the CAA has 'caved in to intense pressure from BAA' (Peston, 2008). Incumbent firms, protecting their own interests, will often attempt regulatory capture, but are rarely open about this strategy; BAA 'expressed disappointment' with elements of the CAA price-setting proposal, but are 'largely trying to spare the blushes of the regulator' (Peston, 2008).

BAA: Common Ownership and Competition

A principal objective in privatizing BAA in 1987 was to provide adequate airport capacity and to support airline competition. But with the current inadequate capacity and the widespread criticism of BAA's management of several of its airports, the Competition Commission, prompted by a reference under the Enterprise Act 2002 by the Office of Fair Trading, was required to carry out a market inquiry. The Competition Commission found, unequivocally, that 'BAA's common ownership of airports in South-East England…gives rise to adverse effects on competition in connection with the supply of airport services by BAA' (Competition Commission, 2009). The scope for competition in this market was determined by 'evidence on competition between airports and the circumstances in which it occurs' and 'evidence on demand substitutability between BAA airports' (Competition Commission, 2009). The report confirmed a significant degree of competition between some non-BAA airports in the UK, for example between Birmingham International Airport and East Midlands Airport; this allowed competition for new routes and on services, such as the provision of facilities for the quick turnaround times required by the low-cost carriers. Crucially, the conclusion that BAA's London airports (Heathrow, Gatwick and Stansted) faced very little competition from non-BAA airports, but that there was significant overlaps in catchment areas and substitutability of demand between the BAA London airports, prompted the suggestion that 'in the absence of common ownership, there would be competition between them' (Competition Commission, 2009).

The continuing BAA monopoly in this market had distinct but related adverse effects; airlines faced higher charges, restricting some potential entrants (mainly in the form of low-cost airlines) and, in turn, passed these increases on to the air passenger. Low-cost airlines, such as Ryanair, are the least able to pass on these increases to their customers, since they risk damaging their main source of competitiveness – their price advantage. No wonder, then, that Michael O'Leary, CEO of Ryanair, has been vocal in his calls for the break-up of the BAA monopoly. Mr O'Leary stated that 'Ryanair has fought long and hard for the break-up of the high cost, inefficient BAA airport monopoly. Over the past 3 years, traffic at Stansted Airport has declined from a high of 24m passengers to just 18m passengers in 2010. This was entirely due to the BAA's high and rising prices and the hopelessly ineffective regulatory regime run by the clueless CAA which was long ago captured by the BAA monopoly' (Dickenson, 2010). O'Leary captures the essence of the problem; both airlines and air passengers, who face the real burden of these monopoly-induced high prices, are the clear losers from such a market situation. BAA, having captured the regulator, are clear winners. Barrett (2000), notes

that, pre-deregulation of the airlines, 'the ease with which cost increases could be passed on to non-competing airlines made life easier for airport managers'; these costs have been significantly reduced by the deregulation of the airline industry and the increased competition within the sector, but, arguably, these costs remain higher than necessary in the presence of a captured regulator stifling competition within the airport market.

Disdain for the regulatory system is not limited to Ryanair alone. BAA's capture of its regulator, the CAA, did not go unnoticed by the Competition Commission; the report concludes that the 'system of economic regulation of airports is a feature which distorts competition between airlines by adversely affecting the level, specification and timing of investment and the appropriate level and quality of service to passengers and airlines' (Competition Commission, 2009). Specifically, it was BAA's common ownership of the three London airports and its effective monopoly in the market that 'exacerbates the inadequacies of the regulatory system, reducing the benefits of regulation and therefore further distorting competition between airlines' (Competition Commission, 2009). BAA may breathe a sigh of relief in this respect; their poor quality of service and high prices are, arguably, not the only factors contributing to adverse effects on competition within the market. But the point remains, whether it is due to the regulatory system or the monopolist, airlines (in the form of higher costs) and air passengers (in the form of higher air fares and poor quality of airport services) are the most impacted by a lack of competition.

Opposition to Competition

We would not expect the monopolist, BAA, to accept arguments for increased competition, even given the mounting 'evidence of poor quality and high charges' (Fingleton, 2006 cited in The Independent, 2006). BAA's CEO, Colin Matthews, warned, in an attempt to counter the move towards a more competitive market, that flight prices were likely to rise if competition authorities forced them to relinquish their monopoly control in London airports: 'clearly a smaller company would have a smaller balance sheet and that would make investment more expensive. It would, theoretically, lead to higher prices rather than lower prices' (Bowers, 2006). The case against competition between airports is not a new phenomenon; in 1976 the UK Department of Trade (1976) proposed that common ownership of airports by BAA would have significant advantages, many of which are now arguably out-dated, since the deregulation of the airline industry has demonstrated the benefits of increased competition. The claim, for example, that competition would result in the concentration of activity at a small number of airports fails to recognise or allow for the popularity afforded to small, low-cost airports. Customers have flocked to small airports, achieving substantial reductions in costs and shocking incumbents who felt they could relax in a non-competitive environment. Other contentions proposed that BAA's common ownership of airports would ensure no 'wasteful competition' and would allow for a common charging and investment policy. However, by

their very nature monopolies are wasteful, since there is little need for efficiency, and a 'common charging policy' merely equates to price fixing.

Foster (1984) and Boyfield (1984) made successful arguments against competition, resulting in the non-competitive and unitary privatisation of BAA airports in 1987. Boyfield's claim is that planning controls and long lead-in times are obstacles to competition within the market, effectively precluding new entrants. Again this is proven out-dated by the evidence: 'the traditional view that airport investment is inevitably large, requires long-run planning and must be part of an overall strategy contrasts with the experience of London City airport in the 1980s. The cost of the project was £7 million and the construction time was forty weeks' (Barrett, 2000). Similarly, Foster's (1984) assertion that, due to economies of scale, only a few airports can be profitable is dispelled by the out-performance of Cheltenham & Gloucester airport (one of the most profitable airports in the UK despite its small volume) and, comparatively, the dismal under-performance and loss made by Stansted (which supposedly benefits from economies of scale as a result of its size). With respect to under-investment, there is no viable economic reason why a market would under-invest; monopolies, in contrast, consistently restrict output, evident in the Competition Commission's report on the inadequate capacity at the three BAA London airports. In the same vein as Boyfield's (1984) initial claim, BAA proposed that planning restrictions and capacity constraints were factors adversely affecting competition in the market and that 'a lack of spare capacity would prevent competition between the BAA London airports, irrespective of ownership' (Competition Commission, 2009). The Competition Commission ultimately concluded, however, that 'BAA, as a result of its common ownership of the three BAA London airports, has itself contributed to the current shortage of capacity' (Competition Commission, 2009). Crucially, the report found that, even in the context of current capacity constraints, 'the airports can rival each other in improving service quality and compete to retain more valuable users.' (Competition Commission, 2009). Arguments that hinge on factors such as under-investment by markets and economies of scale distract from the real benefits – improved quality of service, for example – gained from increased competition.

It is clear that BAA 'failed dismally to meet passengers' needs and that a lack of competition between Heathrow, Gatwick and Stansted had consequences for the levels, quality, scope, location and timing of investment, and levels and quality of service' (Prynn, 2008). As a result of the 'damning report' (Prynn, 2008), the Competition Commission required that BAA divest Gatwick, Stansted and one of either Edinburgh or Glasgow airports, in order to address the adverse effects of its monopoly power on competition and on BAA's customers (both airlines and passengers). The Commission believes that the main benefits from the sale will result from the 'dynamic aspects of competition, improving the way in which the London airports deliver capacity in terms of timeliness, design and cost effectiveness as well as allocation to its users' (Competition Commission, 2009). It allows that current capacity shortages and price control regulations may hinder large benefits in terms of pricing,

but that increased competition would deliver 'quality and modest pricing benefits' (Competition Commission, 2009).

Increased Competition: Success of the BAA Divestment

The success of the remedies and the arguments for increased competition depend somewhat on the performance of divested airports, which serve as a contrast to the performance of those still under monopolistic control. Non-BAA owned airports also function as comparisons between competitive and non-competitive levels of quality, pricing and service. To date, Gatwick is the only BAA airport to have been sold. The airport has seen a year-on-year growth in total terminal passenger numbers of 2.7% between November 2009 and November 2010 (post-divestment), and, although somewhat speculative, 'would have been in the region of 3.7% had it not been for the severe weather disruption' (McCallum, 2010). In contrast, negative growth of -2.4% is shown for the previous year (McCallum, 2010). The strong growth was driven by 'additional capacity, with an increase of 2.2% in average load factors'. Market comparison highlights the impressive growth, with Gatwick's positive air passenger growth far above the total UK market's negative growth of -3.3%, and, similarly, Ireland's -2.8% (McCallum, 2010).

Recognition by airlines of the increased competitiveness of Gatwick airport since divestment by BAA has been noticeable; from February 2011, 'airberlin' will operate from Gatwick, a switch from its current base at Stansted, and the airport will be the only London airport to offer direct flights to Nuremburg. The growth of services and routes from 'Easyjet', 'British Airways' and 'Ryanair', as well as the relocation of 'airberlin' from Stansted is proof that 'competition is working and Gatwick is playing to win' (McCallum, 2010). Increased competitiveness at Gatwick airport allows airlines to compete further with each other on new routes and the services offered. The improved competitive environment necessitates more efficiency and a better level and quality of service; airlines benefit from the lower costs they face from the airport, and customers, in turn, benefit in the form of lower fares.

Global Infrastructure Partners (GIP), the company who purchased Gatwick airport, has spoken publicly about transforming passenger experience at the airport, with upgrades and modernisation at the heart of the transformation. Plans to 'work closely with the airlines to improve performance' have certainly been proven successful at GIP-owned London City airport, which is a 'comparatively pleasant experience for most passengers' (CAPA, 2009). The view that divestment raises the prospect of rival airport owners increasing the level and quality of service offered, or investing in better terminals to win passengers, appears to be proven in the plans for Gatwick to become a 'customer driven airport' (CAPA, 2009).

However, the timescale of such improvements and the move towards a more customer-centric approach 'in the first instance, looks to be about two to five years' (CAPA, 2009). The benefits of increased competition within the market may not be realized instantly, giving scope for those

opposing competition within the market to argue that competition has had little, and less than the desired, effect. The BAA monopoly is slowly being broken up, with Gatwick only the first of its three required divestments, but concerns, such as from Easyjet's CEO, that a 'large monopoly has been replaced by a smaller one' (CAPA, 2009), indicate that proper regulation of Gatwick, and later Stansted, will be needed to protect airline passengers from the new owners exploiting their market power. Many airlines remain cautious of the 'ineffective regulatory system for airports' (CAPA, 2009); BAA managed to create a monopoly and ensure the regulator served its own interests for twenty years, so the incentive for GIP to do the same, if such behaviour is allowed, is high.

Conclusion

BAA were found to have neglected and shown a lack of responsiveness to the interests of its customers, both airlines and air passengers. The common ownership of the 'big three' London airports, Heathrow, Gatwick and Stansted created adverse effects on competition within the market, with a failure of the regulatory system to effectively carry out its duties compounding the problem. As required by the Competition Commission, Gatwick was sold to Global Infrastructure Partners, owners of London City airport, in 2009. The growth in passenger numbers since the sale of Gatwick by BAA suggests that increased competition within the market has allowed for better management and quality of services offered at the airports, increased competition in new routes and services offered by airlines at lower cost, and, thus, clear benefits for airlines and passengers. This is the theory. It must, however, be noted that divestment of Gatwick is still very recent and that the timescale for improvements and transformation of the airport looks to be from two to five years. Any improvements on competition and, thus, on benefits to airlines and passengers may suffer from a time lag before fully being realized. The regulatory system failed airlines and passengers in many respects; proper regulation must be in place to ensure that the new owners of Gatwick do not become 'another BAA' by exploiting their market power at the expense of its customers.

References

Competition Commission, 2009. 'A Report on the Supply of Airport Services by BAA in the UK'. *Competition Commission*, [Online] 19 March. Available at: http://www.competition-commission.org.uk/rep_pub/reports/2009/fulltext/545.pdf [Accessed: January, 2011]

Barrett, S., 2000. 'Airport Competition in Deregulated European Aviation Market'. *Journal of Air Transport Management,* 6(1), pp.13-27.

Baumol, W., 1982. ‘Contestable Markets: an Uprising in the Theory of Industrial Structure’.*American Economic Review,* 72(1), 1-15.

Bowers, S., 2006,. *Airports Operator Faces Break-up in Monopoly Inquiry*. *The Guardian*, [Online] 13 December. Available at: http://www.guardian.co.uk/business/2006/dec/13/transportintheuk [Accessed: January 27, 2011]

Boyfield, K., 1984. ‘Competition and Regulation in Privatising the British Airports Authority’. *Public Money General Series*, 4.

CAPA (Centre for Asia Pacific Aviation), 2009. (2009, October 23). ‘Gatwick Airport Finally Sold while BAA Puts Up a Last Stand to Prevent the Sales of Other Airports’. *Centre for Asia Pacific Aviation*, [Online] 23 October. Available at: http://www.centreforaviation.com/news /2009/10/23/gatwick-airport-finally-sold-while-baa-puts-up-a-last-stand-to-prevent-the-sales-of-other airports/page1 [Accessed: January 27, 2011]

Dickenson, H., 2010. ‘Ryanair Condemns BAA Monopoly's Latest Appeal’. *Flightline Travel News*, [Online] 15 November. Available at: http://www.flightline.co.uk/travelnews/ryanair-condemns-baa-monopoly%E2%80%99s-latest-appeal/ [Accessed: January 27, 2011]

Fiddian, P., 2007. ‘CAA Plans Landing Fee Increases at London Airports’. *Airport News*, [Online 20] 20 November. Available at: http://www.airport-int.com/news/caa-plans-landing-fee increases-at-london-airports.htm [Accessed: January 28, 2011]

Foster, C., 1984. ‘Privatising British airports: what's to be gained?’. *Public Money General Series*, 4.

McCallum, A., 2010. ‘Gatwick November 2010 Traffic Performance Summary’. *Gatwick Airport*, [Online] 13 December. Available at: http://www.gatwickairport.com/business/mediac entre/press-release/gatwick-november-figures.htm [Accessed: January 28, 2011]

Peston, R., 2008.. ‘BAA to Raise Airport Landing Fees’. *BBC News: Business*, [Online] 11 March. Available at: http://news.bbc.co.uk/2/hi/uk_news/7288937.stm [Accessed: 27 January, 2011]

Prynn, J., 2008. April 22. ‘Failing BAA will be forced to sell Gatwick for £3 billion’. *London Evening Standard*, [Online] 22 April. Available at: http://www.thisislondon.co.uk/standard/article 2349714-failing-baa-will-be-forced-to-sell-gatwick-for-3-billion.doc [Accessed: January 27, 2011]

Smith, A., 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations.* London: W. Strahan and T. Cadell.

The Independent, 2006. 'Airport Expansion Schemes Hit by OFT's Call for Ownership Inquiry'. The Independent, [Online] 13 December. Available at: http://www.independent.co.uk/news/business/news/airport-expansion-schemes-hit-by-ofts-call-for-ownership-inquiry-428263.html

UK Department of Trade, 1976. 'An Airport Strategy for Great Britain'. *HMSO*, London.