ISSN (print) 2047-8747
ISSN (online) 2047-8755

Volume 8 Issue 3
(2019)

Interdisciplinary Journal of Economics and Business Law

University of Warsaw
Faculty of Management Press

Submission Guidelines

Books for review should be sent to Prof. Kenneth S Friedman, Regis University, 26 Willow Lane, Black Hawk, Colorado, 80422-4203, USA Subscription rates per volume of 4 issues and online access, postpaid £160/US$250. Special rates for individuals, such as students, who warrant that the journal is for their own use, or those who have no access to company or institutional funds, and order directly from the publisher, postpaid: £70/US$110.

Rates and **publication** dates are subject to adjustments without notice. Subscriptions requests should be sent to ruth.taplin.ijebl@gmail.com

Interdisciplinary Journal of Economics and Business Law
www.ijebl.co.uk

Notes for Contributors

The official language of the journal is English. Papers should be sent by e-mail to the Editor, Professor Ruth Taplin; Email: ruth.taplin.ijebl@gmail.com, along with an abstract of no more than **200** words if possible.

Authors should aim to make their papers understandable to those in other disciplines and wherever possible, purely mathematical/statistical arguments, when they cannot be avoided, should be banished or put into an appropriate appendix. The length of journal articles must not exceed **30 A5** pages (**papers must be submitted in A5 size**) including the references and submitted to IJEBL in black and white so any illustrations should be in shades of grey. Book reviews should be between 1000 and 1500 words. Material will be accepted for review on the understanding that it has not been submitted for publication elsewhere and the author's copyright will be transferred to the publisher in the event of publication. Please see the copyright agreement that must be agreed to by all authors on the website www.ijebl.co.uk.

The pages should be of uniform size (**A5** size) and numbered consecutively. Single spacing should be used for abstracts, references and footnotes. The manuscripts should not contain **more than a few minor corrections**. It is the responsibility of the author(s) to obtain any permission necessary to reproduce figures or tables or quote from published work(s); suitable attribution should always be made. Unpublished manuscripts or letters are also protected.

Titles and Sub Titles should be typed on a separate line without indentation in capitals and lower case. They should not be underlined. Long titles and superfluous capitals should be avoided. Any instructions to the copy editor or printer should be highlighted in red.

The following order should normally be used: First page: title, author(s), name(s), affiliation(s) in this order should be **above** the abstract; present address(es) of author(s), e-mail addresses should be footnoted starting from number one. In the case of multi-authored works, the full address of the person to whom proofs are to be sent should be given in the covering e-mail when manuscripts are sent to IJEBL.

References should be kept within the main body of the text (e.g. author 2019) and full references with pages and bibliographies at the end of the article in alphabetical order. Figures and tables should be prepared in a jpeg format and sent as a separate electronic file to the article. Each figure should have a caption. Footnotes should be only used if it is impossible to incorporate the information in the main body of the text. If used they should be identified by subscript numerals and kept short and at a minimum. An exact page reference should be given in the text for verbatim quotations. Multiple entries by an author or set of authors in the same year should be post scripted a, b, c, (e.g. 2019a, 2019b, 2019c) etc. The list of references should follow the customary conventions.

Proofs will be sent to the authors for correction and should be returned within the requested time by e-mail. Author's alterations in excess of 10% of the typesetting cost will be charged. The Editor reserves the right to publish a paper without the author's own corrections in cases of undue delay in returning proofs.

Papers submitted to Interdisciplinary Journal of Economics and Business Law are refereed.

The journal is covered in the following indexing and abstracting services: American Econ. Lit., International Bibliography of Social Sciences (Pro-Quest), Cabell's Directory, listed in Heterodox Economics Directory 10th Edition 2019 and included in the Guide to Academic Journal Quality (formerly ABS Guide) Australian Business Deans Council Journal Quality List, SSRN and in many other countries' lists.

Aims and Scope

The rationale of IJEBL is to provide a broad international forum for expanding the boundaries of economics and business law seeking solutions to global problems enabling academics and practicing economists, business people and legal practitioners to tackle the problems within this environment and context. We invite both academics and practitioners to contribute to assist in expanding the boundaries of the social sciences which have not as yet come to terms with the profound change in the boundaries of disciplines, especially since the advent of nanotechnology. Prior to this technological development that is based on interdisciplinary research, those of us who supported the interdisciplinary approach were even greater pioneers. The importance of intangible assets in the form of intellectual property is increasing as is the need to understand valuation of these assets so that the true market value of brand, image, copyright and so forth can be understood. Changes in both corporate and company law is crucial for both large companies and SMEs to be successful. This is why we need to include business law which is one of the most rapidly expanding interdisciplinary disciplines.

The aim of the IJEBL is to not only provide an independent forum for solving global economic and business problems but to expose readers to the exciting new developments and free thinking that the interdisciplinary approach continues to offer as an alternative to a narrow and rigid econometric approach. We also wish to provide readers with a taste of the books available on the market that can further such understanding through reviews of cutting edge books that contribute to an interdisciplinary approach in each issue. Articles, comments, short papers, letters and book reviews are all welcome.

Welcome to the world's only open, independent, interdisciplinary forum born out of 20 years of my being Editor of the Journal of Interdisciplinary Economics and a founder member of the late Ken Penney's visionary idea while we were at Exeter University, UK. The original founder members remain on the board of IJEBL and are committed as Ken was to seeing the world through an interdisciplinary perspective. With the continuing lack of ethical responsibility in corporate governance and cross border analysis of intellectual property coupled with the dearth in understanding valuation which can add to the very much needed prosperity of businesses whether large or small the need for an interdisciplinary approach is required more than ever. Those who contribute to IJEBL continue to spurn seeing the world through numbers, compartmentalised disciplines and conformity, making a contribution to creative and challenging solutions to the current ills of the world economy with positive ideas and criticisms that can be implemented on a practical level and ethically. We are interested in innovation, entrepreneurship, solutions within the rule of law, the environment, the health of the planet/wildlife and cultural specificities. We have a global based support network of readers, contributors and subscribers and would like you to join us. This is Issue 3, Volume 8, IJEBL 2019 published January to January although subscriptions can start from October to January each year. An Abstract page to give a broad sample of contents has been added online and updated for Volume 8. Many thanks to all those who support our open, independent, interdisciplinary forum into our eighth year of existence.

IJEBL is published both online and as printed copy quarterly. Additionally, a fifth online issue has been added gratis that focuses on postgraduate papers to afford young contributors to publish for the first time in an international journal. Other extra gratis IJEBL special issues are being prepared too.

Please see www.ijebl.co.uk for guidelines for contributions and subscriptions which we rely on to exist.

Editor

Acknowledgements
We are grateful to the University of Warsaw, Faculty of Management Press for their printing and distribution support of IJEBL.

Interdisciplinary Journal of Economics and Business Law, 2019, Vol. 8,
Issue 3, pp. 8-24

Determinants of Bank Liquidity: The role of Regulatory Capital and profitability

Isaiah Oino[1]

Abstract

Liquidity of banks is paramount for well-functioning financial institutions because it enables them to perform the financial intermediation and credit creation role. Applying pooled fixed effects model on a panel of 100 largest banks from 2008 to 2016 in the European Union area, the results show interesting findings. The results indicate that, better capitalised banks are expected to suffer fewer distortions in lending decisions and perform better. In addition, the findings also indicate that higher capital requirements for banks may provide incentives for the bank to reduce its probability of default by monitoring its borrowers and reduce moral hazard by incentivising banks to invest in less risky assets. Furthermore, there is an inverse relationship between growth in loans and both tier 1 and total regulatory capital. This implies, when the banks are building their capital level, they are likely to reduce their lending. The study suggests that, banks should ensure that the quality of the capital base and the buffers above the regulatory minimum are built up during periods of strong earnings growth during economic boom so that they are available to absorb greater losses in stressful environments.

Introduction

Banking is one of the most important sectors in the economy because of the intermediation role that banks play. In order to perform such role effectively, apart from being efficient, it is paramount that the banks are liquid enough to meet the financial commitments. One may say that liquidity is a product of efficiency. Liquidity may be impaired as a result of sudden unexpected cash outflows by way of large deposit withdrawals, large credit disbursements, unexpected market movements or crystallisation of contingent obligations. The other cause may be because of some other event causing counterparties to avoid trading

[1] Coventry University, Email: ac3497@coventry.ac.uk

with or lending to the bank. Researchers have investigated liquidity risk on the side of the dependent variable (Bunda and Desquilbet, 2008). Liquidity risk has a spiralling effect and often tends to compound other risks such as credit risk and market risk. So, prior approach in assessing liquidity is somewhat static hence the need for a more dynamic approach. This is because as Moore (2010) indicated that liquidity in banking decreased by 8%. This is because among other factors is that, banks were not willing to lend to each other and borrowers were not in position to honour their loan commitments.

Considering banks differ significantly from non-financial firms by the fact that the bulk of the bank's liabilities are subject to payment on call, it is paramount that at all time the resources of the bank provide the means for meeting demands for cash as they are made. It might appear that this would require banks to hold against these liabilities resources which are also payable on call.

The overarching objective of this paper is to assess a host factors that may influence liquidity in the selected banking sector using a host of estimation method in order to draw a comparison. The article analyses how the regulatory capital not only influence liquidity but also how it impacts profitability of the banks. In addition, the procyclical nature of regulatory capital is assessed in the context of 100 major banks in 10 biggest European Union economies.

Literature review

Managing liquidity in banking is one of the most important elements that bank supervisors or regulators keep a close eye to. A bank's liquidity methodology should maintain sufficient liquidity to withstand all kinds of stress events that will be faced. Constant assessment of liquidity risk management framework and liquidity position is an important supervisory action that will ensure the proper functioning of the bank. The banking literature has identified a host of factors that influence the level of liquidity. More significant is that deposit insurance protects banks from runs and facilitates liquidity transformation (Diamond and Dybvig, 1983). Pana et al. (2011) concluded that a higher level of deposit insurance before a bank merger leads to the higher level of liquidity creation after the merger. Merger of banks is likely to lead increase in size. Bigger banks will have more demand than small banks. Indeed, Berrospide (2013) shows that the rate of hoarding of liquidity decreases with the size of banks. That is as the banks grow is size and so is its liquidity holding. However, Laeven, Ratnovski and Tong (2014) argue that the big banks have a better possibility for diversification, which in turn reduces the risk of the portfolio, thereby allowing banks to hold less capital and use less stable funding. On the other hand, there is an assumption that smaller banks create and hold more liquid assets, as they have to rely more on themselves.

In terms of the level of capital and liquidity, Berger and Bouwman (2009) noted that banks with a lower level of equity capital are more likely to commit to monitoring their borrowers and hence allow them to grant loans and to create more liquidity. That is, a higher level of capital increases liquidity creation because a higher level of capital improves an institution's ability to absorb and diversify risk and hence creates more liquidity. Similarly, Pana et al. (2011) also find the positive relationship between capital and liquidity creation for large banks but find no evidence for small banks. Also, Bunda, & Desquilbet (2008), in their study on 1107 commercial banks in 36 emerging economies, find that capitalization measured by the ratio between equity and total assets has a significant and positive relationship with all liquidity measures considered in their study and a significant relationship with inflation rate and growth rate. However, Mukherjee and Pana (2010) report the negative relationship between capital and liquidity creation for credit unions. Other studies have noted that different levels of liquidity can be created by changing the funding mix on the liabilities (Diamond and Rajan, 2001; Gorton and Winton, 2000), and the amount of equity capital impacts the level of lending and the asset portfolio composition, and thus affects liquidity transformation (Thakor, 1996).

In analysing the impact of liquidity, a study of Kumbirai and Webb (2010) investigated the relationship between the liquidity, profitability and credit quality performance of South Africa's commercial banks for the period of 2005-2009. The study concluded that, the global financial crisis had adverse effect on South African banks which had resulted into falling profitability, low liquidity and deteriorating credit quality of the banking sector. Further, Bordelean and Graham (2010) analysed the association between liquidity and profitability for a panel of Canadian and U. S. banks for the period of 1997-2009. Their study indicated that, there is a non-linear relationship between liquid assets and profitability. This implies that profitability is improved for banks that hold some liquid assets but excess of holding liquid assets diminished the bank's profitability.

Other studies have analysed the impact of macroeconomic variables on liquidity. Regarding GDP, Valla, SaesEscorbiac and Tiesset (2006) noted that the appetite for higher creation of liquidity increases with better economic conditions than at a time of economic downturn. Also, Cucinelli (2013) found similar results that liquidity is increased during economic boom. Conversely, Aspachs, Nier and Tiesset (2005), Vodová (2011a) and Moussa (2015) document higher liquidity holdings in a period of economic downturn, when holding is motivated by the principle of precaution from banks, but also by less demand for loans from clients. This demonstrate lack of consensus in terms of impact of economic growth on banking liquidity.

Data and Methodology

For bank level data, this research makes use of Fitchconnect database to extract unconsolidated statement of financial position and income statement data, for a panel of 100 largest banks within 10 biggest economies in the European Union area. Aggregate annual data was used over the period 2008 to 2016. To ensure uniformity we use consolidated data where necessary. Bank holding, private and retail and consumer banks with complete data set that were included to ensure the panel is balanced. World Bank database was used to extract country specific data i.e. real gross domestic product growth. To test the relationship between the bank liquidity and the bank-specific factors and macroeconomic determinants, we estimate a linear regression model, which takes the form:

$$Liquidity = \alpha + \sum_{J=1}^{j} \beta x_{it}^{j} + \sum_{M=1}^{M} Bx_{ts}^{M} + \varepsilon_{its} \quad (1)$$

where $$\varepsilon_{its} = v_i + \mu_{its} \quad (2)$$

In addition, $\sum_{J=1}^{j} \beta x_{its}^{j}$ is a vector of bank-specific variables, the loans to deposit ratio, profitability, loan growth, tier 1 capital, total regulatory capital, growth of the bank, equity to total assets and growth of loans.β^{s} are the coefficients of the variables.

Also, $\sum_{M=1}^{M} Bx_{ts}^{M}$ is the country's macroeconomic variable, gross domestic product growth. βs are coefficients of the variables.

To evaluate the stationarity of the variables in the model, panel unit root test was used, which is applicable to balanced panel data (Fisher-type tests based on augmented Dickey–Fuller). Stationarity mean that the mean, variance and autocorrelation of a variable do not change with time. The results indicate that all the variables are stationary. However, some outliers were identified because of the timeframe that was considered in this research. Without data cleaning to remove or smooth off the outliers, Robust regression method was used. Robust regression methods are designed to be not overly affected by violations of assumptions by the underlying data-generating process.

Assessing the drivers of bank liquidity could be affected by the potential for an endogenous character of certain variables. For instance, the more profitable a bank is, the more likely it is to increase its equity capital. This problem of causality could even move in the opposite direction; for example, higher bank profitability could lead to more employees and less efficiency (García-Herrero et al., 2009). In addition, some characteristics of banks that affect their liquidity are difficult to measure or identify in an equation (the so-called unobserved heterogeneity); if the influence of such characteristics is not taken into account, there could be correlations between some of the coefficients of the explanatory variables and the error terms that bias these coefficients.

That is: $$\mathrm{E}\,(\mu_i / X_i) \neq 0. \quad (3)$$

The dependent variable liquidity is as described in Kashyap and Stein (2000). Liquid assets are composed of cash, reverse Repos, bills and commercial papers and comprise in addition all types of investments securities, such as equities and bonds. This research measure liquidity as a ratio of liquid assets to total assets. This measure is interesting since it informs on the split between liquid and illiquid asset (such as loans) on the bank's balance sheet.

Liquidity estimation model was captured using Robust least square, stepwise regression and pooled fixed effects (FE) after testing for heteroskedasticity and serial correlation using the Breusch Pagan with White test procedure as a possible remedy. The Akaike information criterion for FE is smaller than any other estimation method. Therefore, the FE is a better estimator. Using the Hausman test between the fixed effects and the random effects, the result indicates that the fixed effects is more appropriate, implying that the differences between the banks do not vary greatly from period to period. Although the time effects were not significant, to control specific country effects we use gross domestic growth as this is likely to differ from country to country. The results from the three estimation methods are reported for meaningful comparison. Also, it is paramount to test whether the regression is correctly specified. As shown below using the Ramsey regression equation specification error test (RESET) test, the F-statistic takes the form of:

$$F_{(M;N-k-1)} = \frac{(SSR_{\hat{Y}} - SSR_{\hat{Y}^2})/M}{SSR_{\hat{Y}^2}/(N-K)} = \frac{(SSR_R - SSR_{UR})/M}{SSR_{UR}/(N-K)} \quad (4)$$

where SSRs are the sum of squared residuals for the respective regressions; SSR_R is the sum of squared residuals of restricted regression; and SSR_{UR} is the sum of squared residuals of unrestricted regression. M is the number of restrictions; N is the number of observations; and K is the number of parameters estimated in the unrestricted equation. The F test is compared with the p value; the regression is correctly specified and have not included irrelevant variables. The inclusion of an irrelevant variable may result in correlation with other variables, which could cause multicollinearity.

Table 1. Ramsey RESET test

F-statistic	0.36915	Prob. F (1,24)	0.6742
Log likelihood ratio	0.622340	Prob. chi-square(1)	0.5104

Table 2. Measurements of the variables

Variable	Explanation	Computation
Dependent variable		
Liquidity ratio	Buffer of liquid assets as a share of the balance sheet. It is a measure of the maturity structure of the asset portfolio that can reflect excessive maturity unbalances (Cihak and Poghosyan, 2009).	Liquid assets/ total assets
Loan to Deposit ratio	The loan-to-deposit ratio is used to assess a bank's liquidity by comparing a bank's total loans to its total deposits for the same period. It is a measure of the illiquidity of the asset portfolio that can reflect excessive illiquidity and higher exposure to default risk (Arena, 2005).	Total loans/ total deposit
Idiosyncratic factors		
Profitability	Profitability of the bank as a ready source of liquidity (analysed in the regression in level)	Net profit/ total equity

Variable	Explanation	Computation
Loan growth	Financial constraints. Ability to raise new funds if loan business expands compared to the rest of the balance sheet	% of Loan growth
Economic growth	The status of the economy, Gross Domestic Product Growth	GDPG
Capital adequacy Tier 1	The minimum regulatory capital requirement as a ratio based on the risk-weighted assets. (RWCR). Tier 1 capital includes common equity plus other instruments that are subordinated to subordinated debt, have no fixed maturity and no embedded incentive for redemption, and for which a bank can cancel dividends or coupons at any time.	RWCR = Capital/Risk Weighted assets
Total regulatory capital	This include tier 1 and tier 2. Tier 2 consists of unsecured subordinated debt and its stock surplus with an original maturity of fewer than five years minus investments in non-consolidated financial institutions subsidiaries under certain circumstances.	sum of Tier 1 and Tier 2 capital.
		Total Loans/ Total Deposit
Growth in size of the bank	Changes in the bank size using the log of total assets.	Growth of total assets.
Equity/ total assets	It represents the amount of assets on which shareholders have a residual claim.	Total shareholders' equity/ total assets

Analysis of the Results

Table 3 below indicate that on average the ratio between equity and total assets is 12.3 while that maximum is 162.00 and the minimum 0.01 with a huge standard deviation. This ratio represents the amount of debt and equity used to finance the bank assets. This is an important ratio as it indicates the potential financial risk. A relatively high ratio i.e. 2 commonly indicates an aggressive growth strategy by a company. However, one would expect that with the banks this ratio is quite high compared with non-financial institutions because of the nature of the business that makes use of deposits. In terms of profitability, the results indicate ROE is 2.72 on average while the maximum is 10.76% and the minimum is -151.82%. interestingly as shown in table 4, there is a positive association between profitability and regulatory capital requirements. This confirms Berger (1995)'s study on U.S. commercial banks where he found a positive relationship between the level of capital and earnings in banks. Also, Goddard, Molyneux and Wilson (2009) found that the relationship between the capital-asset ratio and profitability is positive for banks in the Euro-area. Well capitalised banks face lower bankruptcy costs, which reduces the cost of funding and thereby increasing the profitability. Focusing on traditional intermediation role as loans and deposits, the average share of total loans to total deposit in total assets is 95.11%%. As expected, there is a positive association between loans to deposit and ROE.

In terms of bank liquidity, the liquid assets to total assets ratio was employed. This ratio measures the maturity structure of the asset portfolio that can reflect excessive maturity unbalances (Cihak and Poghosyan, 2009). The higher is the ratio, the more liquid an institution is considered. As shown in table 3 below, on average 18% of the total assets were classified as those that can be converted to cash quickly if needed to meet financial obligations; examples of liquid assets generally include cash, central bank reserves, and government debt. To remain viable, a financial institution must have enough liquid assets to meet its near-term obligations, such as withdrawals by depositors. Table 4 indicates that there is a positive association between liquid assets to the both tier 1 and total regulatory capital. This is important for the regulators as it confirm the need to increased monitoring or supervision of regulatory capital that can enhance liquidity in the market. However, there is a negative relationship between liquidity and profitability. This implies that there is a need to strike a balance between liquidity and profitability.

Table 3. Descriptive statistics

	Mean	Maximum	Minimum	Std. Dev.	Obs
EQ/TA	12.298	162.000	0.014	88.202	339
GDPG	0.526	4.079	-5.618	2.089	339
G.LOANS	2.477	202.500	-100	16.963	339
G.TA	4.269	68.890	-57.9	12.339	339
LIQ	18.006	80.6600	0.010	18.922	339
LOANS/DEP	95.108	465.000	0.618	60.432	339
TIER 1	16.114	49.500	5.950	7.407	339
ROE	2.717	40.760	-151.82	14.917	339
T.REG. CAP	17.325	49.500	4.550	6.850	339

EQ/TA (Equity to total assets); GDPG (Gross Domestic Product growth); G.LOANS (Growth of Loans); G.TA (Growth of Total Assets); LIQ (Liquid assets to total assets); LOANS/DEP (Total loans to total deposits ratio); TIER 1 (Regulatory Tier 1 Capital); ROE (Return on Equity); T.REG.CAP (Total Regulatory Capital, Tier 1 plus Tier 2).

As a proxy of bank capitalisation, the ratio of Tier 1 and total regulatory capital to total risk weighted assets were considered. A bank could be more vulnerable when its capital is weaker compared with the volume of its risky assets (Martin, 1977). In this context, bank security buffer could be too weak to absorb losses from bad quality assets. As shown in table 3 above, tier 1 is 16.11% and total regulatory capital is 17.33%. Total regulatory capital is the sum of tier 1 and tier 2. Tier 2 capital consists of capital instruments and subordinated loans and their associated premium accounts. The claim on the instrument or loan must be wholly subordinated to the claims of all non-subordinated creditors, and should not be secured or subject to a guarantee that enhances the seniority of its claim. As of 2017, under Basel III, a bank's tier 1 and tier 2 capital must be at least 8% of its risk-weighted assets. The minimum capital adequacy ratio (including the capital conservation buffer) is 10.5%. The capital conservation buffer recommendation is designed to build up banks' capital, which they could use in periods of stress. Examining the relationship between regulatory capital requirements and lending, the results (table 4) indicate that there is an inverse relationship between

Table 4. Correlation Matrix

	EQ/ TA	GDPG	G. LOANS	G.TA	LIQ	LOANS /DEP	TIER 1	ROE	T.REG. CAP
EQ/TA	1								
GDPG	0.021	1							
G.LOANS	0.003	-0.029	1						
G.TA	0.018	-0.161	0.596	1					
LIQ	0.056	0.002	-0.023	0.114	1				
LOANS/ DEP	-0.019	-0.260	0.057	0.028	-0.128	1			
TIER 1	0.026	0.050	-0.043	-0.053	0.167	-0.004	1		
ROE	0.028	0.063	0.121	0.139	0.199	0.014	0.122	1	
T.REG. CAP	0.022	0.049	-0.075	-0.089	0.155	0.070	0.866	0.030	1

EQ/TA (Equity to total assets); GDPG (Gross Domestic Product growth); G.LOANS (Growth of Loans); G.TA (Growth of Total Assets); LIQ (Liquid assets to total assets); LOANS/DEP (Total loans to total deposits ratio); TIER 1 (Regulatory Tier 1 Capital); ROE (Return on Equity); T.REG.CAP (Total Regulatory Capital, Tier 1 plus Tier 2).

growth in loans and both tier 1 and total regulatory capital. This implies, when the banks are building their capital level, they are likely to reduce their lending. This is in line with Furfine (2000) who noted that a one percentage point increase in risk-based capital requirement results in 5.5% reduction in loan growth. Similarly, Francis and Osborne (2009) find that a one percentage point increase in capital requirements would reduce lending in 2002 by 1.2%.

The status of economy can play a significant role in building bank capital buffer. In terms of economic growth, on average, the major economies in Europe grew by 0.5% between 2008 and 2016. As shown in table 4 above, there is a positive association between GDPG and both tier 1 and total regulatory capital. This indicate the pro-cyclical nature of the regulatory capital which implies the need for the banks to build up their capital level during the economic boom i.e. making the hay while the sun shines.

Using various estimation methods, one is able to assess the significance of various variables. Using various approach, we narrow the reporting to three models. That is using Robust Least Square, stepwise and pooled fixed effects for comparison purposes. As shown in table 5, regression, the Akaike Information Criteria (AIC) for pooled fixed effect model is lower than the rest. AIC estimates the quality of each model, relative to each of the other models. In this case, pooled fixed effects model gives the best fit. The results show that ROE has positive coefficients and significant at 1% across all models. This implies the more profitable the bank is, the less likely to be cash strained. Interestingly, apart from the significance level, the magnitude of the coefficient is quite big. That is a 1% increase in profitability will lead to 0.295 improvement in liquidity. The findings support that of Bordelean and Graham (2010) concluded that profitability is improved for banks that hold some liquid assets but excess of holding liquid assets diminished the Canadian and U.S. banks bank's profitability between 1997-2009.

The theory of financial intermediation highlights various channels through which capital and liquidity are interrelated. Table 5 indicate that both tier 1 and total regulatory capital are significant at 5% and 1%

respectively. Liquid banks are able to meet their financial intermediation role. This implies that forcing banks to hold a significantly higher capital not socially expensive in the long-term. Better capitalised banks are expected to suffer fewer distortions in lending decisions and perform better. Moreover, regulatory approaches based on equity should dominate alternatives (including contingent capital).

Table 5. Regression Results

Variables	Robust Least Square	Stepwise Regression	Pooled Fixed Effects
C	11.393** (3.545)	14.120** (3.361)	14.001** (0.991)
EQ/TA	0.010 (0.012)	0.010 (0.011)	0.010* (0.003)
GDPG	0.246 (0.513)	0.276 (0.486)	0.015 (0.153)
G.LOANS	-0.017 (0.048)	-0.017 (0.046)	-0.144** (0.021)
G.TA	0.229* (0.095)	0.228* (0.090)	0.350*** (0.031)
LOANS/DEP	-0.044* (0.019)	-0.044* (0.018)	-0.035** (0.005)
TIER 1	0.712* (0.335)	0.652* (0.317)	0.478** (0.083)
ROE_	0.296*** (0.072)	0.300*** (0.068)	0.295*** (0.021)
T.REG.CAP	0.560*** (0.168)	0.503*** (0.159)	0.029*** (0.090)
Adjusted R squared	0.107	0.118	0.116
AIC	354.89	8.67	8.59

Dependent variable: LIQ (Liquid assets to total assets).

Independent variables: EQ/TA (Equity to total assets); GDPG (Gross Domestic Product growth); G.LOANS (Growth of Loans); G.TA (Growth of Total Assets); LOANS/DEP (Total loans to total deposits ratio); TIER 1 (Regulatory Tier 1 Capital); ROE (Return on Equity); T.REG.CAP (Total Regulatory Capital, Tier 1 plus Tier 2).

The results imply that higher capital requirements for banks may provide incentives for the bank to reduce its probability of default by monitoring its borrowers and reduce moral hazard by incentivising banks to invest in less risky assets. This is based on Berger and Bowman (2013) argument that higher capital resources enhance the probability of individual bank survival and the maintenance of market share for medium- and large-sized banks during banking crises. However, on the negative side, banks may increase risk-taking activities and conceal them by reclassifying the assets into different risk segments.

Examining the impact of loans to deposit ratio and increase in loans on liquidity indicates a negative coefficient but significant at 5%. This means increased lending may have a detrimental impact on liquidity. That is a 1%-point increase leads to 0.035 decrease in liquidity. As lending is a major source of revenue, there is a need to balance profitability and liquidity.

All the three models also indicate the impact of bank size on liquidity. The results indicate that an increase in total assets leads to improved liquidity. This implies that larger banks have lower liquidity exposure. This is because larger banks have a better reputation and so are less exposed to the liquidity risk. Surprising, although a positive coefficient, GDPG indicate that it is not significant in influencing liquidity. This implies an increase in liquidity in response to stronger growth appears to be only marginally significantly different from zero.

Conclusion and Recommendations

The importance of liquid and well functioning financial institution is undoubtably significant for the economic and social well-being. The study has demonstrated that there is a positive association between profitability of the banks and economic growth. That is, when there is economic boom, the profitability of the banks is likely to be enhanced. This is may influence the private credit to GDP which in turn could influence employment level. Also, the study has shown that banks tend to increase the regulatory capital during the economic growth which demonstrate the procyclical nature of the bank capital. Therefore, banks should ensure that the quality of the capital base and the buffers above the regulatory minimum are built up during periods of strong earnings growth during economic boom so that they are available to absorb greater losses in stressful environments.

The results also highlight that there is a positive significant relationship between growth in total assets and liquidity. This imply that large banks are likely to be more liquid. In other words, as banks grow, there is need for them to develop a model that maps growth and liquidity. Further, a 1% increase in loans grant leads to 0.144 percentage point decrease in liquidity. However, as banks generate much of its income from loans, it is an activity that cannot be avoided. Therefore, there is a need to strike a balance between liquidity and profitability in addition to enhance credit evaluation of the borrowers.

References

Arena, M. (2005). Bank failures and bank fundamentals: A comparative analysis of Latin America and East Asia during the nineties using bank-level data", Bank of Canada, Working paper.

Aspachs, O., Nier, E. and Tiesset, M. (2005). "Liquidity, Banking Regulation and the Macroeconomy. Evidence on bank liquidity holdings from a panel of UK resident banks". Bank of England Working Paper.

Berger, A.N. and Bouwman, C.H.S. (2009). Bank liquidity creation, *Review of Financial Studies*, Vol. 22, pp. 3779-3837

Berger, A.N. and Bouwman, C.H.S. (2013). How does capital affect bank performance during financial crises? *Journal of Financial Economics*, vol. 109, pp. 146–76.

Berrospide, J. (2013). Bank Liquidity hoarding and the Financial Crisis: An Empirical Evaluation. *Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs*, 2013-3: 1-41.

Bordelean, E. and Graham, C. (2010). "Impact of Liquidity on Bank Profitability", Bank of Canada Working Paper No. 2010-xx, pp. 1-24.

Bunda, I. & Desquilbet, J. (2008). The bank liquidity smile across exchange rate regimes. *International Economic Journal*, 22(3): 361-386.

Cucinelli, D. (2013). The Determinants of Bank Liquidity Risk within the Context of Euro Area. *Interdisciplinary Journal of Research in Business*, 2(10): 51-64.

Cihak, M., Poghosyan, T. (2009). "Distress in European banks: An analysis based on a new data set", Working Paper.

Diamond, D.W. and Dybvig, P.H. (1983). Bank runs, deposit insurance, and liquidity, *Journal of Political Economy*, Vol. 91, pp. 401-419.

Diamond, D.W. and Rajan, R.G. (2001). Liquidity risk, liquidity creation, and financial fragility: a theory of banking, *Journal of Political Economy*, Vol. 109, pp. 287-327.

Francis, W. and Osborne, M. (2009). "Bank regulation, capital and credit supply: measuring the impact of prudential standards", Financial Services Authority Occasional Paper Series.

Furfine, C.H. (2000). Interbank payments and the daily federal funds rate, *Journal of Monetary Economics*, vol. 46:2, pp. 535-53.

García-Herrero, A., Gavilá, S. and Santabárbara, D. (2009). What explains the low profitability of Chinese banks? *Journal of Banking and Finance*, 33 (11): 2080-2092.

Goddard, J., Liu, H., Molyneux, P. and Wilson, J. (2009). "Do bank profits converge?" Working Paper.

Gorton, G. and Winton, A. (2000), "Liquidity provision, bank capital, and the macroeconomy", University of Minnesota working paper.

Kumbirai, M. & Webb, R. (2010). A Financial Ratio Analysis of Commercial Bank Performance in South Africa, *African Review of Economics and Finance*, 2(1): 30-53.

Laeven, L., Ratnovski, L. and Tong, H. (2014). "Bank Size and Systemic Risk". IMF Staff discussion note, May 2014.

Moore, W. (2010). "How do financial crises affect commercial bank liquidity? Evidence from Latin America and the Caribbean." MPRA Paper 2010-21473. Munich: Munich Personal RePEc Archive.

Moussa, M.A.B. (2015). The Determinants of Bank Liquidity: Case of Tunisia. *International Journal of Economics and Financial Issues*, 5(1): 249-259.

Mukherjee, T. and Pana, E. (2010). "Credit unions as liquidity creators", working paper.

Pana, E., Park, J. and Query, J.T. (2011). The impact of bank mergers on liquidity creation, *Journal of Risk Management in Financial Institutions*, Vol. 4, pp. 74-96.

Rauch, C., Steffen, S., Hackethal, A. and Tyrrel, M. (2010). "Determinants of Bank Liquidity Creation". Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1343595.

Thakor, A.V. (1996), Capital requirements, monetary policy, and aggregate bank lending: theory and empirical evidence, *Journal of Finance*, Vol. 51, pp. 279-324.

Valla, N., Saes-Escorbia, B. and Tiesset, M. (2006). Bank liquidity and financial stability. *Financial Stability Review*, 9: 89-104.

Vodová, P. (2011a). Liquidity of Czech Commercial Banks and its Determinants. *International Journal of mathematical models and methods in applied sciences*, 5(6): 1060-1067.

Interdisciplinary Journal of Economics and Business Law, 2019, Vol. 8,
Issue 3, pp. 25-51

Changes in labour market policy and the gender wage gap among part-time and full-time workers in South Africa

Colette Muller[1]

Abstract

This paper explores the effect of changes in labour market legislation between 1995 and 1999 and between 2001 and 2006 on the gender wage gap among part-time and full-time workers in South Africa. The research complements the existing literature on gender wage gaps by explicitly differentiating between part-time and full-time workers and by analyzing how the gender wage gap has changed over two distinct periods in the context of legislative changes in South Africa, which aimed, *inter alia*, to reduce gender discrimination. Both the Oaxaca-Blinder and the Juhn, Murphy and Pierce decomposition techniques are used to understand the composition of the gender wage gap, how it differs between part-time and full-time workers, and to identify the factors contributing to the changes in the gender wage gap over time between these groups. The results show that the adjusted gender wage gap is highest among part-time workers but that the total gender wage differential has declined, with the greatest decrease visible for part-time workers. This finding is consistent with gender discrimination declining more among part-time than among full-time workers and is as expected given the amendments to South Africa's labour legislation over the periods interrogated.

Introduction

Investigating and explaining gender wage differentials and gender discrimination is a key area of analysis in the international labour market literature. Extensive research has found that women are typically paid less than men, but that the gender wage gap has narrowed over time (Blau and Kahn 1992; 1997; 2000; 2007; 2017; Hersch 1991; Bernhardt et al. 1995; Brainerd 2000; Manning and Robinson 2004). In South Africa, studies documenting gender differences in pay and the effects of gender-based

[1] School of Accounting, Economics and Finance, University of KwaZulu-Natal, Email: mullerc2@ukzn.ac.za

labour market discrimination are more limited, with much of the research focusing rather on racial wage gaps. Nonetheless, a few studies have documented evidence of gender discrimination in wages – particularly among Whites and Africans (Rospabé 2001; Hinks 2002; Grün 2004; Bhorat and Goga 2013).

Despite extensive international research focusing on gender wage gaps more generally, research that explicitly distinguishes between part-time and full-time workers is limited, and no studies in South Africa make this distinction. A key contribution of this research to the literature is therefore that it analyses the gender wage gap separately for part-time and full-time workers. This distinction is particularly important in the context of legislative reforms in South Africa that occurred in the early years post-apartheid (i.e. post to 1994) and which are likely to have resulted in an improvement in both working conditions and wages, especially in occupations typically associated with women (such as domestic work and other less-skilled jobs). Because these sorts of occupations are overrepresented in female part-time employment in South Africa (Posel and Muller 2008), any observed decline in the gender wage differential should be more pronounced among those working part-time.

Another key contribution of this research to the literature is that the two analysis periods (1995 to 1999 and 2001 to 2006) are chosen intentionally to align with the legislative changes that occurred in South Africa in the late 1990s and early 2000s. By observing changes in the part-time and full-time gender wage gaps over these periods, and by specifically analysing the underlying factors contributing to the changes in these gaps, it is possible to ascertain the impact of South Africa's legislative amendments on the gender wage differential.

The findings show that the adjusted gender wage gap in South Africa is higher among part-time than among full-time workers, but that the total gender gap in wages declined for both part-time and full-time workers across both the time periods analysed. In both periods, the reduction in the gender wage gap is larger among part-time than among full-time workers, suggesting that legislative changes, aimed in part at reducing gender discrimination in South Africa, have had their intended effect.

Theoretical explanations for the gender wage gap

Gender differences in wages may arise, in part, from gender differences in skills and qualifications. If women anticipate shorter and more discontinuous working lives because of household commitments, then they may invest less in formal education and on-the-job training than men do, and even avoid occupations where investments in human capital are required (Mincer and Polacheck 1974). In this case, lower human-capital investments by women[2] will reduce their earnings capabilities relative to those of men. Furthermore, employers who anticipate that women will participate in the labour market intermittently may offer women lower wages (Blau and Kahn 2000).

Labour market discrimination may also account for part of the gender wage gap. According to Oaxaca (1973: 695) “discrimination against females can be said to exist whenever the relative wage of males exceeds the relative wage that would have prevailed if males and females were paid according to the same criteria”. Labour market discrimination can manifest in two forms. Job discrimination occurs when women are segregated into occupations/establishments that pay lower wages. This may be the result of either the initial matching of individuals with jobs, and/or with the process through which promotions are obtained once individuals are employed. Women’s exclusion from ‘male’ jobs can culminate in an excess supply of women in ‘female’ jobs (overcrowding) and lower returns in these occupations. Wage discrimination occurs when, in a given job and within a given establishment, women receive lower wages than men who are equally productive.

Gender differences in skills and occupations, together with labour market discrimination, are typically referred to as the gender specific factors that may account for the wage differential. Wage structure (unrelated to gender) may also influence the magnitude of the gender gap in pay. Blau and Kahn (1997: 2) describe wage structure as “the array of prices set for various labor market skills (measured and unmeasured)

[2] Women’s attainment of human capital may itself be related to discrimination (Peterson and Morgan 1995). This ‘pre-entry’ discrimination occurs outside of the labour market and can result in women’s average productivity being lower than that of men.

and the rents received for employment in particular sectors of the economy”. Human capital theory, for instance, predicts that men have more employment experience than women. Therefore, regardless of gender, the higher the return to experience, the larger the gender wage differential will be. Similarly, if discrimination results in women working in different occupations to men, then the higher the return received by workers (both male and female) employed in predominantly male occupations, the larger the gender pay gap (Blau and Kahn 2000).

Empirical evidence on the gender wage gap

International evidence on the gender pay gap suggests that although the adjusted gap in wages declines when observable differences between men and women are accounted for, a substantial portion of the pay gap (up to forty percent) remains unexplained and is potentially the result of discrimination (Blau and Kahn 2000). However, many studies, particularly for developed countries, have reported a decline in the differential over time (Hersch 1991; Wellington 1993; Blau and Kahn 2000; Blau and Kahn 2017).

Several researchers have investigated the part-time/full-time wage penalty (e.g. Hirsch 2005) but few studies have examined changes in the gender wage differential among part-time and full-time workers. One exception is Preston (2003) who, using data from 1990 and 1998, compared the gender earnings gap among part-time and full-time workers in Australia to determine the effect of decentralised wage bargaining (adopted in 1991) on the pay position of women. Preston’s results show greater convergence in the part-time gender wage gap than in the full-time gender wage gap, a finding attributed largely to changes in productive characteristics, specifically the entry of less qualified and less experienced males into part-time employment. Very recently, Matteazzi et al. (2017) use 2009 data from the European Union Statistics on Income and Living conditions for 11 countries to show that the gender wage gap is typically larger in countries where part-time employment is more prevalent. Their results also suggest that much of the gender wage gap in part-time employment stems from segregation in part-time jobs, while for full-time workers the gender gap in wages remains largely unexplained. Furthermore, these authors highlight the importance of wage-setting institutions in reducing the gender wage gap, particularly among full-time workers.

Very few studies in South Africa have investigated gender wage differentials and none has compared the gender gap in wages among part-time and full-time workers. The available evidence does suggest, however, that having controlled for differences in a range of observable characteristics, women earn less than men (Rospabé 2001; Hinks 2002; Grün 2004; Casale and Posel 2009; Bhorat and Goga 2013). This paper extends existing research on gender wage differentials in South Africa, first by considering evidence of gender wage gaps among part-time and full-time workers estimated at particular points in time, and second, by investigating how the gender wage gap within these groups has evolved over the years.

Legislative changes in South Africa

The early post-apartheid era saw the introduction and implementation of a number of legislative changes likely to affect the gender wage gap among both part-time and full-time workers in South Africa, particularly over the periods of analysis selected here. These include the 1995 Labour Relations Act, which provided guidelines for the resolution of employer/employee disputes and secured the rights of workers to unionise, and the 1997 Basic Conditions of Employment Act (BCEA), which aimed to regulate working hours, overtime pay and basic employment conditions, and which also permits the Minister of Labour to determine minimum wages for employees by sector. Such a determination was made by the Minister of Labour in 2002, when the BCEA was extended to cover domestic services, and a minimum wage for domestic workers was legislated (Department of Labour 2002). The BCEA has also been extended to cover contract cleaning, private security, wholesale and retail trade, agriculture, civil engineering, forestry, hospitality and the taxi sectors (Posel and Muller 2008).

Other legislative additions include the Skills Development Act (SDA) and the Employment Equity Act (EEA) of 1998. The SDA aims to improve the skills of the workforce by raising the level of investment and education in the labour market. While the SDA is not specific to addressing racial and gender disadvantages in the labour market, it is linked to the EEA, which compels employers to implement and extend training measures to individuals from previously disadvantaged groups (including women). The EEA seeks to ensure equal opportunities in the workplace for both men and women by specifically eliminating unfair discrimination in policy and practice and enforcing affirmative action. In addition, the EEA explicitly states that employers must act to reduce disproportionate income differentials. When the implications of the SDA and the EEA are considered jointly, it is possible that previously disadvantaged workers who receive on-the-job or other training as a result of the enforcement of these Acts could anticipate and receive higher wages.

The *collective* implication of these amendments to South Africa's labour legislation is that the gender wage gap in South Africa, as measured in both time periods, should narrow as employers increase compliance and strive to reduce gender discrimination in the workplace. The introduction of protective labour legislation is likely to have improved both working conditions and wages, especially in occupations typically associated with women, such as domestic work and other less-skilled jobs. Because these occupations are overrepresented in female part-time employment in South Africa, the decline in the gender wage differential may be more pronounced among those working part-time. In particular, the introduction of minimum wages for domestic workers in 2002 may have an important impact on the gender wage gap among South Africa's part-time workers over the 2001 to 2006 period.

Data sources and descriptive statistics

This study uses data from the 1995 and 1999 October Household Surveys (OHSs) and the 2001 and 2006 September Labour Force Surveys (LFSs) to investigate the gender wage gap in particular years, as well as to examine changes in the gender wage differential over time. These datasets provide information on the state of the country's labour market prior to the legislative amendments discussed earlier (in the case of the 1995 OHS and the 2001 LFS) as well as following these changes (in the case of the 1999 OHS and 2006 LFS). The change in the gender wage gap is examined from 1995 to 1999 and from 2001 to 2006.[3]

A general concern about comparability, applicable to all the surveys utilised, involves differences in how information is collected over time. Over time, South Africa's official data collection agency, Statistics South Africa (StatsSA) improved the design of the survey instruments, with a view to capturing more information on irregular work. Although these changes may be more likely to influence measures of self-employment (and particularly survivalist and subsistence self-employment), measures of wage employment may also be affected. In particular, because the LFS questionnaires are more comprehensive in defining what constitutes employment, the LFSs are more likely than the OHSs to capture information on individuals (especially women) involved in work that is marginal and poorly remunerated. To help reduce any bias that may result from analysing the change in the gender wage gap over the period that coincides with the introduction of the LFSs the econometric analysis is separated into two parts: a 1995 and 1999 comparison, and a 2001 and 2006 comparison. An additional comparability concern results from using the 1995 OHS. The 1995 OHS is the only survey used which fails to distinguish between actual and usual hours worked. *Actual* working hours are therefore used to calculate hourly earnings and to distinguish part-time from full-time workers in 1995 and in 1999, while *usual* working hours are used in 2001 and in 2006.[4] The 1995 OHS also does not capture information on employees' receipt of benefits (such as medical aid and pension fund contributions) and firm size and it does not permit a distinction between wage employees in the formal and informal sectors.

---

[3] Although more recent data sources are available (e.g. StatsSA's Quarterly Labour Force Surveys) these data are not well suited to this analysis. Not only is the data captured less comprehensive, but there have not been any recent labour market legislation amendments with the potential to influence the gender wage gap, particularly as it pertains to part-time and full-time workers.

[4] There is no significant difference (using a 95 percent confidence interval) between the mean actual and usual hours worked by either men or women wage employees in the 1999 OHS or in the LFSs.

The 1995 and 1999 econometric estimations therefore exclude variables controlling for these characteristics.

Tables 1 and 2 present descriptive statistics of average wages and working hours among the part-time and full-time wage employed by gender. The samples used comprise persons aged older than 15 years with wage employment, who report non-zero working hours of less than 113 hours per week and for whom earnings information is available. The statistics are weighted using population weights, and standard errors are in parentheses.

For all years the statistics show that among the full-time and the part-time employed, women typically earn less than men on average (in terms of both hourly and monthly wages). The average female-male wage ratio has increased over time among those working full-time, indicative of a narrowing in the (mean) gender gap in hourly wages. This trend is somewhat noisier among part-time workers, rising only slightly from 1995 to 1999, falling from 1999 to 2001, and then increasing substantially from 2001 to 2006.

Table 1. Average wages (2000 prices in Rands) and working hours among the part-time wage employed by gender, 1995-2006

	1995		1999		2001		2006	
	M	F	M	F	M	F	M	F
Number of observations	874	1 415	852	1 336	590	1 185	572	1 258
Monthly wage	2276 (115)	1809 (65)	1799 (124)	1581 (190)	1654 (-180)	1136* (70)	1381 (-104)	1457 (-247)
Hours worked	22.57 (0.33)	22.27 (0.24)	18.16 (0.34)	19.99* (0.26)	21.32 (0.43)	21.59 (0.33)	22.00 (-0.55)	22.07 (-0.32)
Hourly wages	28.21 (2.11)	19.84 (1.07)	28.66 (1.92)	20.30* (1.76)	17.48 (1.49)	11.56* (0.64)	16.10 (-1.31)	15.33 (-2.60)
Hourly wage ratio (%) (F/M)	70.32		70.83		66.13		95.21	

Notes: * indicates that means for men and women are significantly different at the 95% level

Table 2. Average wages (2000 prices in Rands) and working hours among the full-time wage employed by gender, 1995-2006

	1995		1999		2001		2006	
	M	F	M	F	M	F	M	F
Number of observations	16 268	10 457	10 405	7 199	10 898	7 755	10 782	7 620
Monthly wage	3205 (37)	2291* (29)	3355 (-112)	2463* (91)	2958 (-73)	2313* (-56)	3264 (-80)	2614* (-72)
Hours worked	46.26 (0.08)	43.96* (0.09)	50.00 (0.13)	47.44* (0.14)	49.72 (-0.14)	47.24* (-0.15)	48.12 (-0.15)	45.58* (-0.17)
Hourly wages	16.23 (0.19)	11.71* (0.15)	16.60 (0.60)	12.80* (0.47)	14.58 (-0.35)	12.12* (-0.28)	16.54 (-0.41)	13.90* (-0.38)
Hourly wage ratio (%) (F/M)	72.15		77.10		83.12		84.03	

Notes: * indicates that means for men and women are significantly different at the 95% level

A comparison of both the part-time and the full-time female-male wage ratios from 2001 to 2006 is suggestive of a larger decline in the gender wage gap among the part-time employed with the increase in the ratio among those working part-time exceeding that among those working full-time.

Estimation technique

Ordinary Least Squares (OLS) is used to estimate separate human capital regressions for men (*M*) and women (*F*) (the process described below is repeated for the respective part-time and full-time samples). For individual i, the following equations are estimated:

$$\ln(W_i^M) = \alpha^M + \beta X_i^M + \varepsilon_i^M \quad (1)$$

$$\ln(W_i^F) = \alpha^F + \beta X_i^F + \varepsilon_i^F \quad (2)$$

W_i represents the real hourly wages of individual i, X_i is a vector of individual, job and industry parameters, and ε_i is the error term.

The Oaxaca-Blinder (OB) decomposition technique is used to identify what portion of any wage gap, estimated at each cross section, is due to differences in observable characteristics, and what portion is the result of differences in the returns to these characteristics.

$$\ln(W^M) - \ln(W^F) = \sum_i \hat{\beta}^M (\bar{X}_i^M - \bar{X}_i^F) + \{(\hat{\alpha}^M - \hat{\alpha}^F) + \sum_i \bar{X}_i^F (\hat{\beta}^M - \hat{\beta}^F)\} \quad (3)$$

The first term represents the portion of the wage differential attributable to measurable factors, i.e. to gender differences in endowments. The second term is the 'unexplained' part of the differential, capturing the effects of differences in the intercepts of the male and female earnings equations and in the estimated coefficients.

Of particular interest is whether the magnitude of the gender gap in wages among part-time and full-time workers has risen or fallen in South Africa over these periods where significant legislative amendments occurred, and what factors may account for any changes observed. When attempting to determine how the gender wage gap, net of differences in observable characteristics, has changed over time it is not possible to just compare the magnitudes of the adjusted (residual) differential estimated at each cross-section. This is because the magnitude of the adjusted gender gap in wages depends not only on gender differences in returns, which can change over time, but also on the underlying characteristics of the workers, $\bar{X}_i^F$, which too can change. For example, a decline in the magnitude of the unexplained gap over time may be the result of women's returns improving relative to men's or it could be the result of women's observable characteristics worsening over the years. Interpreting any change in the magnitude of the adjusted wage gap as evidence of a decline (rise) in the portion of the gender wage gap which remains having controlled for differences in observable characteristics would therefore be misleading as part of what may change, $\bar{X}_i^F$, can in fact be accounted for.

To decompose the *change* in the gender wage differential from one year to the next, a method developed by Juhn, Murphy and Pierce (1991; 1993)[5] (JMP) and subsequently implemented by (amongst others) Blau and Kahn (1997) and Brainerd (2004) is used. The JMP method also provides a way of illustrating how unobservable differences between men and women affect the gender wage gap.

The male wage equation in period *t* is written as:

$$W_{Mt} = X_{Mt}\beta_t + \sigma_t\theta_{Mt} \tag{4}$$

where the dependent variable W_{Mt} is the natural logarithm of real hourly wages, X_{Mt} is a vector of explanatory variables and β_t is a vector of coefficients. The standard deviation of the residual from the male wage equation is represented by σ_t, and θ_{Mt} is the standardised residual of the male wage regression, with a mean of 0 and a variance of 1. The residual therefore consists of two components: θ_{Mt} reflects the percentile that a particular individual occupies in the residual distribution and σ_t reflects the spread of the residual distribution.

This distinction in the components of the residual is exploited by JMP in their decomposition technique. Following Brainerd (2004: 153), the gender wage gap in *t* is written as:

$$D_t \equiv W_{Mt} - W_{Ft} = (X_{Mt} - X_{Ft})\beta_t + (\theta_{Mt} - \theta_{Ft})\sigma_t \tag{5}$$

Note that $\theta_{Ft} = (W_{Ft} - X_{Ft}\beta_t)/\sigma_t$, which reflects the wage that women would earn if their characteristics were rewarded at the same rate as those of men (deflated by the male standardised residual).

[5] Smith and Welch (1989) propose decomposing changes in wage differentials using a double application of the Oaxaca-Blinder decomposition. Their approach yields results identical to Juhn et al. (1991; 1993) bar for their decomposition of the change in the residual wage gap, which is instead decomposed into a portion attributable to changes in observable characteristics, and a portion due to changes in returns. See also Heckman et al. (2000).

The change in the wage gap from *t* to *t'* is written as:

$$D_{t'} - D_t = [(X_{Mt'} - X_{Mt}) - (X_{Ft'} - X_{Ft})]\beta_{t'}$$
$$+(X_{Mt} - X_{Ft})(\beta_{t'} - \beta_t)$$
$$+[(\theta_{Mt'} - \theta_{Ft'}) - (\theta_{Mt} - \theta_{Ft})]\sigma_t$$
$$+(\theta_{Mt} - \theta_{Ft})(\sigma_{t'} - \sigma_t) \qquad (6)$$

The first term, the "Observed X's Effect", reflects changes in the wage gap that result from changes in gender differences in observed characteristics from *t* to *t'*. The second term, the "Observed Prices Effect", shows the contribution of changes in the way observed characteristics of men are rewarded in the labour market, holding constant measurable differences between men and women. As Blau and Kahn (1997: 7) note, the gender wage gap would rise if, for instance, men's return to experience increased and women have less experience than men. The third term, the "Gap Effect", represents the contribution of changes in women's position in the male residual distribution. If women's unobserved labour market skills improve relative to men's, or if labour market discrimination against women declines, they will move up this distribution. Finally, the fourth term, the "Unobserved Prices Effect", measures the change in the gender wage gap resulting from the widening (or narrowing) distribution of male wage residuals while holding constant the gender gap in unmeasured skills.

Aggregating the Observed X's Effect and the Gap Effect reveals the full-effect of gender-specific differences in observable characteristics and gender differences in wage rankings at a particular level of observed characteristics. The Observed and Unobserved Prices Effects together reflect changes in wage structure, i.e. changing returns to both observed and unobserved characteristics.

The interpretation of both the Observed and Unobserved Prices Effects can be complicated by the presence of labour market discrimination. Over time, if women are crowded into certain sectors, and relative wages in these sectors are depressed (even for men), then the

Observed Prices Effect can reflect both job discrimination as well as changes in men's rewards for productive characteristics and rents. This concern is addressed here by estimating the model both with and without controls for occupation. Furthermore, in the presence of discrimination, the Unobserved Prices Effect "in part reflects the interaction between year 0's level of discrimination (which pushes women down the distribution of male wage residuals) and the change in the overall level of inequality, which determines how large the penalty is for that lower position in the distribution" (Blau and Kahn 1997: 8).

Potential concerns

When estimating and decomposing an earnings function for any group parameter estimates based solely on a sample of the employed may be biased if the sub-sample is not representative of the entire sample. In this study, such a problem may occur if women (men) working part-time differ not only from those women (men) working full-time, but also from those women (men) who are unemployed or who are economically inactive. As in studies that investigate the part-time/full-time wage differential, in the gender-wage gap literature the Heckman two-stage procedure is often used to address the sample selection bias problem (Hinks 2002; Grün 2004). Obtaining exclusion restrictions that are not also correlated with earnings can be problematic and in the data used here it was impossible to find such instruments (given the data collected by StatsSA, this problem is endemic in South African studies of this sort). The sample selection problem is further exacerbated by the need to also account for the possibility that part-time and full-time workers differ in terms of both measurable and unmeasured characteristics. Although data from the LFS Panel could be used to address the issue of sample selection between part-time and full-time workers using a fixed-effects estimation (Posel and Muller 2008), because gender is unchanged over time, the effect of gender on any change in earnings would be eliminated with the within-transformation of the data.

Issues of selectivity may also affect the measurement of the change in the gender wage gap over time. Over the two time periods used here women’s labour force participation in South Africa increased rapidly, with research suggesting that women have been pushed, rather than pulled into the labour market (Casale and Posel 2002; Casale 2003). Consequent changes in the unmeasured selectivity of female labour force participants over the years may bias the measurement of the change in the gender wage gap. However, men’s labour force participation in South Africa has been significantly more stable than women’s labour force participation over the analysis periods used and parameter estimates from the male wage equation should be less susceptible to bias introduced by changes in men’s unobservable characteristics over time. The male earnings function, rather than the female, or a pooled, wage equation is therefore used as the reference category for the decompositions.

Another potential concern is that the male and female earnings equations are estimated and decomposed without restricting the comparison to only those individuals whose characteristics are comparable. This problem is typically referred to as a failure to recognise “*gender differences in the supports*” (Ñopo 2008), and may result in an underestimation or an overestimation of the portion of the gap attributable to differences in the returns to individual characteristics.[6] One possible solution can be found in the program evaluation literature where gender is considered as a treatment and matching is used to select sub-samples of men and women with identical observable characteristics (see, for example, Ñopo 2008). While such a non-parametric procedure may assist in solving the ‘gender differences in supports’ problem and is also useful for exploring the distribution of unexplained differences in wages, it is limited in its ability to control for the many explanatory factors that may influence earnings and earnings differences and is therefore not utilised here.

[6] An overestimation (underestimation) of the unexplained wage gap would occur if matched males (i.e. men for whom it is possible to find women with comparable characteristics) typically have wages which are, on average, lower (higher) than those for unmatched males. See Ñopo 2008 for further details.

Results

In the interests of brevity, results from the forty separate OLS estimations used to perform the decompositions are omitted here, but are available from the author on request. Tables 3 and 4 show the decomposition results from 1995 to 1999 for the separate samples of part-time and full-time wage workers, while Tables 5 and 6 provide decomposition results for the part-time and full-time samples from 2001 to 2006. Column I in each table presents results where independent controls for age, job duration, race, education, marital status and location were included, while the in column II, variables controlling for occupation, industry and firm size are added. Column III in Tables 5 and 6 includes further controls for conditions of employment and also distinguishes between employment in the formal and informal sectors. In all the tables, the estimating samples have been restricted to persons aged older than 15 years, who reported non-zero working hours of less than 113 hours per week, and for whom earnings information is not missing. The estimates are calculated using population weights, and estimates as a percentage of the unadjusted/total differential (or as a percentage of the change in the unadjusted/total differential), are in parentheses.

Across the years, and among both part-time and full-time workers, the total gender gap in wages is positive, implying a wage differential in favour of men. In both periods the gender gap in wages persists among the part-time and the full-time cohorts when observable differences between workers are accounted for (although in some specifications the adjusted wage differential is lower than the unadjusted wage gap). In all years, the inclusion of controls for occupation and industry in specification II decreases the magnitude of the residual (unexplained) portion of the wage gap for both part-time and full-time workers, indicating that gender differences in occupational access account for a substantial portion of the gender wage gap.

Table 3. Decomposition of the gender wage differential, 1995 to 1999 (Part-time wage employed)

	I				II			
	1995		1999		1995		1999	
	M	F	M	F	M	F	M	F
Number of observations	799	1 341	811	1 268	775	1 322	765	1 216
R-squared	0.30	0.51	0.34	0.44	0.48	0.79	0.45	0.52
Total (unadjusted differential)	0.841		0.430		0.844		0.402	
Quantity effect	-0.104 (-12)		-0.116 (-27)		0.722 (86)		0.282 (70)	
Residual gap	0.945 (112)		0.546 (127)		0.121 (14)		0.120 (30)	
Change in total differential	-0.411				-0.441			
Change in quantity effect	-0.011 (3)				-0.440 (99)			
Change in residual gap	-0.399 (97)				-0.001 (1)			
Observed X's effect	-0.016 (4)				0.006 (-1)			
Observed prices	0.004 (-1)				-0.447 (101)			
Gap effect	-0.457 (110)				-0.007 (2)			
Unobserved prices effect	0.057 (-13)				0.006 (-1)			

Table 4. Decomposition of the gender wage differential, 1995 to 1999 (Full-time wage employed)

	I				II			
	1995		1999		1995		1999	
	M	F	M	F	M	F	M	F
Number of observations	15 479	9 965	9 858	6 852	15 152	9 822	9 209	6 470
R-squared	0.59	0.59	0.50	0.56	0.72	0.82	0.60	0.66
Total (unadjusted differential)	0.380		0.244		0.380		0.239	
Quantity effect	-0.056 (-15)		-0.050 (-21)		0.214 (56)		0.039 (16)	
Residual gap	0.437 (115)		0.295 (121)		0.166 (44)		0.200 (84)	
Change in total differential	-0.135				-0.141			
Change in quantity effect	0.006 (-4)				-0.174 (123)			
Change in residual gap	-0.141 (104)				0.033 (-23)			
Observed X's effect	-0.017 (13)				0.067 (-48)			
Observed prices	0.023 (-17)				-0.241 (170)			
Gap effect	-0.175 (129)				0.000 (0)			
Unobserved prices effect	0.033 (-24)				0.033 (-23)			

Table 5. Decomposition of the gender wage differential, 2001 to 2006 (Part-time employed)

	I				II				III			
	2001		2006		2001		2006		2001		2006	
	M	F	M	F	M	F	M	F	M	F	M	F
Number of observations	541	1098	550	1206	529	1081	548	1203	483	991	539	1186
R-squared	0.40	0.58	0.38	0.60	0.54	0.64	0.54	0.67	0.62	0.67	0.58	0.68
Total (unadjusted differential)	0.367		0.234		0.345		0.235		0.347		0.222	
Quantity effect	-0.037 (-10)		-0.063 (-26)		0.020 (6)		0.149 (63)		-0.001 (0)		0.188 (85)	
Residual gap	0.405 (110)		0.297 (126)		0.325 (94)		0.085 (36)		0.349 (100)		0.034 (15)	
Change in total differential	-0.133				-0.110				-0.124			
Change in quantity effect	-0.025 (20)				0.129 (-117)				0.189 (-152)			
Change in residual gap	-0.107 (80)				-0.239 (217)				-0.314 (253)			
Observed X's effect	-0.102 (77)				-0.149 (135)				-0.219 (176)			
Observed prices	0.076 (-57)				0.278 (-253)				0.409 (-329)			
Gap effect	-0.062 (47)				-0.215 (195)				-0.306 (246)			
Unobserved prices	-0.044 (33)				-0.023 (21)				-0.007 (6)			

Table 6. Decomposition of the gender wage differential, 2001 to 2006 (Full-time employed)

	I				II				III			
	2001		2006		2001		2006		2001		2006	
	M	F	M	F	M	F	M	F	M	F	M	F
Number of observations	10623	7523	10664	7520	10220	7332	10555	7450	9311	6739	10322	7303
R-squared	0.56	0.64	0.53	0.59	0.68	0.74	0.62	0.69	0.72	0.77	0.67	0.73
Total (unadjusted differential)	0.194		0.162		0.200		0.160		0.202		0.162	
Quantity effect	-0.105 (-54)		-0.085 (-52)		0.008 (4)		-0.020 (-13)		0.020 (10)		0.006 (4)	
Residual gap	0.299 (154)		0.247 (152)		0.192 (96)		0.181 (113)		0.182 (90)		0.156 (96)	
Change in total differential	-0.032				-0.040				-0.039			
Change in quantity effect	0.019 (-59)				-0.029 (73)				-0.013 (33)			
Change in residual gap	-0.051 (159)				-0.010 (25)				-0.025 (64)			
Observed X's effect	0.012 (-38)				-0.019 (48)				-0.016 (41)			
Observed prices	0.007 (-22)				-0.010 (25)				0.002 (-5)			
Gap effect	-0.049 (153)				-0.018 (45)				-0.030 (76)			
Unobserved prices	-0.002 (6)				0.007 (-18)				0.004 (-10)			

In 2001 and 2006, controlling for differences also in conditions of work in specification III further reduces the magnitude of the residual gender wage gap among both part-time and full-time workers (an exception is in 2001, where the adjusted wage gap increases slightly from specification II to specification III for part-time workers).

The cross-sectional decomposition results also show that, in all years and in all specifications, the magnitude of the unadjusted gender gap in wages is greater among part-time than among full-time workers. These results may seem surprising given evidence of a premium to female part-time wage employment in South Africa (Posel and Muller 2008). However, the premium to men's part-time work is even larger than that for women. When the wage estimations control for gender differences in observable characteristics (including occupation and industry in 1999, as well as conditions of work in 2006), the residual gap among the full-time employed in these years exceeds that estimated among the part-time employed. This is suggestive of a greater reduction in wage-based gender discrimination among part-time than among full-time workers from 1995 to 1999, and from 2001 to 2006. To explore these findings further, the JMP technique is used to decompose the change in the gender wage gap over these years.

The decomposition results for 1995 to 1999 point to a decline in the gender wage gap over the period of between 0.411 and 0.441 log points for part-time workers and between 0.135 and 0.141 log points for full-time workers. This suggests that the decline in the gender wage gap over these years was greater among part-time than among full-time workers. Similar results are found for 2001 to 2006: among part-time wage workers the gender wage gap decreased by approximately 0.12 log points (roughly 35 percent, on average), and far exceeded the magnitude of the decline in the wage gap among those working full-time, which ranged between 0.032 and 0.040 log points (or 16 and twenty percent).

The JMP decomposition makes it possible to identify the main sources of the narrowing of the gender wage gap over each period and within each group. For both part-time and full-time workers over the 1995 to 1999 period, the results of the JMP decomposition for the first specification (i.e. when controls for occupation and industry are omitted) suggest that the primary source of the decline in the gender wage gap is the result of gender specific factors (i.e. the Gap Effect). The Gap Effect measures the contribution of changes in discrimination to the change in any wage gap and contributed more than 100 percent to the decline in the gender wage differential over the period. However, when the wage estimations reflect the gender wage gap calculated for women involved in the same occupations and industries as men, the importance of gender specific factors and the Gap Effect is diminishes. In the full specification it is an improving wage structure (the Observed Prices Effect), which is the primary source of the decline in the gender wage gap for both part-time and full-time workers, reducing the gender wage gap by 0.447 log points (101 percent) for part-time workers and by 0.241 log points (170 percent) for full-time workers. The Observed Prices Effect suggests that changes in the prices of skills and/or rents for men have worked to decrease the gender wage gap over the period, ceteris paribus. This result is consistent with the introduction of protective labour legislation over the period (and with the implementation of the Employment Equity Act, in particular), which may have served to decrease the demand for male workers, thereby lowering the returns received by men for their productive characteristics[7]. This result may also be explained by increased occupational crowding among women, which would have lowered the returns not just for women in the occupations into which they were crowded, but also for the men working in these occupations.[8]

For the 2001 to 2006 period, the results of the JMP decomposition for part-time workers shows that between 77 (specification I) and 176 percent (specification III) of the reduction in the gender wage gap among part-time workers can be attributed to an improvement in women's observable characteristics relative to those of men (the Observed

---

[7] In terms of South Africa's affirmative action policy, women are considered a previously disadvantaged group. *Ceteris paribus*, if firms hire women rather than men in order to achieve gender equity targets mandated by the EEA then the price of male labour may decline.

[8] Increased occupational crowding among women is also plausible given the introduction of the EEA.

Xs Effect). In all specifications, the negative sign on the Gap Effect shows that women's position in the residual male wage distribution improved over the period, indicative of a decline in discrimination against women in the labour market and/or improvements in women's levels of unobserved skills relative to men's. Taken together, the Observed X's and Gap Effect reinforce each other and reveal an overall improvement in gender-specific factors for women working part-time, accounting for between 123 (specification I) and 423 percent (specification III) of the change in the wage gap over time.

While these improvements in gender specific factors worked to reduce the gender gap in wages among those working part-time, a deteriorating wage structure worked to increase the gender wage gap. This is indicated, in part, by the positive signs observed on the Observed Prices effects. In contrast to the 1995 to 1999 period, the Observed Prices effect shows that the prices of skills or rents have changed from 2001 to 2006 so as to *increase* the male-female wage gap among part-time workers in South Africa. This finding may also reflect increased occupational crowding among women working part-time. As a result, despite women's position in the part-time male residual wage distribution typically improving from 2001 to 2006 (as shown by the negative sign on the Unobserved prices effect in all three specifications), the overall widening of the part-time wage distribution over the period offsets the gains made in gender-specific factors by between 0.03 and 0.4 log points.

Among full-time employees the results of the decomposition of the change in the gender wage gap from 2001 to 2006 over time are similar to those among part-time workers. Gender specific factors are shown to account for between about 93 and 117 percent of the reduction in the gender wage differential among full-time workers, with a worsening wage structure offsetting some of these gains. Overall, a far greater improvement in gender specific factors is found among those working part-time than among those working full-time. The Gap Effect, which illustrates changes in discrimination and/or unobservable characteristics, contributes more to the reduction of the gender wage differential among part-time workers (more than 190 percent in specifications II and III) than among the full-time wage employed (less than 80 percent in specifications II and III.) This is consistent with improvements in labour legislation impacting particularly upon part-time workers, and where a reduction in discrimination may be greater than among those working full-time. This result may also reflect the effects of potentially larger improvements in the unobservable characteristics of women working part-time as compared to those of women working full-time over the period.

Concluding comments

Few studies of the gender wage gap in South Africa have investigated changes in gender wage differentials over time, and none have distinguished between part-time and full-time employment. The evidence here suggests that the gender wage gap in South Africa is considerably higher among part-time wage employees than among those working full-time.

To investigate the change in the gender wage gap in post-apartheid South Africa the analysis distinguished between two periods: 1995 to 1999 and 2001 to 2006. These two periods are well suited to analysing changes in the gender wage gap as they provide information on the South African labour market both prior to and following the implementation of various legislative changes targeted specifically at improving women's access to jobs and pay.

The results from 1995 to 1999 and from 2001 to 2006 show that the gender gap in wages is typically higher among part-time than among full-time workers. Over time, however, the gender gap in wages has narrowed. Moreover, the decline in the gender wage differential over both periods is more pronounced among part-time than among full-time workers.

Identifying the primary source of the decline in the gender wage differential over time is complicated by the inability to account for various sources of potential selectivity. Nevertheless, from 1995 to 1999, the results from the full specification of the wage equations suggest an improvement in the structure of wages that served to decrease the gender wage gap among both part-time and full-time workers. This points to a possible reduction in the demand for male workers following the introduction of legislation aimed at improving women's access to employment, working conditions and pay, which may have worked to decrease the returns to men's characteristics over the period.

In contrast, the results from 2001 to 2006 suggest that the decline in the gender wage differential is the consequence mainly of improvements in gender-specific factors. The magnitude of the Gap effect, which may reflect changes in discrimination and/or unobservable characteristics, is larger among those employed in part-time jobs. Although there is descriptive evidence suggesting that certain employment benefits (like medical aid and pension fund contributions) have been lost by workers over the years while others (like contributions to the unemployment insurance fund) have been gained, this finding is consistent with employers increasing compliance with the legislative changes implemented over the time periods analysed. The findings of this paper are therefore consistent with policy changes having real labour market effects.

Although this paper makes an important contribution to the understanding of gender wage differentials and how labour market legislation can assist in narrowing the gender wage gap in South Africa, scope for additional analyses remains. In particular, a deeper understanding of how legislative changes may have affected men's' and women's' wage distributions and the gender wage at points along these wage distributions is still required. In addition, investigating how the gender gap in wages differs between public and private sector workers is also important, not least to understand in which of these sectors any decline in gender discrimination stemming from amendments in labour market legislation has been the greatest.

References

Bernhardt, A., Morris, M. and Handcock, M.S. (1995). Women's gains or men's losses? A closer look at the shrinking gender gap in earnings. *The American Journal of Sociology*, 101(2): 302-328.

Bhorat, H. and Goga, S. (2013). The gender wage gap in Post-Apartheid South Africa: A re-examination. *Journal of African Economies*, 22(5): 827-848.

Brainerd, E. (2000). Women in transition: Changes in gender wage differentials in Eastern Europe and the former Soviet Union. *Industrial and Labor Relations Review,* 54(1): 138-162.

Blau, F.D. and Kahn, L.M. (2017). The Gender Wage Gap. Extents, Trends and Explanations. *Journal of Economic Literature*, 55(3): 789-865.

Blau, F.D. and Kahn, L.M. (2007). The gender pay gap. *Economists' Voice*, June, 1-6.

Blau, F.D. and Kahn, L.M. (2000). Gender differences in pay. *The Journal of Economic Perspectives*, 14(4): 75-99.

Blau, F.D. and Kahn, L.M. (1997). Swimming upstream: Trends in the gender wage differential in the 1980s. *Journal of Labor Economics*, 15(1): 1-42.

Blau, F.D. and Kahn, L.M. (1992). The gender earnings gap: Learning from international comparisons. *The American Economic Review*, 82(2): 533-538.

Casale, D. (2003). "The rise in female labour force participation in South Africa: An analysis of household survey data, 1995-2001." Unpublished Ph.D thesis, Economics, University of Natal, Durban.

Casale, D. and Posel, D. (2009). Unions and the gender wage gap in South Africa. *Economic Research South Africa Working Paper*, No. 113.

Casale, D. and Posel, D. (2002). The continued feminisation of the labour force in South Africa: An analysis of recent data and trends. *South African Journal of Economics*, 70(1): 156-184.

Erichsen, G. and Wakeford, J. (2001). Racial wage discrimination in SA before and after the first democratic election. *Development Policy Research Unit Working Paper*, No. 01/49.

Grün, C. (2004). Direct and indirect gender discrimination in the South African labour market. *International Journal of Manpower*, 25(3/4): 321-342.

Heckman, J.T., Lyons, T.M. and Todd, P.E. Understanding black-white wage differentials, 1960-1990. *The American Economic Review*, 90(2): 344-349.

Hersch, J. (1991). Male-female differences in hourly wages: The role of human capital, working conditions and housework. *Industrial and Labor Relations Review*, 44(4): 746-759.

Hirsch, B.T. (2005). Why do part-time workers earn less? The role of worker and job skills. *ILR Review*, 58(4): 525-551.

Hinks, T. (2002). Gender wage differentials and discrimination in the New South Africa. *Applied Economics*, 34(16): 2043-2052.

Juhn, C., Murphy, K. and Pierce, B. (1993). Wage inequality and the rise in returns to skill. *The Journal of Political Economy*, 101(3): 410-442.

Juhn, C., Murphy, K. and Pierce, B. (1991). "Accounting for the slowdown in black-white wage convergence." In M. Kosters, *Workers and Their Wages*. Washington D.C.: AEI Press, 107-143.

Manning, A. and Robinson, H. (2004). Something in the way she moves: A fresh look at an old gender gap. *Oxford Economic Papers* 56, 169-188.

Matteazzi, E., Pailhé, A. and Solaz, A. (2017). Part-time employment, the gender wage gap and the role of wage-setting institutions: Evidence from 11 European countries. *European Journal of Industrial Relations*.

Mincer, J. and Polacheck, S. (1974). Family investments in human capital: Earnings of women. *The Journal of Political Economy*, 82(2), Part 2, S76-S108.

Muller, C. and Posel, D. (2004). Concerns with measuring informal sector employment: An analysis of national household surveys in South Africa, 1993-2001. *Journal of Studies in Economics and Econometrics*, 28(1): 1-21.

Ñopo, H. (2008). Matching as a tool to decompose wage gaps. *The Review of Economics and Statistics*, 90(2): 290-299.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 14(3): 693-709.

Peterson, T. and Morgan, L.A. (1995). Separate and unequal: Occupation-establishment sex segregation and the gender wage gap. *The American Journal of Sociology*, 101(2): 329-365.

Posel, D. and Muller, C. (2008). Is there evidence of a wage penalty to female part-time employment in South Africa? *South African Journal of Economics*, 76(3): 466-479.

Preston, A. (2003). Gender earnings and part-time pay in Australia, 1990-1998. *British Journal of Industrial Relations*, 41(3): 417-433.

Rospabé, S. (2002). How did labour market racial discrimination evolve after the end of apartheid? An analysis of the evolution of employment, occupational and wage discrimination in South Africa between 1993 and 1999. *The South African Journal of Economics*, 70(1): 185-217.

Rospabé, S. (2001). "An empirical evaluation of gender discrimination in employment, occupation attainment and wage in South Africa in the late 1990s." Unpublished mimeo, University of Cape Town.

Smith, J.P. and Welch, F.R. (1989). Black economic progress after Myrdal. *Journal of Economic Literature*, 27: 519-564.

Wellington, A.J. (1993). Changes in the male/female wage gap, 1976-85. *The Journal of Human Resources*, 28(2): 383-411.

Interdisciplinary Journal of Economics and Business Law, 2019, Vol. 8, Issue 3, pp. 52-78

Rothschild, Warburg, and the Mitsui family: A case of ownership and organisational reforms under the modern legal constraints in early 20th century Japan

Shunsuke Nakaoka[1]

Abstract

We aim to explore how one notable Japanese family business sought to model their corporate governance on European family businesses as they were urgently looking for suitable examples to transform their traditional system of family ownership in this period of radical change to business and economic conditions. Specifically, this analysis is focused on the case of the Mitsui family, who successfully established one of the most influential business conglomerates, or *zaibatsu*, in modern Japanese history. Firstly, we assess the general background and framework of reforms from the late 19th to the early 20th century, which culminated in the foundation of the Mitsui *Gōmei Gaisha* (Mitsui Holdings). From the European businesses we focus on the House of Rothschild since reliable sources demonstrate that Mitsui considered the Rothschild family as an ideal model for Mitsui's ownership and business reforms. Secondly, we support our arguments empirically through the documents of business trips made by Mitsui executives and Mitsui family members, which sought out examples of family ownership and corporate systems in Europe and the US. This includes the contents of meetings with several foreign business advisors and their contributions to this quest, in particularly those of Max Warburg.

Introduction

Since the late 19th century, as a result of the abolition of the long-term restrictions on foreign trade by the previous Tokugawa government, Japan confronted serious challenges in the process of reconnecting with the world economy through trade and business. This was followed by a series of political crises which contributed to the establishment of a new social

[1] Faculty of Political Science and Economics, Kokushikan University, Tokyo, Japan. Email: nakaoka@kokushikan.ac.jp

order after the Meiji Restoration period (Jansen 2000). While opening to trade with other countries created opportunities for Japanese businessmen and merchants, the adaptation of an European institutional framework was necessary for their survival under this new circumstance. This was largely due to the necessity for the transformation of their activities – from organisational structure to day-to-day business – under globally based new economic and social conditions that were dominated by western countries. The changes to Japanese business in the late 19th century were further facilitated by governmental policies, since the government wanted to gain autonomy from the European and US tariff regime which was a result of an unequal treaty that Japan had signed with western countries at the end of the Tokugawa era. This led to the introduction of a new legal system based on the European model in the late 19th century; when Japan codified its laws, as this was one of the conditions for foreign countries to abolish their unequal trade treaties with Japan (Akita 1967). This new legal framework certainly proved useful to Japanese family businesses in the modification of their ownership and managerial systems.

It was during this period of historical transformation, the first major Japanese family business (*zaibatsu*) focused on European family businesses to identify a business model, as they were urgently seeking suitable examples to assist them in moving away from their own traditional system of family ownership. The Mitsui family had successfully transformed its traditional mercantile business into one of the most influential business conglomerates, or *zaibatsu*, in modern Japanese history (Taplin 1995, 2016). In the late 19th century, this family was confronted with serious but inevitable challenges that necessitated reforms in the form of the reorganisation and redevelopment of their businesses. This occurred while social and political connections with the new Meiji government simultaneously led to rapid expansion and diversification of their business throughout this period (Mitsui Bunko 1980). The family ownership structure comprised a combination of traditional business and household systems that facilitated the ownership of business assets commonly among the members of the clan (*Sōyusei*). The aforementioned reforms included reshaping this system of family ownership to fit into modern corporate and legal systems. As a result of the success of their transformation, the Mitsui family business was then considered a model for many Japanese mercantile families that were in similar situations pertaining to family ownership. Several other such families began to utilise this model for the continuity and survival of their businesses. Therefore, it can be said that the Mitsui case illustrates survival strategies for family businesses at a time when Japan was confronted with both internal and external pressures for the internationalisation of commerce. Their progress and the struggle they faced in implementing these reforms also indicate that seeking a suitable model within European family business cases certainly did mitigate the burden of business reorganisation within a vacuum.

Extensive studies both in Japanese and English have analysed the Mitsui business model and reforms during this period.[2] However, one of the purposes of this article is to add another perspective to the existing contributions by reconsidering the Mitsui household reforms from the perspective of one of their attempts, that is, the exploration of European family business systems, which was undoubtedly connected to their business reorganisation needs. First, after presenting a concise history of the Mitsui family business, which originated in the period of the Tokugawa Government, the general background and the framework of reforms between the late 19th and the early 20th centuries which culminated in the foundation of the Mitsui *Gōmei Gaisha* (Mitsui Holdings) will be examined in exploring the primary sources and documents related to the household reforms from the Mitsui Archives (*Mitsui Bunko*). Special focus will be on a particular European business family, namely the House of Rothschild, which was a leading banking and finance house with great influence worldwide, during this period. It was a fact that some of the British media from these times, such as newspapers, termed the Mitsui family as 'The Rothschild of Japan' and

---

[2] Due to the huge number of studies, only a few important ones have been noted. See, for example, S. Yasuoka, *Zaibatsu Keiseishi no Kenkyu* (Studies on History of Zaibatsu's Management: Kyoto, 1970), A. Chimoto, 'Mitsui ni okeru Choki Kinzoku Shoreisaku no Shiteki Kosatsu'(A historical analysis of tactics adopted by the Mitsui House to encourage long-term employment), *Keieishigaku* 23-4 (1989), M. Kasuya, *Gosho no Meiji: Mitsui Kagyo Saihen Katei no Bunseki* (Wealthy Merchant and Meiji Japan: an analysis of the process of business reorganization in the House of Mitsui: Nagoya, 2002), H. Morikawa, *Zaibatsu: the rise and fall of Family Enterprise Groups in Japan* (Tokyo, 1992).

Mitsui Holdings was also certainly conscious of this title.[3] These media sources, other primary sources, and published documents illustrate why the Mitsui family considered the Rothschild family as an ideal model for their own business reforms. Therefore, we explore Mitsui's attempts to collect detailed information on the Rothschild family while formulating drafts of its own reforms. This had a decisive impact on a certain stage of the Mitsui family business reforms since they managed to seek successful methods to adapt their business to the new Japanese social, economic, and legal frameworks, which were crucial for their survival. Based on documents chronicling the business trips made by the Mitsui executives and family members, that sought examples of family ownership and corporate governance systems in Europe and the US, the details of their meetings and social contact with several members of the Rothschild family will be discussed below briefly.

Secondly, from recent studies and works concerning the Rothschilds, the comparison of the Rothschild and Mitsui families with respect to the framework of their business, ownership, and organisational systems will be analysed briefly. Recent academic contributions, such as Ferguson's monumental work on the history of the Rothschild family have revealed the details and reality of their household and business systems by exploring primary sources from the family's archives. This comparison also aims to investigate and consider whether or not the Rothschilds could be a model for the Mitsui family during the period of their business and organisational reforms.

A concise history of the Mitsui business and how it confronted problems during the late 19th century

The Mitsui business had a long tradition since its establishment during Tokugawa Japan. Its particular characteristics in terms of ownership and asset holdings had developed during a period of business expansion in

[3] See 'The Rothschilds of Japan: A firm established three centuries ago', January 21, 1895: *Pall Mall Gazette*. This article was translated into Japanese by the Mitsui Bank, one of affiliates of Mitsui Zaibatsu, and is currently kept at the Mitsui Archive under the title 'Nihon no Rosuchairudo', *Tsui-856-7*, Mitsui Archive (date unknown).

the early 18th century. The commonly owned business assets system, which became the core of their business, originated during its founding period. The founder, Mitsui Takatoshi, formalised this rule before his death since he foresaw that a cooperative relationship within the Mitsui family would be a core value for the maintenance of long-term ownership and management of the family business (Nakata 1988). This led to the original system of ownership, which became, nine Mitsui families (later eleven) owning the business assets commonly with each family receiving a portion of the profits annually based on their investment ratios (Mitsui Bunko 1976: 124-138). Thus, the status of each family unit within the larger Mitsui clan was defined through the extent of investments made. The family of the eldest son of the founder, whose investment was the largest among the families, was named *Soryōke.* He became the hereditary head of the Mitsui business. The other relatives of the founder's son were defined as *Honke*, which consisted of five houses and whose investment ratios and statuses were higher than the remaining families, and *Renke*, which consisted of three (later five) houses of the founder's daughters and other relatives (Mitsui Bunko 1976: 126-130). This complex ownership system comprised the core of the Mitsui family business until its dissolution after the Second World War.

After the Meiji Restoration, Mitsui had already undergone these household reforms because of the formation of a strong connection with the new Meiji government in terms of financing and banking (Mitsui Bunko 1980: 4-24, 387-430). However, the household reforms that began from the end of the 19th century were strongly related to the ownership and managerial reforms of the Mitsui business. One of the motivating factors for this reform was the compilation and enactment of the Japanese Civil Code that the Meiji government promoted enthusiastically, since it included as a condition the revision of the unequal trade treaties with European countries and the US (Gordon 2003: 46-50, 91-92, 119). The process followed to enable the enactment of the Code undoubtedly influenced the inauguration of the Mitsui reforms. While it was possible to exploit some of the content of the new Civil Code to maintain the traditional business strategy, for example, child adoption which was necessary to keep the multi-generational ownership of the family business intact, including the Mitsui families (Nakaoka 2011). However, the enactment of the Commercial Code, which contributed to shaping the modern Japanese business and firm system, had a strong impact on the Mitsui family since it had maintained the traditional merchant house style of business and management. Thus, the aim of Mitsui's business structure in this period was to fit Mitsui's business and management into the modern Japanese legal framework while retaining the traditional family ownership system from the Tokugawa period which maintains the business assets as commonly owned by multiple Mitsui families. As may be expected, it was difficult and time-consuming for those who were involved in the planning and implementation of these complicated reforms.

One of the leaders who played a significant role in the reforms was Inoue Kaoru, who was an important politician in the Meiji government and had a close relationship with the Mitsui family and its business concerns (Mitsui Bunko 1980: 51-71). Although he was appointed later as the Supervisor by the Mitsui families to oversee and reconcile their business and family affairs, he had already exercised substantial power over the family's decision-making process before gaining a formal position within the family. His strong connection with the Mitsuis was demonstrated by the fact that documents and papers related to Mitsui's important business decisions were always sent to Inoue.[4] It was commonly understood among many of Mitsui's managers in this period that without Inoue's strong initiative and leadership in conducting the reforms, and navigating the ensuing conflicts of power with regard to ownership and management within the family, they would have failed.[5]

Nevertheless, at the initial stage, the reforms made slow progress, such that the managers and advisers who participated in the process became frustrated. Some evidence suggests that one of the senior

---

[4] These documents and papers, which were copied by hand and sent to Inoue immediately, are preserved at Mitsui Bunko (Mitsui Archives) as multiple volumes titled *Inoue Koshakuke yori Kofu Shorui* (Documents delivering from the House of Marquis Inoue).

[5] For instance, see the comment of Masuda Takashi, one of the influential Mitsui business leaders during this period on Inoue's leadership in overseeing the household reforms in Mitsui Bunko (ed.), *Mitsui Hachiro'emon Takamine Den* (Mitsui Hachiro'emon Takamine: A Biography: Tokyo, 1988), p. 219.

managers of Mitsui submitted a letter of resignation during this period, which upset the Mitsuis.[6] Later, he confessed that the purpose of this attempted resignation, which did not materialise, was to facilitate the necessary reforms (Mitsui Bunko 1980: 437-438). Thus, meetings among managers and members of the Mitsui clan became more frequent after this incident and a framework of household reforms was formed in addition to organising commitment to reform among families and managers.[7] This incident also had a strong impact on the formalisation of business and managerial intervention by Inoue, since he became directly involved in the Mitsui reforms after this event (Mitsui Bunko 1980: 438-441). It should also be noted that he considered the compilation of *Mitsui Kaken* (Mitsui Household Rules), which contributed to the regulation of the ownership and managerial involvement of various Mitsui families, as the main focus of Mitsui reforms. He started this task immediately after he took charge of the Mitsui family and business affairs (Mitsui Bunko 1980: 458-460). Therefore, considerable progress in the reforms of the household began with the compilation of *Mitsui Kaken* from 1891 onward, under the leadership of Inoue. Although there were several issues that delayed the full compilation of *Kaken*, they managed to complete the inner household rules and issued it in 1900.[8]

From the process of revising the draft to completing the full version of *Kaken*, it is apparent that the main aim of making this private but

[6] This letter was written by Masuda Takashi, one of the important senior managers of Mitsui between the late 19th and the early 20th centuries. Masuda also played a significant role in the Mitsui household and business reforms which are discussed later in this study. 'Masuda Takashi Jihyo' (Letter of resignation: Masuda Takashi), 1st October 1890, I-Ko 1, Mitsui Archive.

[7] The original version of this oath was archived as 'Inoue Haku ni Teijo Dozoku narabini Juyaku Seiyakusho' (Presentation to Count Inoue: written oath of family members' and managers) November 1890, I-Ko 4, Mitsui Archive. This was also republished in Mitsui Bunko (ed.), *Mitsui Jigyoshi Shiryohen* (History of Mitsui's Business, Appendices) Vol. 3 (Tokyo, 1974), pp. 149-150. And it should be noted that Inoue firstly ennobled as a Count in 1884, and later became a Marquis in 1907 so that documents also reflected changes in his aristocratic status.

[8] For a more detailed explanation of several issues that delayed the full compilation, see, Mitsui Bunko (ed.), *Mitsui Jigyoshi* (History of Mitsui's Business) Vol. 2, pp. 458-464, 516-521; and also Mitsui Bunko (ed.), *Mitsui Hachiro'emon Takamine Den*, p. 218.

binding covenant was to prevent the possibility of the dissolution of the Mitsui business as a consequence of serious disputes within the family.[9] It is assumed that the strict regulation of the freedom of the Mitsui families through *Kaken* was largely due to the pressures of maintaining a traditional joint business system under a modern legal framework. Further, it was clear that by instituting this structure, the core of the Mitsui family business required the establishment of a common understanding not only for families, but also for managers and advisers.[10] However, since this system could not completely fit into the newly emerging legal framework, especially in the context of the Civil and Commercial Codes, the continuity of this core system inevitably led to the imposition of limitations on the freedom of the Mitsui families. Thus, any potential eventuality causing a serious ownership crisis had to be pre-empted by giving the power of constraint to *Mitsui Kaken*, which also implied that the 'Mitsui Family Constitution' applied in Japan independent of all public laws. It was strictly forbidden for Mitsui family members to take any of their family disputes to court.[11] This indicates that the Mitsui families had sacrificed their legal rights based on the Civil Code, including their property and household rights, in order to adhere to a 'traditional' ownership system and a 'modern' business system, which in reality caused major contradictions in practice.

---

[9] For more details on this procedure, see, 'Mitsui Kaken: Dai-ichi Soan' (the Mitsui Household Rule; 1st Draft) (undated), I-Ko 58, Mitsui Archive. Especially, see Section 3 (Marriage and Adoption) and Section 5 (Inheritance). The content of *Dozokukai* was found in Section 2 (Family Council). It should be noted that *Dozokukai* itself was officially established in 1893, before the proclamation of *Kaken*. The details of *Dozokukai* are presented in Mitsui Bunko (ed.), *Mitsui Jigyoshi* Vol. 2, pp. 535-541, and the codified rules for *Dozokukai* 'Mitsuike Dozokukai Kisoku' (regulations of Mitsui Family Council) are found in Mitsui Bunko (ed.), *Mitsui Jigyoshi Shiryohen* Vol. 3, pp. 281-283.

[10] The details of this system are in, Nakaoka 'Binding emotions for long-term continuity of family business? The foundation of family rule and Mitsui's business in the late 19th and early 20th century Japan', *Entreprises et Histoire* 91 (2018).

[11] This restriction was already included in the second draft of *Kaken* (see 'Mitsui Kaken Dai-ni Soan' (the Mitsui Household Rule; 2nd draft) Section 2), and was enforced further by deleting the part on the possibility of appealing before the court in the final version (Mitsui Kaken: Section 2 (Duties of the Families), Article 14).

Searching for an adoptable and useful model to organise family business in the new era

The *Mitsui Kaken* undoubtedly contributed to shaping and defining the in-house rules for ownership and management by the Mitsui families. However, other important issues, in particular, the reorganisation of the Mitsui business within the context of a modern legal framework, still remained unresolved. It became crucial for the families and members of the senior management to decide on a way to adapt their traditional asset-ownership system to the new Commercial Code since it was totally beyond the scope of the existing legal framework. Therefore, exploring the feasibility of maintaining its system under the new legal order became Mitsui's priority. This explains why they urgently required support and advice from foreign legal and business experts, since the new Japanese legal system was based on European (mainly French) case law.

Yet, from the beginning, they faced difficulties in their business system. Several legal advisers, who also contributed to the codification of the Japanese Civil Code attempted to deal with these problems in light of Mitsui's uniqueness. The experts' embarrassment and confusion was reflected in their response to Mitsui's questions concerning their corporate reorganisation.[12] The issue was complicated further by the enactment of the Commercial Code. This was largely due to the fact that the regulation of limited liability businesses was not included in the Code which legally defined the types of new companies.[13] This caused great problems

[12] See 'She Boasonado Shi no Henshin' (Reply from Mr. Boasonade) (undated) I-Ko 99, and 'Roesureru Shi no Hento' (Reply from Mr. Roesler)(undated) I-Ko 100, Mitsui Archive. Both Boasonade, the French lawyer, and Roesler, the German lawyer, played significant roles in the formation of the Japanese legal system as formal legal advisors of the Meiji Government. The former was involved in the enactment of the Civil Code, and the latter was involved in the drafting of the Imperial Constitution. Some evidence suggests that the opinion of Roesler was considered important for Mitsui's corporate reorganization process. See Mitsui Bunko (ed.), *Mitsui Jigyoshi* Vol. 2, pp. 453-455.

[13] Another possible option for business reorganization, for instance, the establishment of a limited liability company, was not legalized in Japan until 1938. This option was already available in Germany, which was one of the models for the new Japanese Commercial Law, by the end of the 19th century. For more details, see, T. Sakamaki, 'Yūgen Gaisha Hōsei no Hatten to Gaikoku Hō no Eikyō' (The Development of a Legal

for large Japanese businesses that experienced rapid expansion and diversification, including Mitsui, and thus, they had to reorganise their businesses. The only option for firms in this situation was to set up a 'Gōmei Gaisha' or 'Goshi Gaisha,' a type of Japanese partnership which had been approved of in the form of an unlimited liability business under the Commercial Code.[14] For Mitsui, the reorganisation of their business into a Japanese-style partnership had a potential negative impact on its asset-ownership system for if its affiliate company became bankrupt, that meant the family had to repay any debt under a system of unlimited liability.[15]

In addition, there was another reason which forced the acceleration of business and organisational reforms. As part of the major taxation reforms at the end of the 19th century, the government formally introduced a corporate income tax regime. The government formally instituted a lower tax rate for this newly introduced tax to encourage organisational reforms among merchant houses and businessmen.[16] This issue was further complicated by the outbreak of the Russo-Japanese war. On the one hand, the Japanese government prioritised financing this war through the international capital market (Sherman 1983). On the other hand, the government also eventually introduced a special corporate income tax

---

System to establish a Limited Liability Company and Impact from the Foreign Legal System), *Hikaku Hōgaku* 2-1 (1966), pp. 89-118.

[14] For more details on 'Gomei Gaisha', see, Article 80, Commercial Law, which defined unlimited liability for capital investors. Though 'Goshi Gaisha' partly accepted limited liability for the partner, those partners could not engage in business management of or decision-making for the company, since these rights were retained solely by the unlimited partner under the 'Goshi Gaisha' arrangement. See, Article 146, 156 and 157, Commercial Law.

[15] This formed a temporal introduction of an eclectic strategy for business expansion and diversification during the period of reorganization. Under this strategy, affiliate companies of Mitsui were formally founded by the head of each house of Mitsui so that only one head bore the responsibility of investment, and therefore did not influence the commonly-owned business assets in the Mitsui families. See, Mitsui Bunko (ed.), *Mitsui Hachiro'emon Takamine Den*, pp. 185-187.

[16] As a result of these tax reforms, non-corporate merchant houses and firms applied the personal income tax rate directly, which was higher than the corporate income tax rate, and thus, that also had a negative impact on their business. For details, see, M. Katsu, *Nihon Zeisei Enkakushi* (Fiscal History of Japan: Tokyo, 1938), esp. pp. 237-38.

to finance the war, which became a permanent tax that remained even after the war (Ōkurashō 1950: 1000-48). This nascent tax regime also caused serious problems for Mitsui since it had completed a fixed process of applying lower tax rates for corporations while other forms of companies, including partnerships were required to pay higher taxes (Yoshida 1998: 45-76).

It became inevitable, therefore, that after the completion of *Mitsui Kaken* and the supplementary rules, the families and members of the senior management focused on organisational reform of Mitsui. The necessity for this business reorganisation was further accelerated by the development of affiliates that made the maintenance of the traditional ownership framework more complex and difficult (Mitsui Bunko 1985: 248-50). Nevertheless, their efforts towards and attempts at reorganisation within the options available based on the Japanese legal framework, such as the plan to establish new 'Gōmei Gaisha' to integrate all of Mitsui's business and affiliates, failed largely due to disputes with regards to their plans among the senior management (Mitsui Bunko 1980: 732-741). Thus, they recognised that they needed to find suitable models from European and American family businesses to solve this complex and serious problem. Consequently, several business trips were made to these countries to discover reliable business and organisational models to learn from, as well as notable businessmen or legal professionals who could act as advisers.

One noteworthy business trip, in particular, took place in 1907. One of the Mitsui businessmen, Masuda Takashi, who was then in senior management, led the visits to four countries, namely Britain, France, Germany, and the US, along with several assistant staff to conduct interviews and collect information.[17] Others who accompanied them included a representative of the Mitsui family, Mitsui Takakage, the housemaster of one of *Honke*, along with his wife, daughter and relative,

[17] Detailed documents on this business trip were prepared after their return, which was maintained as the 'Oubei Shuccho Fukumeisho' (Report on official trip to western countries) 1907, I-Ko 180-187, Mitsui Archive (also reprinted in Mitsui Bunko (ed.), *Mitsui Jigyoshi*: *Shiryohen*, pp. 487-581). These documents were probably written under Masuda's instructions.

in addition to other influential senior managers.[18] Notable European and US businessmen who were consulted during this visit included Nathan and Alfred de Rothschild, Sir Ernst Cassel, the Gibbs family, James Rothschild, Albert Kahn, Max Warburg, Jacob Schiff, and James Farrell.[19] Some Japanese diplomats and ambassadors supported this trip directly and indirectly since they understood the importance of this task and the significance of the Mitsui enterprises for Japan's business and economy.[20] It may be assumed that during this business trip, Masuda had understood that the Rothschild case was a possible option to rely on in implementing further reforms of the Mitsui business and household.

There are several reasons for this. First, there were some similarities in the family business systems followed by Mitsui and Rothschild, that is, a cooperative relationship among the members of multiple families who served to manage and control the family business. Though the framework and content of the family business network was considerably different between both families, as will be discussed later, the nominal similarity was sufficient for Masuda, who was urgently seeking an applicable case from among the existing European and American business families.[21]

[18] Though the remaining documents did not refer to any member on this trip except Masuda, the newspaper articles mentioned some of the other members who participated in this significant event. However, most of the articles considered this trip as either 'leisure travel' or a 'formal visit'. This indicates that the real purpose was not revealed by Mitsui to the media. For example, see 'Shiberia mawari Ryūkō' (Travel via Siberia becomes fashionable), *Asahi Shinbun* 10 June 1907, p. 2; and 'Mitsui Fujin Ikko no Kinshin' (Current situation of Mrs. Mitsui and her party), *Yomiuri Shinbun* 23 June 1907, p. 3.

[19] The full list of businessmen and professionals they met can be found in 'Oubei ni Okeru Kyuka Gouka Iji no Hoho oyobi Ginko Kinyugyo no Soshiki Keiei ni kansuru Chosa Kiyo' (Bulletin of strategy for household maintenance in western notable wealthy families and report on management of organization of Banks and Enterprises) 1907, I-Ko 180, Mitsui Archive.

[20] As an example of their assistance, see 'Futsukoku ni okeru Meishi tono Kaikenroku' (Report on meetings with French notable businessmen) 1907, I-Ko 182, Mitsui Archive (reprinted on *Mitsui Jigyoshi Shiryohen* Vol. 3; pp. 534-545).

[21] Masuda's report on this business trip also displayed his interest in the Rothschild family as a possible ideal solution for Mitsui's problem. See, 'Rosuchairudo ke no Soshiki oyobi Eigyo no Hoho' (Memorandum of Rothschild's Business organization and management) in *Oubei Shuccho Fukumeisho* (reprinted version; pp. 506-508).

Second, since the Mitsui family was confronted with difficulties pertaining to the maintenance of its traditional asset holding system under a newly established and westernised Japanese legal framework, the house of Rothschild was considered an ideal model on which to base its strategies for resolving its problems. In particular, Masuda was very interested in the relative long-term continuity of the House of Rothschild's business within a complicated network of families which occasionally led to serious disputes among the members of these families.[22] Mitsui interpreted this as being a case of a similarity that it shared with the house of Rothschild, within the context of their historical experience of the complex family ownership system. Thus, it was believed that the House of Rothschild was a model that was applicable for its own reforms.

Third, although this was a misunderstanding, Masuda assumed that European and American business families had their own codified inner house rules or contract clauses equivalent to the household rules of Mitsui, *Mitsui Kaken*, which governed the organisation and control of their family business.[23] Since one of the major aims of this business trip was to collect these documents for reference on organisational reform strategies, Masuda and his assistant staff struggled to discover even a few western examples of such documents, while coming to understand that such examples were rare in both European countries and the US cases.[24] It was also plausible that they interpreted the Rothschild case as the best example, if the Rothschild household rule had existed in reality.

Consequently, the meeting with several noteworthy members of the house of Rothschild during this business trip, including Nathan and Alfred of the House of London, James of the House of Paris, and Alon who was a member of the board of trustees in Paris, ended with some disappointment for Masuda and his assistants. Most of the discussions

---

[22] Ibid, p. 506.

[23] *Oubei Shuccho Fukumeisho*, pp. 508-509.

[24] However, he attempted to explore some rare cases with the assistance of his diplomatic association with the Japanese Ambassador. Ibid; p. 509.

with Nathan concentrated on the explanation of the business framework followed by Mitsui, and the meeting with Alfred was not fruitful at all.[25] Both of them showed no interest in explaining information specific to Rothschild's business and organisational system during the meetings. In contrast to the case of the House of London, James, the head of the House of Paris, displayed a friendly attitude towards Masuda. However, he was also reluctant to accept Masuda's request, and the memorandum he sent to Masuda after this meeting, which aimed to give him some advice with regard to Mitsui's business reform, was certainly incomprehensive and unsatisfactory.[26] Alon's comment on the House of Rothschild and its business was limited to the framework of their business investment.[27]

Consequently, the documents on the meetings with the members of the House of Rothschild indicated that the results were far from the expectations of the members of the Mitsui delegation. Even when compared with other documents on meetings with notable businessmen or financiers, the contents of these were considerably less useful for the Mitsui business reform, which led to the foundation of its own holding company, *Mitsui Gōmei Gaisha*. Documents chronicling this business trip suggested that the attitudes of the members of the House of Rothschild were suspicious of the intentions of Mitsui representatives. The latter believed that relevant information was kept secret as Masuda still made attempts to obtain more information on the House of Rothschild even after the business trip. However, it also raised another question as to whether the House of Rothschild would truly be applicable to Mitsui's reforms, and accordingly, if it would make a difference for Mitsui to seek complete and comprehensive information on the House of Rothschild and its business and organisational systems.

---

[25] The details of the meetings with Nathan and Alfred were mentioned in 'Eikoku ni okeru Shomeishi tono Kaikenroku' (Report of meeting with British notable business-men) in *Oubei Shuccho Fukumeisho*, esp. pp. 510-512.

[26] See, 'Futsukoku ni okeru Shomeishi tono Kaikenroku' on ibid; pp. 542-543.

[27] 'Futsukoku' on ibid; pp. 537-539.

Image and reality: could Rothschild be a model for Mitsui to reorganise business?

Although Masuda was convinced that the house of Rothschild would be some form of a model for the solution to Mitsui's business and legal problems, it should be noted the period of Mitsui's visit coincided with a period of upheaval and crisis for the House of Rothschild. Before their visit, the House of Rothschild was no longer a partnership of five houses that had originated with the founder. The House of Naples, one of the five, had already closed their business in the 1860s and re-settled in their birth place, Frankfurt (Barbagallo 2000: 306). The House of Frankfurt ended in 1901 due to the lack of a male heir (Kuper 2001: 277). This was largely due to their strict patrilineal system of succession in combination with the practice of choosing their marriage partners within the inner family circle (often cousin marriages) (Kuper 2001). This consequently limited the continuity of their business and family ownership in its original form. Therefore, although Masuda urgently sought detailed information on the House of Rothschild even after this visit, his knowledge and information of this family had already been out-of-date and thus impracticable. However, some questions remained as to whether or not the House of Rothschild could ever be a model for Mitsui's household and business reforms if Masuda (or the others that accompanied him) had been able to access detailed and accurate information on the House of Rothschild.

Recent studies of the Rothschild archives suggest that there are some problems in accessing the more significant primary sources, which may provide better opportunities for a more detailed analysis of this event, since correspondence or letters with interviewees, in particular Nathan, were not stored or retained (Gray 2009: 34). This indicates a difficulty in tracing or exploring personal impressions, such as for example, the response and evaluation of Mitsui's visit from their exchange of information among private participants, such as personal letters. However, even from the results of the limited survey of the secondary sources, it can be seen that there was considerable differences between Mitsui and Rothschild with respect to their household and business systems. Thus, it would have been difficult to apply the House of Rothschild's business organisation to Mitsui's business reforms, even if Masuda and his assistant was able to glean relevant information that would be useful to Mitsui's need for reform.

First, their business characteristics, such as their core business or the level of diversification, most likely was a core difference between their business strategies; as was their business and governmental relationships. As discussed briefly, Mitsui rapidly formed a huge business complex during this period, and entered into various business fields including industrial, commercial, and financial sectors. In contrast, the House of Rothschild's business was almost entirely focused on the fields of banking and finance. Even in their mining sector activity, their business interests focused solely on precious metals such as gold, silver, and mercury which was necessary for their silver refining operations (Ferguson 2009: 12). The advice of Nathan Rothschild's to Masuda conflicted with experience of Mitsui. Nathan recommended to Masuda the withdrawal of Mitsui from the coal mining sector, which was an early core business of Mitsui being given to Mitsui by the Japanese government without any international involvement. According to Nathan's advice, coal was a 'highly risky sector' which probably was correct in Europe but not in Japan where coal mining sites were granted directly by the Japanese government to Mitsui which lowered any risk taking. In addition, being integrally linked to both the domestic and global political levels of the economy meant that the House of Rothschild had their core business dependent on government financing and the development of the international bond market; which inevitably led to both direct and indirect pressure on committed nations which underpinned their core business strategy (Knight 1984, Weller 2015). This clearly demonstrates a contrast with Mitsui's case since they kept a distance from the Japanese government by operating the coal mines as Mitsui enterprises separate from government political control. Mitsui involvement with the political sphere was relatively limited and distant when compared to the House of Rothschild.

Second, with regards to the legal framework that defined their family business, this difference influenced the method of profit sharing strongly. From the inception of its business, the profit sharing method of the House of Rothschild was based on a frequently revised partnership agreement among the five houses (Ferguson 1999: 267). In contrast, while it took longer to form a hierarchal and complex family structure in Mitsui, the profit sharing ratio among the 11 families was largely based on the household rule that had prevailed in the Tokugawa period (Mitsui Bunko 1980: 126-128). While the familial and business disputes within the families influenced the changes and the revisions made to the agreements in the House of Rothschild (Ferguson 1999: 268-272), no considerable revisions were made in the Mitsui case although some serious family disputes or business crises had shaken their fraternal harmony (Mitsui Bunko 1980: 326-338). It may be assumed that senior managers of Mitsui, including Masuda, who had implemented the business reforms, would never be accepted within the Rothschild system of profit sharing since it had the potential to destroy the Mitsui household and business systems.

Third, there was a certain difference in the levels of cross-border business activities between the Houses of Rothschild and Mitsui. The House of Rothschild was unique in this context in that its banking and financial operations were multinational from its inception. Although they originated in Frankfurt, their business success depended on cities abroad, such as London or Paris, where the founder's sons had settled (Cassis 2006: 21-22). On the other hand, although they had evolved into multinational business activities, especially after the late 19th century, Mitsui's business headquarters remained in the capital city of their home country.[28] In the case of Rothschild, the nature of their transnational business activities were also linked with that of the Europeans, and in later international political risks, which directly influenced their earnings and performance.[29] Additionally, with their position as an ethnic and religious minority, these risks required a certain measure of flexibility in their business activities, so that may have possibly influenced the profit sharing method as discussed above. Compared to Rothschild, Mitsui's

[28] The only case to move their business headquarters within Japan was in the late 19th century, when they decided to move their headquarters from Kyoto, the old capital, to Tokyo, the new capital. See, 'Dōmyō Icchi Kesshin Seimeisho' (Statement for Family's agreement) in Mitsui Bunko (ed.), *Mitsui Jigyoshi: Shiryohen Vol. 3*, pp. 136-137 (the original is kept at the Mitsui Archive, *Zoku2390*-3).

[29] See, Ferguson, *The House of Rothschild: Money's Prophets*, pp. 437-483, in the case of being confronted with the crisis of the 1848 European revolution.

business was more national rather than transnational, and therefore was solely conscious of internal changes and governmental reforms, including domestic legal and taxation reforms.

Fourth, there were considerable differences in the family networks of both houses. Although both gained prestigious positions as wealthy businesses, which, in turn, resulted in the gradual expansion of their family networks beyond the business world through marriage alliances with nobility and political elites, the influence of 'the aristocratisation of social networking' was certainly different in both cases. In the Rothschild case, there was an acceptance of a gentlemanly lifestyle that had a strong impact on their day-to-day business (Ferguson 2000: 234-235). With newer generations, many of the family's daughters married outside their faith and family circles, regardless of countries in which they settled (Ferguson 2000: 244-256). The cognate relationships within the five houses became diluted and resulted in the demise of a few of the branch lines, such as the Frankfurt and Naples families (Ferguson 2000: 237, 241-242; Kuper 2001). Consequently, the marriage network helped the family survive and become diverse and aristocratic, which in turn accelerated the weakening of family relationships.

On the other hand, while some testimonies from the Mitsui family showed that they interpreted their marriage alliances after the modern period as aristocratic, and that although several heads of the Mitsui families were granted the title of 'Baron', they were still conscious about maintaining their family network through intermarriage within the Mitsui families (Yasuoka 1979). They also strategically utilised adoption, which formed a tradition that they included within the modern Japanese legal system – wherein they adopted children from within and outside the family network to maintain the continuity of the household and business.[30] In the late 19th century, when Mitsui was in the preliminary stages of household reforms, they decided to revive several defunct

[30] For example, see, Mitsui Bunko(ed.), *Mitsui Hachiro'emon Takamine Den*. Takamine, who was the head of Mitsui *zaibatsu* during the late 19th to the early 20th century, was also adopted by his childless brother, the former head, as the next heir of Mitsui. The description of the adoption system (in particular, the case of the Japanese elite) is presented in T. S. Lebra, *Above the Clouds* (Berkeley, 1993), esp. pp. 107-132.

houses to maintain the original ownership system among the 11 families by utilising the adoption system within the family network instead of reorganising it entirely (Mitsui Bunko 1988: 70-71, 183-184). This return to historical origins with regard to household organisation influenced the particular structure of Mitsui and the content of future reform in the early 20th century.

Finally, some differences can be seen if we compare the roles of the female members in both houses. Though the daughters of the Rothschild and Mitsui houses both played important roles in expanding social networks through favourable marriages there were marked differences. In the case of the Rothschilds the sons of the founder formed five houses to conduct the family business while the Mitsui daughters of the founder also became formal members (Mitsui Bunko 1980: 48-49). This indicates that the position of women with regard to business and household activities differed, since the female line and sons-in-law were excluded from the Rothschild's business and inheritance (Ferguson 2000: 6). After the modern period, female members of the Mitsui families were still considered important members in maintaining the family line through endogamy and as a spouse of an adopted heir of 11 families, even if female members were not absorbed into the family business (Yasuoka 1979: 115).

Though it is necessary to explore and discuss this further, the evidence and the differences described above suggest that the House of Rothschild could not have been a suitable model for Mitsui. Apart from the role of women, the introduction of the gentlemanly lifestyle of the heads of the House of Rothschild among the Mitsui men, for example, would have been met with objection, especially from Mitsui's supervisor, who criticised the pursuit of luxury among Mitsui families members in strong terms. The decision-making process concerning the profit sharing ratios in the House of Rothschild may have probably been recognised as a serious risk with the potential to lead to the destruction of Mitsui's business as a result of serious family disputes. Rothschild's transnational business framework would not have fit with Mitsui because their managerial and organisational systems would have required coordination with institutional frameworks in various countries. Therefore it was only assumed that the knowledge of or information on the House of Rothschild and its business system was invaluable for Mitsui to apply to its own business reform standards. The question that remained was whether or not Mitsui was successful in proceeding with its own reforms without any useful model from among European or US family businesses to rely on.

Conclusion

Though the members of Mitsui that went on the 1907 business trip could not find a suitable and applicable model from among the western business families, they managed to reorganise Mitsui's business. This led to the establishment of their own holding company in the form of the Japanese-style partnership, *Gōmei Gaisha*, in 1909.

This system was chosen largely due to its utility as a holding company, which worked under the Japanese Commercial Code.[31] Given that they shaped Mitsui's managerial and organisational reforms, and established the holding company, this business trip was quite meaningful. First, by exploring business models that were applicable and adaptable, they deepened their understanding of the contemporary western corporate system. Some aspects of the corporate systems in these countries, for example, the private companies in Britain, the Advisory Board (*Aufsichtrat*) system in Germany, and the holding company system in the US, were identified by Mitsui's managers as useful examples for their business reforms to follow.[32]

Moreover, they became fully aware of the special nature of Mitsui's business model with respect to the system of family ownership. For example, during this trip, they urgently sought out European businessmen who managed their own family businesses and had in-house family rules

[31] See 'Oubei Shisatsu ni yori Eigyō Soshiki ni kansuru Hiken' (Personal opinion for organizational reforms based on western business trip) in *Mitsui Jigyōshi Shiryōhen* Vol. 3 (Tokyo, 1974), esp. pp. 584-585.

[32] See, 'Oubei ni Okeru Kyuka Gouka Iji no Hoho'.

similar to the *Mitsui Kaken*.[33] Thus, Mitsui's managers were very interested in the business system of the House of Rothschild and suspected that there were some hidden and secret family rules that were kept by other major business families, such as Krupp, a German industrial giant. However, since primogeniture or entailment formed the basis of inheritance of family fortunes, including the business, such family rules did not exist in these countries.[34]

Mitsui was fortunate to have received the most useful advice during this business trip from Max Warburg, a prominent banker from Hamburg who befriended Takahashi Korekiyo, a well-known politician and financier in modern Japan.[35] It is assumed that this personal contact helped them make an appointment with Max Warburg.[36] Reliance on his advice by Mitsui's manager was largely due to his sincere effort to give meaningful advice to them. The memorandum he sent to Mitsui contained important business advice, including ways to split Mitsui's main businesses, such as trading, banking, and mining, into independent joint-stock companies. Warburg also provided instructions for setting up a holding company that would own and manage these separate companies.[37] In addition, the memorandum pointed out that the Rothschild family was not a useful model for Mitsui to follow, and presumably reflected his apprenticeship experience at New Court, the headquarters of the London House in the 1890s (Ferguson 2000: 234). He was the first person to have been allowed to work at the business heart of the Rothschild family, despite being an outsider (Ferguson 2000: 518). Thus, it is assumed that while Mitsui failed to obtain inside information on the members of the Rothschild

---

[33] See, 'Rothschild-Ke no Soshiki oyobi Eigyo no Hoho'.

[34] However, they doubted the existence of in-house rules in these countries. Thus, they requested that Japanese diplomats in Europe continue this investigation. See, *Oubei Shuccho Fukumeisho* (esp. pp. 508-509).

[35] The details of the relationship with Takahashi can be found in his memoir. See M. Warburg, *Aus Meinen Auszeichnungen* (New York, 1952), pp. 19-20.

[36] See the introductory part of 'Dokkoku Meishi Kaikenroku' (Report of meeting with German notable businessmen) 1907, I-Ko 183, Mitsui Archive.

[37] See 'Mitsuike no Soshiki ni kansuru Makkusu Woruborugu Shi no Ikensho' (General suggestion from Max Warburg on Mitsui's business organization), (reprinted on *Mitsui Jigyoshi:Shiryohen Vol. 3*, pp. 561-565).

family, they did have an important adviser, who also happened to be a rare insider in the Rothschild business concerns during this period.[38]

By exploring foreign business systems and organisations, Mitsui deepened their understanding of western business, which was considered the most advanced system in the eyes of contemporary Japanese businessmen. Through the experience of seeking out business and managerial information from western countries, Mitsui recognised the variety and diversity of such systems which were based on a combination of tradition, historical experience, and differences in the development of various social and legal systems. It is assumed that through the process of learning and collecting information during this business trip, they had understood that the case studies of western business families, including the House of Rothschild, were not applicable to Mitsui's business and household reforms.

Therefore, instead of relying on the case of a particular western business family, they successfully applied what they had learned generally to their organisational reforms. For instance, they learned about splitting their business into independent joint-stock companies. Founding a holding company through Japanese partnership (and unlimited liability) took the form of a combined solution that applied both Warburg's advice and some of their own revisions through the adaptation of their business traditions. Sources and reports on this reform, in particular, the documents and papers from their business trips, indicate that the intellectual skills and understanding of the influential senior managers of Mitsui, such as Masuda Takashi, reached higher levels after their visit.

However, evidence related to the establishment of *Mitsui Gōmei Gaisha* also suggests that the reform did not progress smoothly. Masuda needed a prolonged time to obtain the consent of the stakeholders, including the members of the family and the managers of the business

[38] Masuda also mentioned that most important information on the House of Rothschild during the 1907 business trip was gained through the meeting with Max Warburg. See, 'Hanburuku ni oite Makkusu Wobagushi tono Kaiken' (Report on meeting with Max Warburg in Hamburg) in *Oubei Shuccho Hokokusho* (reprinted on *Mitsui Jigyoshi: Shiryohen Vol. 3*, pp. 545-553). However, he did not mention his experience at New Court during this meeting.

to implement this plan.[39] The remaining documents also indicate that Inoue Kaoru, Mitsui's influential supervisor, played a crucial role in silencing and limiting the objections, in addition to monitoring the reorganisation.[40] His speech on the day of the inauguration of *Mitsui Gōmei Gaisha* for the members of the Mitsui Family, stressed the necessity of this reform for the stakeholders of Mitsui: a seeming reflection of some reluctance from the family members about implementing these reforms.[41]

The drafts and memorandums concerning the creation of the Corporate Charter of *Mitsui Gōmei Gaisha* demonstrate the importance of *Mitsui Kaken*. Evidence suggests that this reform was given priority so as not to contradict the contents of *Kaken* in order to avoid any possibility of revision or overhaul of this in-house rule. For instance, in the drafting stage, the Corporate Charter contained articles relating to corporate dissolution that were omitted from the final version.[42] Further, articles relating to the legal disputes among the members of the holding company were deleted from the Corporate Charter.[43] It may be assumed that this deletion was because these contents contradicted *Kaken*. This is especially applicable for the content relating to legal disputes, since it would have undermined the articles of *Kaken* that prohibited legal disputes within

[39] The letter from Masuda to Inoue Kaoru, Mitsui's Supervisor, during the period of the reforms indicates some difficulties in implementing this reorganization. Letter from Masuda Takashi to Inoue Kaoru, 17th June 1909 (quoted on *Mitsui Jigyōshi* Vol. 2, p. 759).

[40] It is certain that the draft, plan, and details for the establishment of *Mitsui Gomei Gaisha* were in Inoue's hands. Thus, Inoue had great influence and power on the entire decision-making process. These documents are also compiled on volume 26 of *Inoue Koshakuke yori Kofu Shorui* in Mitsui Archive.

[41] 'Mitsui-Ke Eigyososhiki Kaisei ni tsuki Inoue Koshaku Enzetsu' (Speech from Marquis Inoue with regard to Mitsui's organizational reform) 8 October 1909, I-Ko 238, Mitsui Archive.

[42] See, 'Mitsui Gomei Gaisha Teikan' (Corporate Charter, Mitsui Holding co.) (undated) I-Ko 218, 219, Mitsui Archive. These drafts contained the articles that defined the process of corporate dissolution; however, this information was not a part of the final version of the Corporate Charter.

[43] See, ibid; the section 8 'Chusai' (Arbitration) which was about the legal dispute of the holding company. This was also deleted from the Corporate Charter.

the Mitsui family. This assumption was further reinforced by other documents containing basic instructions governing the business reforms for the foundation of *Mitsui Gōmei Gaisha.* These documents indicated that the future revision of *Kaken* in order to implement this reform was strictly forbidden. This was in addition to the notice that the general meeting of *Gōmei Gaisha* was, in reality, substituted by *Dozokukai*, the family council defined by *Kaken*, informally.[44]

This evidence indicates that the conformity of the newly established *Gōmei Gaisha* with *Kaken* and its consistency with the law were given importance. Though advice from foreign businessmen, especially from Max Warburg, undoubtedly contributed to the formulation of new legal and business frameworks for Mitsui, the newly emerged business complex had to shape and redefine itself since it had to explore its own route towards finding a solution to maintain the traditional ownership system, which was difficult for foreign advisers to understand. Ironically, their commitment towards the maintenance of the traditional ownership system, in addition to their in-house rule, *Kaken*, during this process, presumably indicates that even if they had been able to access sufficient information and details on the Rothschild's household system, they would not have utilised it as an ideal or an applicable model for their reform.

References

Archival Sources (*Mitsui Bunko* (Mitsui Archive)):

Inoue Koshakuke yori Kofu Shorui (Documents delivering from the House of Marquis Inoue; abbreviated as *I-Ko*)

Mitsui-Ke Kirokubunsho: Betsu (Documents on the House of Mitsui and businesses: supplementary materials; abbreviated as *Betsu*)

[44] See, 'Jikkojo no Kokoroe' (Instruction), I-Ko 228 (reprinted on *Mitsui Jigyoshi Shiryohen* Vol. 3, pp. 597-598). This document is the most basic instruction for the establishment of the holding company. It was secret information and was maintained in-house.

Published Sources:

Mitsui Bunko (ed.) *Mitsui Jigyoshi: Shiryohen* 3 (History of Mitsui's Business: Appendices Vol. 3), Tokyo, 1974.

Newspaper Database:

The British Newspaper Archive (https://www.britishnewspaperarchive.co.uk)

Kikuzō (article database of *Asahi Shimbun*: https://database. asahi.com)

Yomidasu Rekishikan (Historical database of *Yomiuri Shimbun*: https://database.yomiuri.co.jp/rekishikan/)

Secondary Sources:

Akita, G., *Foundations of Constitutional Government in Modern Japan 1868-1900*, Cambridge Mass. 1967.

Barbagallo, F., The Rothschilds in Naples, *Journal of Modern Italian Studies*, Vol. 5-3 2000.

Cassis, Y., *Capitals of Capital: A History of International Financial Centres, 1780-2005*, Cambridge 2006.

Cassis, Y. and Cortrell, P. (eds.), *The World of Private Banking*, Ashgate 2009.

Chimoto, A., 'Mitsui ni okeru Choki Kinzoku Shoreisaku no Shiteki Kosatsu' (A historical analysis of tactics adopted by the Mitsui House to encourage long-term employment), *Keieishigaku* 23-4 1989.

Ferguson, N., *The House of Rothschild: Money's Prophets 1798-1848*, New York, 1998.

The House of Rothschild: The World's Banker 1848-1999, New York 1998.

'The Rise of the Rothschilds: the Family Firm as Multinationals', in Cassis Y. and Cortrell P. (eds.), *The World of Private Banking*, Ashgate, 2009.

Gordon, A., *A Modern History of Japan: From Tokugawa Times to the 4 Present*, Oxford 2003.

Gray, V. and Aspey, M., 'The Rothschild Archive', in Cassis and Cortrell (eds.), *The World of Private Banking*, Ashgate 2009.

Jansen, B.M., *The Making of Modern Japan*, Cambridge Mass., 2000.

Kasuya. M., *Gosho no Meiji: Mitsui Kagyo Saihen Katei no Bunseki* (Wealthy Merchant and Meiji Japan: an analysis of the process of business reorganization in the House of Mitsui), Nagoya, 2002.

Katsu. M., *Nihon Zeisei Enkakushi* (Fiscal History of Japan), Tokyo, 1938.

Knight, G.A., 'The Rothschild-Bleichröder Axis in Action An Anglo-German Cooperative 1877-1878', *Leo Baeck Institute Yearbook* Vol. 28-1 1983.

Kuper, A., 'Fraternity and Endogamy: The House of Rothschild', *Social Anthropology* Vol. 9-3 2001.

Lebra, T.S., *Above the Clouds: Status Culture of the Modern Japanese Nobility*, Berkeley, 1993.

Mitsui Bunko (ed.) *Mitsui Hachiroemon Takamine Den* (Mitsui Hachiroemon Takamine: A Biography), Tokyo 1985.

Mitsui Jigyoshi (History of Mitsui's Business) *Vol. 1*, Tokyo 1976.

Mitsui Jigyoshi (History of Mitsui's Business) *Vol. 2*, Tokyo, 1980.

Morikawa, H., *Zaibatsu: the rise and fall of Family Enterprise Groups in Japan*, Tokyo, 1992.

Nakaoka, S., 'Yōshi Sōzoku to Familii Bijinesu: Kindai Nihon Fuyū Kigyōka Shōninsō no Jirei kara' (Adoption, inheritance, and modern Japanese family businesses: the case of the wealthy economic elite) *Shakaikeizaishigaku* 76-4 2011.

'Binding emotions for long-term continuity of family business? The foundation of family rule and Mitsui's business in the late 19th and early 20th century Japan', *Entreprises et Histoire* 91 2018.

Nakata, Y., *Mitsui Takatoshi* (Mitsui Takatoshi: A Biography), Tokyo 1988.

Ōkurashō (ed.), *Meiji Taisho Zaiseishi* (Fiscal History in the Meiji and Taisho Period) Vol. 6, Tokyo 1950.

Sakamaki, T., 'Yūgen Gaisha Hōsei no Hatten to Gaikoku Hō no Eikyō' (The Development of Legal System to establish Limited Liability Company and Impact from the Foreign Legal System), *Hikaku Hōgaku* 2-1 1966.

Sherman, A.J., 'German-Jewish Bankers in world politics: the financing of the Russo-Japanese war', *Leo Baeck Institute Yearbook* Vol. 28-1 1983.

Taplin, R., Decision Making and Japan: A study of corporate Japanese Decision-Making and its Relevance to Western Companies, Routledge: Oxon 1995 and 2016.

Warburg, M., *Aus Meinen Auszeichnungen*, New York, 1952.

Weller, L., 'Rothschilds "Delicates and Difficult Task": Reputation, Political Instability and the Brasilian Rescue Loans of the 1890s', *Enterprise and Society* Vol. 16-2 2015.

Yasuoka, S., 'Mitsui-Ke Dōzoku no Kon'in to Sōzoku: Mitsui Reiko Shi tono Taidan' (Marriage and Inheritance of the House of Mitsui: An Interview from Mrs. Mitsui Reiko), *Doshisha Shōgaku* 30 5-6 1979.

Mitsui Zaibatsu no Hitobito: Kazoku to Keieisha (People of Mitsui Zaibatsu: Compiled Records of Interviews from Mitsui's Family Members and Managers), Kyoto 2004.

Zaibatsu Keiseishi no Kenkyu (Studies on History of Zaibatsu's Management), Kyoto, 1970.

Yoshida, J., *Nihon no Kaishaseido Hattatsushi no Kenkyu* (Studies on History of the Development of Japanese Corporate System), Ryūgasaki 1998.

Interdisciplinary Journal of Economics and Business Law, 2019, Vol. 8,
Issue 3, pp. 79-92

Assessing Debt Sustainability in Zambia

M. Chengo[1] and J.P.S. Sheefeni[2]

Abstract

Debt has been a major concern in most developing countries including Zambia and it can be traced back to their dependence on mineral resources. Majority of these nations have landed into unsustainable debt which has greatly hindered economic growth and development. This paper looked at what sustainable debt is and the different ways it can be determined. It also aimed at determining where Zambia's current debt stands and to compare two different approaches of analysing debt sustainability. Majority of the analysis on debt sustainability has been conducted using debt ratios such as the debt-to-GDP ratio, the debt-to-exports ratio, the debt-to-government ratio. This study however used a different approach called the accounting approach which is recommended by HIPC. HIPC stands for Heavily Indebted Poor Countries and was formed by the International Monetary Fund (IMF) and World Bank to give debt relief to its member countries. The study covers the period 1980 to 2017. The study found that Zambia's current level of debt is unsustainable based on the accounting approach and that there is no difference between HIPC and IMF results.

Introduction

Zambia is a landlocked country situated in Southern Africa and it was formerly known as Northern Rhodesia during the time of colonialism (Pillay, 2002). The author further states that it shares its borders with Zimbabwe, Malawi, Democratic Republic of Congo, Tanzania, Namibia, Botswana, Mozambique and Angola. In addition, the country gained its independence in 1964 and since then has experienced structural changes in its political, social and economic systems. The country has established relationships with the rest of Africa and the world at large through its

[1] Department of Economics, University of Namibia

[2] Department of Economics, University of the Western Cape, Email: peyavali@gmail.com

membership in organisations such as the Common Market for Eastern and Southern Africa (COMESA), Southern African Development Community (SADC), World Trade Organisation (WTO), the African Development Bank (ADB), and the United Nations (UN) (Pillay, 2002).

Zambia is rich in copper especially in a region called the Copperbelt which is near its border with the Democratic Republic of Congo. According to Masengo, Cheela, Banda, Chileshe, Alexeev and Siamimwe (2013) in 1964 when the nation gained its independence it was led by Mr Kenneth Kaunda who became the first president of the Republic of Zambia and believed in socialism. At this time copper prices were high and the country was prospering. It was perceived that this would always be the case but today Zambia is among the poorest nations in Africa. Mid 1970's copper prices started falling and it was at this time that Zambia started borrowing (Masengo et al., 2013). Furthermore, the country borrowed heavily like it anticipated copper prices would escalate in the future.

More than 80% of Zambia's external debt was forgiven as the country reached the Heavily Indebted Poor Countries (HIPC) completion point in April 2005 (Sungwema and Odhiambo, 2018; Masengo et al., 2013). The total amount of debt forgiven was US $14.3 billion as at 2005. The country borrowed this money to finance its budget deficits and its main lenders were the International Monetary Fund (IMF), World Bank and the Paris Club. At that time these organisations gave for various purposes unlike today where they give for specific projects only. The country found it hard to pay back its debt as copper prices behaved contrary to its expectations.

Heavily Indebted Poor Countries (HIPC) classification was launched in September 1996 by the World Bank and the International Monetary Fund (IMF) (Canuto & Moghadam, 2010). Canuto and Moghadam (2010) further state that it was launched with an initiative to help the poor and heavily indebted countries and to reduce their debt burden by 90%. It created a framework for creditors to forgive such countries to allow them to use the funds for economic growth and poverty reduction. Canuto and Moghadam (2010) also further state that in 1999 HIPC was enhanced (and this marked the beginning of HIPC II which is still in use until today) to provide faster debt relief and strengthen the connection between debt relief, poverty reduction and social policies. It was also realised that the HIPC countries increased their poverty reduction expenditures due to the decline in debt service.

Debt sustainability can be defined as the capacity or ability of a country to service its debt without hindering economic growth (Development Finance International, n.d.). A study conducted by the IMF classified Zambia as a debt sustainable country (Zambia Country Debt Profile, 2015).

In the year 1986 Zambia spent about 86% of its total export earnings to service its outstanding debt. This means that, only 14% of what the country exported was left for local consumption (Zambia Country Debt Profile, 2015). In addition to this in 1999 the country used three times the funds allocated to health, education and social services to finance its debt. Later in 2000 the country's external debt to exports ratio was 652% and in 2006 the nation received debt relief under the Multilateral Debt Relief Initiative (MDRI) which reduced its external debt balance by 77.9%. At the end of 2006, the government acquired six new loans amounting to US $79.7 million dollars to finance its infrastructure, social service provision and poverty reduction projects.

Debt must be serviced and in most cases it is paid back with interest (Roubini, 2001). Arnone, Bandiera and Presbitero (2008) state that increased debt service soaks resources from a nation's government budget. They further state that if debt accumulates it may lead to a debt overhang. A debt overhang is situation where investment reduces because foreign creditors tax away returns and it becomes less profitable to invest as this was observed in Latin America in the 1980s. Zambia has been depending on both domestic and external debt to finance its budget deficits since the 1970s when copper prices started falling (Masengo et al., 2013). The nation has incurred a lot of debt and is now classified as a 'heavily indebted poor country'. Therefore, the study draws its primary interest from this to ascertain whether or not debt in sustainable in Zambia. The paper is organized as follows: the next section presents a literature review. Section 3 discusses the methodology. The empirical analysis and results are presented in section 4. Section 5 concludes the study.

Literature Review

According to Roubini (2001), assessing debt sustainability is not easy but the author suggests a number of methods that can be used, which include foreign debt-to-GDP ratio, debt service-to-GDP ratio, debt service-to-exports ratio, debt service to government revenue ratio debt-to-exports ratio and debt-to-government revenue ratio. Furthermore, the author recommends the debt-to-GDP, debt-to-exports and debt-to-government revenue despite their shortcoming. The author further cautioned that irrespective of the indicator used in analysing external debt, they are all affected by the exchange rates in the sense that when domestic currency depreciates the value of the external debt in relation to the specific indicator increases.

Debt sustainability analysis is a method of assessing how a country's existing debt level and prospective future debt affects its ability to service its debt in the future and it is important to analyse the risk of a country to debt distress (World Bank, 2006). Furthermore, debt burden indicators include debt service, debt stock and repayment capacity of the indebted country. These indicators are similar to those of Roubini (2001) in the sense that the study also suggests that variables like debt service, debt stock and GDP must be taken into account when assessing debt sustainability.

When a country faces challenges servicing its external debt, it is a sign of external debt burden (Kasidi & Said, 2013). They further state that this burden can be measured by the proportion of current income that is devoted to servicing external debt, the greater the proportion of current income devoted to servicing external debt the greater the external debt burden. The greater the external debt burden, the higher the level of unsustainability. They further state that a debt sustainability analysis is forward looking and a number of factors must be put into place to determine whether a country will service its debt. As in the case of Roubini (2001), they place emphasis on variables like GDP, exports, government revenue and debt service. Kasidi and Said (2013) also include factors like prospective path of deficits and associated debt service. They further state that alternatively a country can find what the optimal level of debt should be in order not to run into difficulty when servicing its debt.

Ekanayake (2011) suggests that an over borrowing hypothesis approach that involves calculating a benchmark ratio and comparing the current level of debt to this benchmark ratio can be used in assessing a nation's repayment capacity. If the current level exceeds the benchmark, it is considered as a situation of over borrowing. This means the nation is borrowing beyond its capacity. This approach is represented by the following equation:

$$d^* \cong ps/\,r - g = rev - pexs/\,r - g \tag{1}$$

where d^* is the benchmark debt level, *ps* is the constant primary surplus as a percentage of GDP expected in the future, *rev* is the average revenue ratio, *pexp* is the average primary expenditure ratio, *r* is the real interest rate and *g* is the economic growth.

According to International Monetary Fund (2013), debt is sustainable when the debtor is able to service the debt without affecting its income and expenditure balance. Furthermore, debt sustainability analysis is a reflection of a country's solvency liquidity and adjustment capacity. A country is solvent if the present value of its current and future income is in excess of its debt, it is liquid if it has control of its maturing debt and its adjustment capacity is determined by whether it can adjust spending and revenue. Just like Roubini (2001), World Bank (2006) and Kasidi and Said (2013), debt-to-GDP, debt-to-exports and debt-to-government revenue ratios are recommended in the IMF approach of analysing debt sustainability. However, IMF (2013) states that Debt Sustainability Framework (DSF) differs from the Debt Sustainability Analysis (DSA) under HIPC in the sense that the one under HIPC focuses on decreasing existing debt levels to substantial levels while the DSF focuses on determining the level of debt distress of a country. They further advise that for accuracy it is better to use present value to determine whether a country will face challenges in paying its debt in the future.

One of the approaches of determining debt sustainability is the accounting approach which measures the ability of a country to service current and future external debt (Arnone, Bandiera and Presbitero, 2008). This approach requires a country to find the level of primary surplus or deficits that stabilizes the debt-to-GDP ratio, a stable debt-to-GDP ratio indicates debt sustainability. The equation is given as $SURP = (r–g/1+g)b$, where $SURP$ is the government budget surplus or deficit, r is the interest rate, g is the growth rate of GDP and b is the debt-to-GDP ratio.

Burda and Wyplosz (2013) introduce the aspect of tax revenue. They state that a country's ability to repay a debt depends on its economic size. Given the same tax rate, a larger country will obtain greater tax revenue and have higher chances of repaying its debt than a country with a smaller economy. This is due to the fact that, in almost all world economies the main source of government revenue is taxation. Hence the stabilization of the debt-to-GDP ratio is a better tool for the debt sustainability analysis than the stabilization of debt itself.

There is voluminous empirical literature on the subject of debt sustainability. Roubini (2001) carried out a debt sustainability analysis in Argentina in 2000. The author used the debt-to-GDP, debt-to-exports and debt-to-government ratios. Using the debt-to-GDP ratio which was equal to 50%, the results showed that Argentina was found to be a solvent state. On the contrary, using the debt-to-exports ratio the country was found to be insolvent with over 400%. Similarly, using the debt-to-government the country was found be insolvent as the ratio lie in the range of 200% and 250%.

IMF and IDA (2002) conducted an analysis on debt sustainability in Uganda. They used debt and debt service indicators in relation to GDP, exports of goods and services and government revenue converted in to US dollars. They looked at how the economy performed in the 1990s and how this affected the debt indicators. Between 1990 and 1998 when the nation had strong macroeconomic performance it was able to maintain high levels of GDP, export earnings and government revenue as a percentage of GDP also increased. Debt was found to be sustainable during those years.

In Rwanda, IMF and IDA (2010) conducted debt sustainability analysis. The main debt indicators used in this analysis were the debt-to-revenue, debt-to-exports, debt-to-GDP, debt service-to-revenue and the debt service-to-exports. The debt-to-government revenue threshold (250) that was used is the same as the one suggested by HIPC. The debt service-to-exports threshold that was used is 20, 5%, higher than the one under HIPC. The debt service-to-revenue threshold that was used is 30%, also 5% higher than the one under HIPC. However, debt was found to be sustainable as the debt indicators were all below the IMF thresholds.

In Zambia, IMF and IDA (2015) conducted a debt sustainability analysis and they made future projections from 2015 to 2035 as well as analysing the level of debt sustainability as 2015. The external debt indicators used were the debt-to-GDP, debt-to-exports, debt-to-government revenue and debt service-to-government revenue. All the ratios were expected to lie below the thresholds except the debt service-to-government revenue ratio and hence debt was found to be sustainable in Zambia.

IMF and IDA (2016) conducted debt sustainability analysis for Kenya. In Kenya apart from the Central Bank and the government, the private sector was also borrowing externally. The risk of external debt distress was found to be low for the country using the IMF thresholds. The external debt indicators were also below the IMF thresholds and debt was found to be sustainable in Kenya. Additional empirical literature is in table 1 appendix 1.

There are a number of lessons to be learnt from the literature review above. From the theoretical literature review of debt sustainability analysis, it is clear that the most common methods of carrying out this analysis are the debt ratios. In addition to this, the debt-to-GDP ratio is crucial in analysing debt sustainability as it captures all the other variables (exports and government revenue). However, it is not enough to base conclusions on this indicator alone as it is vulnerable to external forces such as exchange rates and can change value over time. For accuracy it is essential to examine its movement over time and very few authors like Roubini (2001) and Arnone, Bandiera & Presbitero (2008) have pointed this out.

From the empirical review, first, most of the analysis done in most countries including Zambia has focussed on the debt ratios like the debt-to-government revenue, debt-to-exports and debt-to-GDP. Second, different ratios give different results and it is hard to make the definite conclusion of whether debt is sustainable in a country or not. In addition to this, the findings of the IMF and IDA done in Uganda in 2002 show that debt was found to be sustainable from the year 1990 to 1998 because the country had the ability to maintain high levels of GDP, exports and government revenue. It is therefore not enough to make conclusions on the current ratios alone but to actually examine the stability of these ratios over time especially the debt-to-GDP ratio. Furthermore, from the study done in Uganda an important lesson to note is that a country should not exclude economic growth in analysing debt sustainability.

From the work of Blominvest (2015), it can be noted that to analyse the level of debt sustainability of a country over time each year must be looked at individually. This means that the debt ratios are analysed yearly not over the entire period as a whole. This may be due to the fact that most countries borrow to finance their budget deficits and national budgets are prepared yearly. In addition to this, debt incurred in a certain year may reduce the following year as a nation may have serviced it or due to appreciation of exchange rates. Furthermore, as mentioned earlier on, before the IDA can give grants to countries yearly it demands knowledge of the country's current level of debt sustainability.

Overall, it can also be concluded that the accounting approach of HIPC has not been exhausted in many countries including Zambia. For accuracy a country should employ many techniques before making its final conclusion and should ensure debt sustainability analysis is conducted annually.

Methodology

Analytical Framework

This study followed an approach by Arnone, Bandiera & Presbitero (2008) called the accounting approach. The approach uses the equation which can be expressed as:

$$SURP = \{\frac{R-G}{1+G}\}b \tag{2}$$

where *SURP* is the government budget deficit or surplus needed to stabilise the debt-to-GDP ratio, *R* is the interest rate on debt, *G* is economic growth and *b* is the debt-to-GDP ratio. If Zambia's current budget deficit or surplus does not equal the one given by the equation, then the debt-to-GDP ratio will not stabilise over time and it will be concluded that debt is unsustainable. This is based on the analysis under HIPC that a fiscal deficit is sustainable if it generates a constant debt-to-GDP ratio.

Arnone, Bandiera & Presbitero (2008) state that an increase in economic growth increases the level of debt sustainability in a country and thus, it essential to include economic growth for an appropriate debt sustainability analysis. A country may currently have a low GDP, and debt may look unsustainable in relation to GDP but this GDP has the potential to grow and stabilise the debt-to-GDP ratio. In addition, an increase in GDP tends to improve the nation's budget deficit. Given the fact that Zambia borrows mainly due to budget deficits, equation (2) was deemed as most appropriate for analysing debt sustainability in Zambia.

Lastly, the same equation was used to compare IMF and HIPC results from 1970 to 2016. IMF debt sustainability analyses have already been conducted in Zambia so this research paper only focused on using the HIPC approach to obtain results under HIPC to compare year by year. According to the IMF, debt is sustainable in a country if the debt-to-GDP ratio is less than or equal to 50% but under HIPC it is not enough to make conclusions on this threshold. Therefore, this research paper also aimed at finding if at all the results differ.

Data and Data Sources

The sample range of 1980 to 2017 was used for analysis and this indicates a sample size of 37. The reason for this is that Zambia started borrowing heavily only in the 1970s and debt only became a big problem in the 1980s. The data was collected from the International Monetary Fund offices in Lusaka. The data was observed yearly from 1980 to 2017 to determine debt sustainability in each particular year.

Empirical Findings and Analysis

Firstly, the study established whether the budget deficit is equal to the government budget deficit or surplus needed to stabilise the debt-to-GDP ratio as generated by equation (2). According to IMF, the budget deficit was projected to be 7% of GDP, while the value generated (sustainable surplus) by equation (2) is 25.74. The estimated budget for 2017 was estimated to stand at U$ 1.5 million, hence, as of 2017 Zambia's current debt is unsustainable. Therefore, the country will not be able to service both its current and future external debt without compromising economic growth. Saungweme and Odhiambo (2018) categorically put it that the maturing of Eurobonds as well as the recent or newly contracted non-concessionary loans, Zambia's foreign debt service costs are likely to double by 2020. This is coupled with the rising international interest rates on debt. Thus, posing anticipated future challenges in debt repayment for the country.

Secondly, the study went further to compare the HIPC accounting approach of analysing debt sustainability to the IMF approach of analysing debt sustainability. The results for IMF showed that Zambia's debts is unsustainable. In this study, all the calculations are reported in table 2 in appendix 1. Under the IMF if the debt-to-GDP ratio is less than 50% then debt is sustainable. From table 1 in appendix 1 it can be observed that out of the 36 years, HIPC and IMF results were identical for 27 years. In other words the results further validate the hypothesis that Zambia's debts is unsustainable since there is no difference between HIPC and IMF results.

The findings of this study is also supported by recent activities by Zambia, as there is still no sign of slowing down in terms of borrowing. Particularly, the results for Debt Sustainability Analysis prepared by World Bank and IMF staff in revealed that in October 2017, Zambia was classified at 'high risk of debt distress'. This is due to rapid accumulation of debt resulting from non-concessional loans or financing. One such example is non-concessional loans from China as well as from the international debt markets. However, there are issues regarding the transparency of these loans due to less or limited public data. In addition, Zambia had other alternative sources from countries such as from India, the Gulf and other emerging markets. It is for this reason that Saungwema and Odhiambo (2018) states that public debt servicing challenges being experienced Zambia are as a result of economic and financial policy choices, global financial developments, and scanty foreign direct investment inflows.

Conclusion

The main objective of this study was to analyse whether or not debt is sustainable in Zambia while the specific objective was to compare HIPC and IMF approaches of analysing debt sustainability to determine whether or not they arrive at the same results. The main reason that this study was conducted was to help policy makers implement effective debt management strategies. It also looked at how other authors define debt sustainability and their approach of analysing a sustainable debt. Furthermore, it also looked into account the empirical finding of previous studies done in the same area in Zambia and other countries. The IMF conducted a debt sustainability analysis in Zambia in 2015 using the debt ratios and concluded that was sustainable in Zambia. This paper however used a different approach which is recommended by HIPC to analyse the sustainability of debt in Zambia.

It was found that debt is not sustainable in Zambia as at 2017 and that both HIPC and IMF approaches yield the same results. Furthermore, it was observed that HIPC has played a very important role in relieving Zambia from its debt because it was straight after the nation's debt was forgiven that the debt-to-gdp ratio fell tremendously. This shows that the management of debt in Zambia is reliant on foreign assistance hence the need for effective debt management strategies by policymakers within the country.

Burda and Wyplosz (2013) state that there are many techniques of achieving debt sustainability. One of these is a reduction in the interest rates at which countries borrow both home and away. This is not easy to achieve as interest rates are determined on the capital market and the government of a single country cannot influence the interest rate at which all other countries borrow. Secondly, the government can reduce its public spending but this however will be resisted by those directly affected civil servants, pensioners and public construction contractors. Going back to the birth of debt in Zambia, decline in copper prices and the country trying to diversify its economy through infrastructure development, there are certain measures it can take to suit the nature of its economy.

Based on the findings, the authors recommends the following. One policy the government should implement is the maximisation of tax revenue from the booming sector (mining) and use it to grow the lagging sectors such as the tourism sector, agricultural sector and entertainment industry. This will be a good investment given a scenario that minerals are completely depleted. Furthermore, the government can direct some of its infrastructure development projects to the mining companies instead of borrowing funds abroad to develop infrastructure in the nation. This is something the mines will view also as a benefit on their part as it will also attract investors to their location.

The government can also reduce its total spending without reducing provision of public goods and services because Zambia is a third world country and many people are living in absolute poverty. This can work if the government empowers young people to venture into entrepreneurship than relying on the government to create jobs. It can do this by mentoring and coaching young people. In addition to this, it can also help them grow their businesses through marketing assistance. Supporting entrepreneurial activities should become a government priority as it will reduce the amount of government revenue directed towards salaries and wages of government workers. These young entrepreneurs can also take up some of the government's responsibility in terms of public goods provision as Cooperate Social Responsibility (CSR). Furthermore, the government can further merge some ministries and cut down on the number of government positions as this will reduce expenses incurred to send government officials overseas for various meetings and offer various benefits to them.

References

Arnone, M., Bandiera, L., & Presbitero, A.F. (2008). *Debt Sustainability Framework in HIPCs: A Critical Assessment and Suggested Improvements.* Retrieved from http://utenti.dea.univpm.it/presbitero/pubblicazioni/Debt%20sustainability_rev.pdf

Blominvest Bank. (2015). *Debt Sustainability Analysis for Lebanon*. Retrieved from http://blog.blominvestbank.com/wp-content/uploads/2015/06/Debt-Sustainability-Analysis-for-Lebanon-2015-20201.pdf

Burda, M., & Wyplosz, C. (2013). *Macroeconomics: A European Text.* Oxford University Press.

Canuto, O., & Moghadam, R. (2010). *Heavily Indebted Poor Countries (HIPC) Initiative and Multilateral Debt Relief Initiative (MDRI) - Status of Implementation.* International Development Association and International Monetary Fund.

Development Finance International. (n.d.). *Debt Sustainability.* Retrieved from http://www.development-finance.org/en/topics-of-work/debt-strategy-information/debt-sustainability.html

Ekanayake, D.G.D.I. (2011). *Assessing Government Debt Sustainability in Sri-Lanka.* Retrieved from http://www.cbsl.gov.lk/.../_docs/IRC/2011/Govt.Debt

International Monetary Fund. (2013). *Public Sector Debt Statistics: Guide for Compilers and Users*. Washington DC, USA.

IMF., & IDA. (2002). *Updated Debt Sustainability Analysis and Assessment of Public External Management Capacity*. International Development Association and International Monetary Fund.

IMF., & IDA. (2010). *Heavily Indebted Poor Countries (HIPC) Initiative and Multilateral Debt Relief Initiative (MDRI) – Status of implementation*. International Development Association and International Monetary Fund.

IMF., & IDA. (2013). *Sixth Review Under the Three-year Arrangement Under the Extended Credit Facility – Debt Sustainability Analysis.* International Monetary Fund. 7U.

IMF., & IDA. (2015). *Consultation - Debt Sustainability Analysis*. International Monetary Fund.

IMF., & IDA. (2016). *Second Reviews Under the Standby Arrangement and the Arrangement Under the Standby Credit Facility, and Requests for New Twenty-Four Month Standby Arrangement, and a New Twenty-Four Month Arrangement Under the Standby Credit Facility – Debt Sustainability Analysis Update*. International Monetary Fund.

Kasidi, F., & Said, A.M. (2013). Impact of External Debt on Economic Growth: A Case Study of Tanzania. *Advances in Management & Applied Economics* 3(4): 59-82.

Masengo, P.C., Cheelo, C., Banda, B., Chileshe, M.P., Chipili, J., Alexeev, M., & Siaminwe, L. (2013). Issues on the Zambian Economy. *Bank of Zambia Reader*, 01(09): 15-17.

Pillay, P. (2002). *The Role of the State in Economic Development in Southern Africa*. Berlin, Germany. Dialogue of Globalisation.

Roubini, N. (2001). *Debt sustainability: How to Assess Whether a Country is Insolvent.* Stern School of Business, New York University.

Saungwema, T., & Odhiambo, N.M. (2018). An Analysis of Public Debt Servicing in Zambia: Trends, Reforms and Challenges. *Croatian International Relation Review* XXIV(81): 113-136.

World Bank. (2006). A Guide to Debt LIC Debt Sustainability Analysis. Retrieved from http://siteresources.worldbank.org/.../DSAGUIDE_EXT200610.pdf

Zambia Country Debt Profile. (2015). Retrieved from http://www.afrodad.org/phocadownload/publications/Country_Profiles/zambia.pdf.

Interdisciplinary Journal of Economics and Business Law, 2019, Vol. 8,
Issue 3, pp. 93-113

Managing the establishment of Business Continuity Planning by Greek SMEs: can it safeguard strategic viability of courier services during turbulent times?

Alexandros Garefalakis,[1] Kanellos Toudas,[2] Panagiotis Ballas[3], Spinthiropoulos Konstantinos,[4] and Athanasios Zisopoulos[5]

Abstract

The aim of Business Continuity Management (BCM) is to manage risks that impact on the normal operation of an organization maintaining enterprise and operational continuity. The contemporary business environment is so erratic and versatile too that companies should continuously cater for practices to eliminate the impact of risks on financial losses, reputational issues, natural hazards, and terrorist attacks, to name a few. On top of the above, globalization of operations accompanied by the holistic use of internet in commerce and transactions contributes to the emergence of new types of risks. Thus, top management, especially in SMEs, should carefully examine them through environmental scanning and evaluate their implications to the going-concern of the business. Adopting the case study approach, our study focuses on the services sector in Greece and explores the stance of a Greek courier company towards risk management and the need to adopt BCM practices as a plan to avoid disruption of its operations. In order to evaluate top and middle management readiness to respond and recover from unforeseen events, both top managers and employees participated to the data collection regarding company risks so as to determine whether a formal plan to overcome the aforementioned issues existed. The study fills a gap in existing literature on real-life case studies on business continuity planning in SMEs in Greece.

---

[1] Accounting and Finance Department, Technological Educational Institute of Crete, Greece. Email: agarefalakis@staff.teicrete.gr

[2] Economics Department, National and Kapodistrian University of Athens, Email: kstoudas@gmail.com

[3] Laboratory of Accounting and Financial Management, Accounting and Finance Department, Technological Educational Institute of Crete, Greece, Email: p.ballas@gmail.com

[4] Department of Business Administration, Technological Education Institute of Western Macedonia, Greece, Email: spinthiropoulos@gmail.com

[5] Applied Mathematics, Faculty of Mathematics and Informatics, University of Sofia "St. Kliment Ohridski", Sofia Bulgaria, Email: zisopoulos.athanasios.dim@gmail.com

Introduction

Everyday business environment undergoes major changes, many of which bear risks for the achievement of various targets set by an organization. Potential sources of risk may stem either from the internal (e.g. operations, personnel, power outages, equipment breakdown, IT malfunction, software errors) or the external environment (e.g. natural disasters, changes in laws/ regulations, cybercrime, terrorism, economic conditions) of an organization. Thus, companies should consider ways to safeguard themselves against major types of business risks, which may a) impact on the successful implementation of their strategies, b) force them to comply to certain laws and regulations, c) hinder their day-to-day operations and d) cause serious damage to their reputation. The numerous sources of risk accompanied by significant consequences and losses in monetary terms has led organizations to take proactive actions in order to sustain survival as a going concern entity.

In recent years, Risk Management constitutes a management tool of paramount importance due to multidimensional uncertainty conditions. Also, the concepts of risk, uncertainty and impact are critical in contemporary management practices. Elaborating on the relationship among these terms, risk refers to the variability of results/consequences and is the product of uncertainty and impact. Uncertainty depicts the ignorance of possible events that could lead to disruption of operations and the possibility of their occurrence. Impact shows the cost of these negative events. Various studies have focused on the negative implications to the organization of the lack of a proper plan against various disruptive events (e.g. Hiles, 2010; Losada et al., 2012).

There are many suggestions as to which Risk Management application is the best with Project Risk Analysis and Management (PRAM) (APM, 2004) and Shape, Harness, And Manage Project Uncertainty (SHAMPU) (Chapman and Ward, 2003) being among the most preferred ones. SHAMPU manages risk into three major steps; namely:

- Shaping sources uncertainty at a strategic level.
- Harness uncertainty by developing risk efficient tactical plans.
- Manage implementation process to eliminate risk.

Even though companies in the past developed disaster contingency recovery plans, Business Continuity Plans offer the advantage of safeguarding an organization before the undesired condition occurs (Cerullo and Cerullo, 2004). Business Continuity Management (BCM) is a vital component of enterprise risk management and its purpose is to sustain continuity. BCM is applicable both to for profit and non-profit organizations to prevent the negative implications of unforeseen events. Its goal is to prevent, if possible, or overcome disruption through quick recovery and through facilitation of long-term performance. Provided the advantages, it is surprising that companies fail to take continuity risks into account when performing their risk analysis. According to Maier (2005), who performed a study in the field and found that 25% of the 200 companies in the sample had not developed a BCP.

BCM aims at managing discontinuities caused by various events and incidents, while ensuring the continuity of operations. In this study we explore to what extent does a Greek SME have a positive stance towards the implementation of a BCP and whether managers' decision making takes it into consideration. Specific types of risks that modern corporations face are examined, such as financial losses, terrorism attacks, reputation crisis. In methodological terms, we employed a mix of questionnaires and face-to-face interviews conducted with top managers and employees to collect our data. The objective was to assess both top managers' as well as employees' awareness concerning various types of risks that the company faces. We also investigate whether the construction of a formal BCP could be regarded as helpful too to the prevention of a fatal risk from happening ensuring "business as usual".

This paper presents arguments for the successful application of BCM by a Greek SME since most of them have little understanding of what it involves. To that end, we analyse the relevance of BCP in a Greek SME making at the same time inferences regarding its usefulness as a tool that could contribute to strategic success. The focus of the paper is placed on whether top management of a Greek courier company exhibits awareness of BCM advantages. It was concluded that BCM was nonexistent in the company with the employees feeling that the implementation of the specific approach could act as a helpful tool towards effective risk management.

The originality of our study lies with the fact that there is limited body of research to investigate real cases of businesses with regards to their decision to focus on the development of continuity plans. Our aim is to focus on services and specifically on a company that offers courier services in Greece and explore managers' and employees' understanding of risks and whether BCP could offer advantages to an entity like that.

The paper is structured as follows: section 2 discusses relevant literature on risk management and BCM, section 3 presents the methodological approach followed, section 4 the analysis of data, and section 5 proceeds to the concluding remarks.

Literature Review

The daily operation of an organization in a global business environment is associated with dangers and risks, some of which could affect not only their day-to-day operations but more importantly their viability. That is the reason why contemporary management practices highlight the importance of risk management as a practice of utmost importance. Depending on the size of an organization, risk management is performed either by specialised personnel or a whole dedicated department. Also, taking advantage of business opportunities as well as the formation of business strategy involves the acceptance of a certain level of risk. However, both the internal and external business environment does not remain stable but undergoes continuous transformations. Thus, considering the existence of factors that impact on the stability of the entrepreneurship, the board of directors should review business strategy in light of these changes and proceed to the evaluation of the agreed strategic moves considering the impact on the flow of operations, on their effectiveness, on competitiveness of the entity and ultimately on the overall viability of the business. Globalization and evolution in information technology, political and social instability, and asymmetric threats together with climate change are pivotal factors that may hinder the implementation of business strategy. In many cases, changes to the environment are associated with alterations to strategic goals, which are associated with the consumption of valuable business resources (not only monetary).

In recent years, the economic crisis, the rise of terrorism, natural disasters and accidents have given cause to firms to search for ways to prepare for and anticipate the effects of multifaceted types of risks. The rationale of BCM is to provide companies with the foundations to avoid discontinuities and safeguard the continuity of operations and, ultimately, the entity. According to the Business Continuity Institute (BCI):

> *"Business Continuity Management is a holistic management process that identifies potential impacts that threaten an organisation and provides a framework for building resilience and the capability for an effective response that safeguards the interests of its key stakeholders, reputation, brand and value creating activities".* (BCI, 2007)

Further to the above and considering the above in the realm of internal audit and corporate governance, COSO (2004) had identified the importance and necessity of effective management of risks suggesting that organizations should adopt a culture to eliminate negative consequences of risks against the successful implementation of strategy and processes throughout the organization should be designed in a way to provide reasonable assurance towards risk management. The above depicts the need for experienced and competent risk management employees, whose duty would involve the formation of BCP tailored to the company's strategy and resources.

Business Continuity Planning

The underlying rationale for the BCP is that it operates proactively, forming plans on how the company should act to avoid discontinuities and how to overcome the problems that are likely to arise in case of a unfavourable event, while minimizing the impact due to potential losses keeping at the same time the firm operational within a particular timeframe. Hiles and Barnes (1999) illustrate the above:

> *"Business continuity planning is the identification and protection of critical business processes and resources required to maintain an acceptable level of business, protecting those resources and preparing procedures to ensure the survival of the organisation in times of business disruption".*

Elliot et al. (1999) recognised the importance of internal and external threats and proposed that business continuity plans should include both hard and soft assets in preventing risk and recovering from a serious disaster maintaining their competitive advantage. Moving a step further, Shaw and Harrald (2004) acknowledged the need for organizations being proactive towards risks and to the need for determined actions to prevent, mitigate, anticipate, respond to, recover and restore from a potential risk.

An effective BCP should comprehend main functions of a business including operations, production, finance, and IT since in the case of an unfavourable event in any of them there could be negative impacts on revenue generation, quality of customer services, and corporate image and reputation. Especially for IT, the need for the development of a recovery and a business continuity plan is of outmost importance (Myers, 1993). Predicting potential events that can lead to disruption, the goal of the BCP is to formulate plans to enable the firm to remain functional in case of events of emergency (Smith & Sherwood, 1995). What is important to highlight is the distinction between two types of plans that a company should develop; namely, a BCP and a Disaster Recovery Plan (Wunnava, 2011). Both of them differ in the time horizons they are carried out. The former is developed with an aim to restore operations at a minimum level to continue offering services/producing, whereas the latter is a plan for full recovery at prior-disaster level (ISO: 22310, 2012). In the early days of BCM, planning was mainly associated with risks stemming from information technology (Savage, 2002). Investigating BCM in supply chains, Burt et al. (2003) mentioned that maintaining a normal flow of inputs to the transformation/production process of an organization is a major source of risk.

The construction of a business continuity plan should include specific actions to identify, assess, rank and manage potential risks (Gilbert and Gips, 2000). What is also important though is the time to restore operations and return to normality after an unfavourable incident. The rationale for the development of a continuity plan is to a) identify risks, b) plan how to eliminate their impact on operations, c) communicate the plan into the organization among employees and test for its effectiveness (Cerullo and Cerullo, 2004). The necessity for risk management and planning lies at the fact that various types of risks may take place either individually or simultaneously; thus, the impact may be enforced in the latter case.

Another factor that affects BCP is the company culture. Embedding BCM in the organization culture should involve initially an understanding of the organization, which will dictate the different options to manage business continuity plan and in the end the plan should be reviewed, communicated and applied. Ward (2005) describes culture as a ‘conventional umbrella’ which protects and harvests behavioral and moral patterns. The culture of an organization affects the way employees behave within an organization and perform. When a new member of staff joins a company, he/she struggles to familiarize him/herself with the company culture, fit in and become an instrumental part of the entity (Botha J. and Rossouw, 2004). Risk management and BCP, in particular, can be greatly influenced by cultural traits embedded in a company either in a positive or a negative way.

BCP and BCM practices provide a variety of benefits to companies and, most importantly, create a sense of security to various stakeholders irrespective if they are customers, shareholders or employees, by creating a failsafe environment (Elliot et al., 2002). BCP is a tool to recognise and handle possible risks and leads to the development of an information flow system that facilitates the understanding of a business environment which can influence the extent to which a company can be competitive and its overall performance. In the event of an emergency, the existence of a plan to tackle the challenges that arise can be the lifesaver to overcome the difficulties in a prompt way and can contribute to a company’s survival in extreme circumstances. In the absence of such plan, the chances of organizational failure are increased; managerial decisions are made under pressure and there is no time to process all the data and parameters to reach an informed decision that will safely guide the company to safe grounds. It is also important to stress the advantage of training the staff on how to optimally react to a possible crisis since in that way they can be prepared and have enough time at their disposal to reach informed decisions to tackle an unfavorable event (Smith & Sherwood, 1995).

Even though BCM requires capital investment in the long run, it can lead to financial benefits for companies through risk anticipation and reduction of the negative effects that may incur (CCTAa, 1995). It also enables companies to minimize financial casualties by optimizing the response time to restart their operations during a crisis period providing shareholders a sense of security. BCP aims for the elimination of negative effects of an unanticipated event to the daily operation of a business. Zsidisin et al. (2005) focused on such planning on supply chains highlighting the negative impact of disruptions to meeting customer demand. Herbane (2010), acknowledging the economic importance and vulnerability of small businesses, urged future research to focus on the development of plans to ensure continuity. Moreover, bearing in mind the flexibility of business environment, the number and nature of risks change; thus, the BCP should undergo continuous evaluation and update when needed.

Business Continuity Cases

Acknowledging that theories should be applied and tested in real world cases, there is a number of studies on BCP in recent years, neither large nor extensive though. Combining strategic planning and cultural context, Perano et al. (2018) investigated BCP using a sample of 50 organizations from various sectors (i.e. banking, industrial and insurance sectors) in Albania and identified a positive relationship between strategic planning and BCP. Stark et al. (2016) focused on the impact of Ebola outbreak across West Africa in 2014 on business continuity highlighting the negative impact on logistic services and transportation of people and goods as a result of the adoption of specific travel and safety policies and regulations. Sarmiento et al. (2015) used a sample of 1200 entities based in America from 3 sectors (food and agriculture, tourism, construction), investigated their level of readiness to various types of risks specifically applicable to each of these sectors concluding: a) the lack of a business continuity plan for more than half of these organizations, b) the skepticism regarding the priority to develop a BCP, c) the increased need for small businesses to develop such plans compared to their large counterparts. Asgary et al. (2012) explored the impact of the 2010 flood in Pakistan on small SMEs and found that 90% returned to operations after six months but operated with loss due to the implications on personnel, supply chain, logistic services and damage to the infrastructures. Focusing on the banking sector of the United Arab Emirates, Randeree et al. (2012) developed a BCM maturity framework for banks.

The review of existing literature identified the gap for additional studies on the application of BCM in business entities, especially SMEs. Also, there is lack of research studies in the particular field to capture the Greek market. Hence, our aim to focus on the case of courier services in Greece is justified.

Methodology

The study adopts the case study approach (Yin, 1994) collecting not only quantitative, but also qualitative data via a series of interviews and a survey conducted within the particular courier company (Ghauri P. & Grønhaug K., 2002). The case selected is a Greek listed company, GR Courier Company SA, and the rationale is to explore the views of managers and employees on the value of a BCP to protect it against various types of risks and ensure continuity of its operations. The extent to which the company is capable of adopting such a plan is also examined. The analysis of data revealed fruitful results and insights into the topic of BCM allowing for propositions for future research (Eisenhardt, 1989).

Seeking to gain insight from all aspects of organization operations, two managers from the Sales department, one from the IT department, the International Business department along with a manager from the Sales Department were interviewed after giving their consent and after being presented with the questionnaires. Descriptive information was included in the questionnaires giving interviewees time to become familiar with the subject examined and also it was made possible for them to have access to necessary information and details necessary as our data collection evolved.

The reason why questionnaires was the preferred method to collect data is the fact that they could be easily and quickly distributed to the selected members of staff of the courier company either via email or by post (where an email address was not applicable) and, more importantly, contributed to the collection of a larger number of responses in comparison to interviews given time restrictions (Hussey and Collis, 2003). In order to get a deeper understanding of our topic and to sketch a complete picture of the particular case, we constructed a semi-structured questionnaire and proceeded to the implementation of a number of interviews with top management of the company in question. Moreover, we collected secondary data after analysing information provided in financial statements, management reports, company website and articles that provided information on the external to the company environment.

Analysis of the results

This section of the study presents and discusses the data collected. Bearing in mind the aim of the study, our emphasis is placed on understanding participants'/respondents' stance towards business continuity for the specific courier company and is followed by an analysis of data to pinpoint the advantages and disadvantages that originates from the application of a BC plan.

Analysis of the interviews

As described earlier, part of our data was collected via semi-structured interviews with managers from particular departments of the business; namely, Sales, Finance, IT, International Division and Greek Subsidiaries respectively. The selection of departments was made partly on the basis of their importance and deterministic role to the operational success of the entity, taking into consideration the possibility of suffering severe consequences in the event of an emergency/unfavourable event/risk. The interviews were conducted with the purpose to determine:

- The extent to which the company was aware of risks to its internal and external environment.
- How well informed the managers of the various departments were about these risks and of their potential implications.

- The overall attitude the company exhibited with regards to the identified risks.

The results showed that in spite of the satisfactory level of awareness of the relevant risks pertaining to their jurisdiction, departmental managers lacked an understanding of the greater picture on risks that affected the whole entity. Even though they could identify potential sources of danger, they were keen to take action in their department without cooperating with the rest of the departments in either the sharing of experience in identifying risks or in placing proper controls to prevent them from happening or even to cooperate in constructing an effective BCP to capture the whole entity.

An important finding from the discussion with respondents was the negative impact of outsourcing of the risk assessment process. Actually, the courier company under investigation depended for the review of risks and the development of an effective management plan partly on an external services provider. The only department that performed risk analysis and developed a crisis and recovery plan in the case of a disaster/ an emergency was IT. However, the absence of a risk management department within the company had the disadvantage that the actions taken to manage risks were not smoothly coordinated. Moreover, even though the company had not developed a BCP, all respondents agreed to the significance of performing risk analysis in-house.

Analysis of questionnaires and secondary data

In addition to the collection of data from interviews, we distributed a questionnaire for employees to fill in. Our objective was to determine how informed members of staff were about company risks and potential prevention matters in place. We collected 40 questionnaires, which is acceptable for the type and the size of the entity (Ritchie J. and Lewis J., 2006).

An important finding is that even though all respondents believed that the construction of a BCP would be beneficial for the going concern of the business, only 20% of them were aware of specific benefits of such planning. The same situation was justified when respondents were asked about risk management practices and 84% of them were confident there were such measures in place, but 89,3% of them was unable to specify any of them.

The following table ranks findings from the coding of questionnaires and the calculation of percentages presenting which risks were regarded as important by respondents and the recovery time in case such risk materialized sometime in the future. Respondents were asked to choose, according to their view, the most important risk factor and then depict the recovery time as a result of this risk.

Table 1. Ranking of risks relative to their importance and recovery time in months

	Risks	IT systems	Financial	Reputation	Natural hazards/ terrorism
Recovery time in months	> 24 months	0%	**72.6%**	12%	**72%**
	13-24 months	0%	20%	15.6%	11.3%
	7-12 months	6.7%	4.5%	**56.8%**	12.3%
	4-6 months	10%	0%	28.2%	3.2%
	< 4 months	**85%**	0%	2.3%	0%
Importance	80-100%	**72%**	**66.7%**	3.3%	0%
	60-80%	14.1%	21.2%	**68.5%**	2.3%
	40-60%	14.3%	3.1%	11.3%	14.7%
	20-40%	4.2%	3.3%	6.3%	21.3%
	0-20%	0%	2.4%	11.3%	**66.6%**

In line with previous studies, most of respondents associated risks with IT and Finance matters. Specifically, 72% of respondents consider risks relating to IT and Finance as being the most important ones with reputation risks following in third place. This is in line with Lemonakis et al. (2018) who asserted that a combination of financial soundness in the form of adequacy of equity together with the establishment of Enterprise Resource Planning tools are pivotal to the viability of Greek SMEs in a period of economic crisis. Natural disasters and terrorism appeared in the last place of ranking as 66% of the respondents failed to recognize them as major risk factors. The recovery time in case of an IT related risk is short and this is crucial, because the courier company is a technology intensive operation that is heavily dependent on information systems. The long recovery time after a financial risk striking the company should alarm the board of directors. The recovery time of over 2 years is worrying and requires further analysis on the impact to the ability of the company to cover its short- and long-term liabilities. This requires the cooperation among departments and, definitely, the Chief Financial Officer is responsible to construct a plan to protect the company's assets and safeguard the going concern of the business. Also, this is an area where the consulting side of internal audit could make a significant contribution. An additional dimension of risk identified was reputational damage, which was supported as being important by nearly 70% of respondents. Moreover, around 57% of respondents believe that in case of reputational damage, the company will struggle for a period that ranges between 7 to 12 months until its reputation regains the quality characteristics of the past. The last dimension of risk is identified as related to natural hazards and terrorism. This dimension was regarded as being less important compared to the previous ones, but it would require more than 24 months for operations to be restored in case of any risk from this category hit the courier company.

An important remark is that none of these categories of risk had occurred in the last two years prior to the data collection. This may have built respondents' confidence that nothing could have such an important implication as to bring operations to a halt. On the other side, the influence from both the global financial crisis in 2007-2008 together with the crisis in Greece definitely impacted in a negative way the sense of insecurity from the side of respondents; thus, they judged the financial dimension of risk as being very important.

Our data collection elaborated on the above asking respondents if they knew whether the main types of risks identified were managed in some way from within the entity. The table below exhibits both managers' and employees' beliefs regarding which risks were insured and whether each of these risks was managed internally or externally by third parties.

Table 2. Insurance measures relative to various types of risks

	Are those risks insured?			**Source of insurance**		
Types of risks	Yes	No	Do not know	Internally	Externally	Do not know
Financial losses	**60%**	17.3%	13.2%	30%	**39.3%**	27.2%
IT disruptions	**83.5%**	0%	5.6%	**38.6%**	33.3%	20.6%
Loss of personnel	**41.2%**	36%	19.7%	10%	**81.2%**	3.3%
Natural catastrophes (fire & earthquake)	**76.2%**	3.3%	23.4%	10%	**87%**	0%
Reputational loss	3.1%	19.9%	**64.8%**	**81.2%**	10%	6.1%
Strategic failures	12.3%	15.7%	**68.4%**	**74.3%**	6.6%	21.3%
Terrorism attacks	0%	91.6%	3.3%	0%	**97.1%**	3.3%

To sum up, the risks we identified as more important to the viability of the entity were the ones that threatened the going concern of the business and posed a significant obstacle to the achievement of strategic targets. However, managing risks in designing a BCP requires that even risks with potentially a low possibility of happening, should be tackled effectively due to their fatal consequences and disastrous effects. The fact that such a circumstance may have not occurred until the period of data collection, should not prevent companies from taking preventive actions through the implementation of a BCP. Identifying weaknesses in dealing with disruptions is of paramount importance for the entity considering the set-up costs after a serious risk delays operations. Once the above conditions have been met, the ground has been prepared for a BC plan to be realized.

Concluding remarks

It is essential for every organization to develop and adopt a BCP due to the numerous risks that have been present in recent years. Globalization, intense competition among companies and recessions, for instance, create an unstable business environment enforcing the need to take actions to safeguard their viability by establishing plans to tackle these challenges. BCM offers the foundations for it and could be used as a tool for top management in eliminating risk factors and their impact to the day-to-day operation of an entity. In an attempt to investigate the acceptance of BCM by SMEs in Greece, this study presented the shortcomings that stem from the negative attitude exhibited by companies to adopt such tools that will enable them to sustain continuity and manage risks effectively. Although the leading company examined in this article has made notable attempts to manage its risks, it is far from reaching the standards necessary in order to develop a satisfactory continuity plan to protect its operations effectively. The establishment of an effective BCP can be facilitated through the analysis of the information accumulated with regards to the existing circumstances of the Greek company and environmental scanning. The aim of the study was to identify the need that exists to form a BCP, the extent to which such a plan could be applied to a specific company and the advantages that the company can enjoy through its adoption.

The analysis of our data revealed that the courier company should adopt a BCP and both managers and personnel agreed with this. Operating in a highly competitive market with many competitors having already made steps to form continuity plans, the company of interest operates within a business framework where such plans gradually become a necessity.

Additionally, its strong cash flow position can ensure that the development of such plan can be established. Further to the above, efficient cash management could lead to saving of financial resources since the company could save on insurance costs and invest in a BCP. An advantage of developing a continuity plan internally could save costs offering at the same time the required flexibility and adaptation to future changes; thus, making it more effective. In order to achieve the above, executive recruitment and selection is considered of vital importance since staff competence is pivotal for the company's success. The courier company has already proceeded with the reorganization of its management structure to optimise company effectiveness. It can be concluded that the company has the foundations to develop a BCP internally offering at the same time the training required to its executives.

However, some weaknesses and flaws in the particular courier company, pinpoint the need for the establishment of a BCP. Firstly, discussion with top managers revealed the absence of competent employees to be able to construct an effective BCP. That is why risk management practices are initiated by a third party. Further to the above, dimensions and sources of risk should incorporate the whole spectrum of the entity. Even though specific functions within an organization may have a major impact, a BCP should involve an evaluation of potential dimensions of risk that stem from all parts of the business. In addition to the above, internal processes should be re-examined to clarify the line of responsibility for the various types of risks that affect the entity. Moreover, our data revealed another issue applicable to many SMEs with regards to preventive measures against risks; namely, insufficient and in some cases non-existent training of employees to manage operations in cases of emergency. The explanation for this is from one side due to existing business culture with regards to risk management and preventive measures, whereas on the other SMEs and the particular courier company should utilise their resources wisely – especially during a prolonged crisis period – to establish an effective BCP. Besides, this is in line with previous literature that accepts as a critical point to the success of a BCP the support offered by senior management given the limitations that stem from resources and strategic planning (Cervone, 2017). In addition to the above, organizations should always consider that risk management and the subsequent development of a business continuity plan is not a one-off event. It requires continuous evaluation of the business environment (both internal and external) and commensurate adjustments to BCP to avoid the latter becoming outdated. Towards this direction, the effectiveness of Internal Audit practices through the evaluation of the relevant Internal Controls (e.g. Lemonakis et al., 2018) as well as Corporate Governance policies (e.g. Ballas et al., 2018) could have a positive contribution to make.

What is important to point out is the difficulty in capturing all potential facets of risk an organization may face in the future accompanied by the possibility of them actually taking place; thus, there will always be an inherent discrepancy between BCP and reality. As a potential avenue for future research, we strongly support the need for embodying theories of the business with business continuity such as the resource-based view together with the resource allocation in case of a disaster together with complexity theory. In addition to the above, linking economic sustainability and environmental protection (in the form of law, regulations, directives) with business continuity management aiming for a holistic conceptual framework could prove a fruitful avenue for further research (Sariannidis et al., 2018). Also, information provided in the Management Commentary Reports could provide pivotal information on the above, which could be further be utilised in the identification of sources of risk (Garefalakis et al., 2018).

Finally, our study is not without limitations. Since we adopted the case study method, our findings refer to the particular entity and should be carefully scrutinised with regards to their applicability to companies other than this specific sector. In other words, the generalisation of results should be made carefully and definitely after the validation of the findings to other business settings.

References

Asgary, A., Anjum, M.I., and Azimi, N. (2012). Disaster recovery and business continuity after the 2010 flood in Pakistan: Case of small businesses. *International Journal of Disaster Risk Reduction*, *2*, 46-56.

Association for Project Management (APM) (2004). Project Risk Analysis & Management (PRAM) Guide (second edition), APM Publishing, High Wycombe, Bucks UK.

Ballas, P., Garefalakis, A., Lemonakis, C., and Balla, V. (2018). Investigating the effects of the implementation of IFRS and Corporate Governance practices on the quality of financial reports in a crisis era: evidence from the Greek Banking sector. *Interdisciplinary Journal of Economics and Business Law*. Forthcoming.

BCI (The Business Continuity Institute) (2007). 'Good Practice Guidelines 2007: *A Management Guide to Implementing Global Good Practice in Business Continuity Management*', Version 2007.2 (15th of March 2007).

Botha, J. & Rossouw, Von S. (2004). A cyclic approach to Business Continuity Planning', *Information Management and Computer Security*, 12 (4): 328-337.

Burt, D.N., Dobler, D.W. and Starling, S.S., (2003). "World Class Supply Management: The Key to Supply Chain Management", 7th edition, McGraw-Hill: New York, NY.

CCTAa (1995). An introduction to business continuity management. The government centre for information systems, IT infrastructure Library, HMSO: London.

Cerullo, V., and Cerullo, M.J. (2004). Business continuity planning: a comprehensive approach. *Information Systems Management*, 21(3): 70-78.

Cervone, H.F. (2017). Disaster recovery planning and business continuity for informaticians. *Digital Library Perspectives*, 33(2): 78-81.

Chapman, C., Ward, S., (2003). Project Risk Management: Processes, Techniques and Insights, Second Edition. John Wiley & Sons.

Chatsivasileiadou. S. (2007). Business Continuity Planning in the Greek Business Environment. *Library Management University of Southampton*, 46-50.

Committee of Sponsoring Organizations of the Treadway Commission (COSO) (2004). Enterprise Risk Management – Integrated Framework.

Eisenhardt, K.M., (1989). Building Theories from Case Study Research. *Academy of Management Review*, 14(4): 532–550.

Elliot D., Swartz E. and Herbane B. (2002). "Business Continuity Management: A crisis management approach", Routledge: London.

Elliott, D., Swartz, E. and Herbane, B. (1999). Just waiting for the next big bang: business continuity planning in the UK finance sector. *Journal of Applied Management Studies*, 8(1): 43–60.

Garefalakis, A., Ballas, P., and Dimitras, A. (2018). Determinant factors of the Quality Management Commentary Reports, in Zopounidis, C., Christopoulos, A., and Kalantonis, P. (Eds.), *Perspectives, Trends, and Applications in Corporate Finance and Accounting*, IGI Global (pp. 301-315), June 2018 [doi: 10.4018/978-1-5225-6114-9].

Ghauri, P. and Grønhaug, K. (2002). "Research methods in business studies: a practical guide", 2nd edn., Pearson Education Limited: Essex, UK.

Gilbert, G.A. and Gips, M.A. (2000) Supply Side Contingency Planning. *Security Management*, 44(3): 70–73.

Herbane, B. (2010). Small business research: Time for a crisis-based view. *International Small Business Journal*, 28(1): 43-64.

Hiles, A. and Barnes, P. (1999). 'The definite handbook of Business continuity management', John Wiley & sons Ltd: West Sussex, UK.

Hiles, A. (2010). *The definitive handbook of business continuity management*, John Wiley & sons Ltd: West Sussex, UK.

Hussey, R. and Collis, J. (2003). "Business Research: a practical guide for undergraduate and postgraduate students", 2nd edn., Palgrave MacMillan: Basingstoke, UK.

ISO:22310. (2012). Societal security – Business continuity management systems – Requirements. Terms and definitions. In *Terms and definitions*. Switzerland: International Organization for Standardization.

Lemonakis, C., Ballas, P., Balla, V., and Garefalakis, A. (2018). Audit fees and pricing strategy: Do restatements of internal control reports and earnings matter? *Risk Governance and Control: Financial Markets and Institutions*, 8(2): 63-73. [http://doi.org/10.22495/rgcv8i2p4]

Losada, C., Scaparra, M.P., and O'Hanley, J.R. (2012). Optimizing system resilience: a facility protection model with recovery time. *European Journal of Operational Research*, 217, 519-530.

Maier, S. (2005). ‘Corporate Governance Ranking: How global is corporate governance?’ EIRIS August 2005 (accessed online 5/2/2008: http://www.eiris.org/files/research%20publications/howglobalisgoodcorpgov0 5.pdf).

Malliara, N. (2000). ‘Risk management is gaining significant importance in developing the structure and the operation of banks’, *To Vima*, 24/12/2000, pp. D06, No. 13152 (accessed online 6/2/2008: http://tovima.dolnet.gr/print_article.php?e=B&f=13152&m=D06&aa=1).

Myers, K.N. (1993). “Total Contingency Planning for Disasters: Managing Risk, Minimizing Loss, Ensuring Business Continuity”, John Wiley & Sons, New York, NY.

Perano, M., Hysa, X., and Calabrese, M. (2018). Strategic Planning, Cultural Context, and Business Continuity Management: Business Cases in the City of Shkoder. In *Geopolitics and Strategic Management in the Global Economy*, 57-77. IGI Global.

Randeree, K., Mahal, A., and Narwani, A. (2012). A business continuity management maturity model for the UAE banking sector. *Business Process Management Journal*, 18(3): 472-492.

Ritchie, J. & Lewis, J. (2006). “Qualitative research practice: a guide for social science students and researchers’”, SAGE Publications Ltd: London, UK.

Sariannidis, N., Garefalakis, A., Ballas, P., and Grigoriou, E. (2018). Eco-Efficiency, Sustainable Development and Environmental Accounting in the Tourism Industry During a Crisis. *Corporate Board: Role, Duties and Composition*, 14(3): 58-64 [http://doi.org/10.22495/cbv14i3art5].

Sarmiento, J.P., Hoberman, G., Ilcheva, M., Asgary, A., Majano, A.M., Poggione, S., and Duran, L.R. (2015). Private sector and disaster risk reduction: The cases of Bogota, Miami, Kingston, San Jose, Santiago, and Vancouver. *International Journal of Disaster Risk Reduction*, 14: 225-237.

Savage, M. (2002). Business continuity planning. *Work Study*, 51(5): 254-261.

Shaw, G.L. and Harrald, J.R., (2004). Identification of the core competencies required of executive level business crisis and continuity managers. *Journal of Homeland Security and Emergency Management*, 1 (1): 1-14.

Smith, M. and Sherwood, K. (1995). Business Continuity Planning, *Computers and Security*, 14: 14-23.

Stark, C.G., and Willemse, A. (2016). Maintaining Business Continuity During the Ebola Crisis. In *SPE International Conference and Exhibition on Health, Safety, Security, Environment, and Social Responsibility*. Society of Petroleum Engineers.

Ward, S. (2005). Risk management organisation and context', 1st edn. Witherby & Co. Ltd: London, UK.

Wunnava, S. (2011). *Application of protection motivation theory to study the factors that influence disaster recovery planning: An empirical investigation*. Louisiana Tech University.

Yin, R.K., (1994). "Case Study Research: Design and Methods", 2nd edition. Sage Publications: Thousand Oaks, CA.

Zsidisin, G.A., Melnyk, S.A., & Ragatz, G.L. (2005). An institutional theory perspective of business continuity planning for purchasing and supply management. *International Journal of Production Research*, *43* (16): 3401-342.

Interdisciplinary Journal of Economics and Business Law, 2019, Vol. 8,
Issue 3, pp. 114-136

Competition, Consolidation, Collusion, Constraint

Kenneth S. Friedman[1]

The Example

Consider a society with 1,000 agents. Each is economically rational. Each is endowed with perfect knowledge. Of these agents, 999 have a net worth of €1,000 each. The thousandth has a net worth of €1,000,001.

A virulent plague strikes the society, infecting every agent. Untreated, it has a 100% mortality rate. The only cure is a naturally occurring substance that grows in plentiful supply, but only on the property of one of the agents.

The agent, being economically rational, seeks to price the cure to generate the maximum profit [which equals the maximum revenue, given the near-zero cost of obtaining the substance]. The economically rational price is €1,000,001. In this scenario, the remaining 998 agents die. Presumably, long-term GDP suffers as well.

Now add a minor wrinkle. The substance grows on properties of a few of the non-wealthy. These few agents seek the strategy that will maximize their profits. That optimal strategy is to collude, charge €1,000,001 for the cure, and split the profits. The remaining agents (who have neither the €1,000,001 nor the substance growing on their property) die. Presumably, GDP suffers.

The significance of this example stems from its providing those conditions – an endowment of economic rationality plus perfect knowledge – sufficient for an economic maximum, according to orthodox neo-classical economic theory. But despite these endowments, the reality falls short.

[1] Regis University, Denver, Colorado.

This is, of course, an extreme example, perfect knowledge and economic rationality (in this case, collusion among oligopolies) leading to social and economic disaster. Yet it provides a perspective that contrasts to orthodox economic theory.

Orthodox economic theory begins – and ends – with the paradigm of perfect competition, the realm of the invisible hand. If there were reason to believe that competition is the rule and that collusion is a rare exception, one might regard the orthodox theory as at least a good approximation.

However, the opposite appears to be true of most modern economies. A sufficient size advantage gives any one agent an advantage over others. There are typically economies of scale. With size there may be a budget for research and development, for specialization, for market research, for advertising, for training. With size comes greater consumer recognition. With size may come access to less expensive financing. With size there may be greater flexibility and a better ability to withstand localized adverse events. With size comes an ability to buy out smaller competitors and become even larger. [the]

The classical or neo-classical paradigm of a level playing field with many equal competitors is not stable, but rather meta-stable. If one starts out with such an economy, differences in relative size may be regarded as random fluctuations – up to a point. However, once differences exceed that tipping, or bifurcation, point, new dynamics come into play, dynamics that favor size. The uniform playing field naturally consolidates into one of oligopolies.

Even in finance, larger college endowments regularly outperform smaller ones (Piketty, 2013). Differences in size naturally tend to become greater over time.

This explains a widely-recognized economic phenomenon. Nearly every sector of the economy is dominated by a handful of global giants. What, other than macro-forces favoring size, could explain this palpable regularity?

The success of an explanation in terms of advantages that accrue to size creates a difficulty for orthodox theory. For it calls into question one of its essential paradigms. Orthodox economic theory may generate neat models based on economic rationality, perfect (or at least symmetric) knowledge, and (implicitly assumed) pure competition. Reality, however, is messy, and very different, especially with respect to competition vs. collusion. The Example may not be so unrealistic. Oligopolies are natural. Under a wide range of conditions, they are more stable than is perfect competition. As competition (including indirect competition stemming from substitution) lessens, size advantage generates positive feedback which increases that size advantage, reducing the number of players, magnifying the advantages of oligopoly.

However, while oligopolistic policies may be economically rational and good for the oligopolies, they are bad for economies. It is not just The Example. There is substantial evidence to the effect that mergers and acquisitions typically lead to higher prices and poorer quality.

One of the significant advantages of oligopoly is rent-capture that stems from collusion. As practiced in contemporary economies, collusion is often implicit and passive, one company raising a price to see if the few others will follow, prepared to rescind the increase if they do not. The few others quickly learn that it is more profitable to follow with their own price increases than to attempt to gain market share at the risk of a price war. It is economically rational to focus on keeping prices higher than they would otherwise be.

> '*Mehrotra (the CEO of Micron Technology) dismisses concern that the industry might flood the market with chips… He says his handful of rivals, such as Samsung Electronics Co., are squarely focused on profit, not market share.*' (Bloomberg Businessweek, Dec. 19, 2018 – Jan. 6, 2019, p. 30)

This has a ratcheting effect on prices – increasing in upward phases of the business cycle, stable in downward phases.

The economy naturally trends toward The Example – different oligopolies colluding and seeking to extract rents. In such cases, in which the invisible hand has a limited range of motion, or is completely paralyzed, it is essential to provide supplemental constraints.

Leviathan

For Hobbes, the central role of government is to provide a framework of justice and to maintain civic order. Today, a government unable to perform these roles is deemed a failed government.

Yet as economies have become less rudimentary and societies more complex, governments have filled additional roles. Most national governments issue and control their own currency. They have central banks that manage interest rates and may serve as domestic lenders of last resort. They provide social safety nets and infrastructure, from education to transportation. Many even impose industrial policies, seeking (often successfully) to pick and nurture winning sectors.

There is a school of thought, libertarianism, that opposes this, that would roll back government toward its Hobbesian minimum. Superficially, the theory sounds nice, if utopian. But it may be asking the wrong questions. The alternative to intervention is not free and fair competition, but exploitation through the collusion of oligopolies to extract wealth, an economy of The Example. There is significance in the fact that *every* industrialized country has (and may need) a non-libertarian government.

One should ask *how*, rather than *whether*, government should intervene. Over centuries, those countries that have forged economic miracles: Great Britain of the Industrial Revolution; the U.S. under the influence of Alexander Hamilton; Bismarck's Germany under the influence of Friedrich List; post World War II Japan; the East Asian Tigers; the Peoples' Republic of China; have done so on the back of active interventionist, dirigiste, government, imposing tariffs and implementing industrial policies.

Generalizing from many examples, Acemoglu and Robinson (2012) conclude that strong (if pluralistic) central government is necessary to national success. A careful reading of U.S. history provides insight into how specific government interventions aided, and may have been necessary to, the economic – and democratic – success of the country (Cohen and DeLong, 2016).

Questions as to what sorts of intervention are appropriate and functional are important, if difficult. An insistence that *any* intervention is inappropriate both flies in the face of history and clouds the essential issue of distinguishing intelligent interventions from foolish ones.

Government intervention is widely accepted as not only desirable, but necessary, in a range of areas. Certain general features are characteristic of these areas.

1. One area in which government intervention (to the extent of provision of goods and services) is readily justified involves the provision of institutions that constitute a foundation essential to society. Economic considerations here are tertiary, at best.

A judicial system which awards verdicts to the highest bidder is not a system of justice. Without a system of justice there is no functional notion of private property, a notion that may be necessary to a functional economy. In a similar vein, a national defense force that provides its allegiance to the highest bidder is incompatible with national security. Failure or inability to intervene in this sector is associated with failed government. Failed government has been closely linked to failed economies.

2. A second area meriting government intervention – again, to the extent of provision of goods and services – is where externalities figure prominently, especially where immediate economic benefit is difficult to calculate. Pigovian considerations are necessary.

In the context of positive (even rival and excludable) externalities, a society benefits from having a literate and numerate citizenry. But immediate economic benefits of widespread literacy and numeracy are often difficult to discern, much less measure. And attempts to increase productivity of educators, typically by increasing class size, are often counter-productive. (Though many education institutions could be made more efficient by reducing the number of administrators.) Education is commonly provided by government. Even in the U.S. most primary and secondary education is public. And many of its recently-formed private colleges have skated close to the edge of fraud in their pursuit of profits at the expense of students. Immediate profitability is not an appropriate gauge of long-term broad-based benefit.

Similarly, a society may benefit from an efficient system of roads, bridges, public transportation, even if the benefits are too diffuse to enable private corporations to capture enough to justify investment. It benefits from traffic lights, though there may be no direct way for makers of traffic lights to recoup their costs. New technologies, from pharmaceuticals to semiconductors, may fit this category. Health care delivery also fits into this area, and its obsession with privatization may explain why the U.S. lags in many measures, including life expectancy, despite spending so much more per capita than do countries that produce more favorable outcomes. In the context of negative externalities, government can regard these as endangerment and criminalize them (as in dumping toxic waste into streams). Or government can merely attempt to discourage or limit them by taxing them (as in a carbon tax or a tax on tobacco).

3. A third area, and the focus of this paper, also justifies government intervention, if only as regulator. Government acts not to replace, but to constrain, the private sector in a particular area – oligopoly pricing that distort the market and the economy.

There is a range of measures that have been available to and widely used by governments. An examination of these, their benefits as well as their limitations (and risks), may provide insight in an area that is economically important, but below the radar of orthodox theory.

Interventions

Intervention$_1$ – Prevention

If there are no oligopolies/monopolies, there is no oligopolistic/ monopolistic collusion and pricing. That is the rationale for anti-trust legislation, common in industrialized countries. Short of criminalizing collusion (which may be difficult to prosecute because it is difficult to identify intent, especially in passive collusion that merely goes along with others' price increases), government could seek to prevent companies from growing to the size and dominance that facilitates collusion or monopolistic rent-capture.

This may be appropriate in certain circumstances. Yet oligopolistic pricing is not the only advantage to size. Economies of scale are often real. This is most pronounced in the area of natural monopolies, an area that is wider than that suggested by orthodox economic theory. Breaking up such monopolies results in unnecessary duplication of effort and a loss of net efficiency, to the detriment of the economy.

Intervention$_2$ – Subsidies

In The Example, if government had sufficient resources, it could enable the survival of the 998 citizens otherwise doomed to perish by subsidizing the cure enough to change the pricing incentive. If there are government resources, N (> €1), with government providing €N/1,000 to each agent, it would be economically rational to price the cure at €(1,000 + N/1,000).

This resolves one problem, but raises another. While all agents would survive in this scenario, the subsidy makes it economically rational to price the cure to transfer all the assets of government to the owner(s) of the cure. This may be an extreme example. But it illustrates the tendency of subsidies to raise prices, to transfer assets from government to oligopolies/monopolies.

A striking example of this is portrayed by government and private insurance subsidizing pharmaceutical expenses. In 2015 Turing Pharmaceuticals acquired the rights to Daraprim, a drug treating life-threatening parasitic infections. The drug, originally discovered in 1952, had previously sold for $13.50 per tablet. Turing promptly raised the price to $750 per tablet.

In 2016 Valeant Pharmaceuticals acquired the rights to the cardiac drug, Isuprel, which had first been approved in 1947. It promptly raised the price from $440 per dose to $2,700 per dose. Similarly, Rodeles Therapeutics increased the price of cycloserine, discovered in 1954 and used to treat drug-resistant tuberculosis, from $17 per pill to $360 per pill. Sun Pharmaceutical and Mylan increased the price of albuterol tablets, an asthma prescription, from 13¢ to $4.70 per pill.

These examples are the tip of an iceberg. In their egregious move toward economically rational pricing, they make the somewhat less extravagant price increases demanded by major pharmaceutical companies seem restrained in context. This is partly responsible for the rise in costs of U.S. health care delivery significantly outpacing the rise in inflation, for the U.S. lagging countries that spend far less per capita in health care outcomes (World Health Organization). In effect, well-intentioned policies have encouraged rent-seeking and enhanced rent-capture.

Intervention$_3$ – Price Controls

This form of intervention is more controversial. In intervention$_2$ government appears to be no more than a generous uncle, giving money away. The source of the money tends to be ignored. In intervention$_3$ government is interfering with an agent trying to make a profit. Moreover, there is a standard argument (Hayek, 1948) that governments do not have the information necessary to rationally price a good or service. On this view, attempts to arbitrarily set prices inevitably send the wrong signals to consumers, producers, investors.

This argument, though widely accepted, carries diminished weight because monopolistic or oligopolistic pricing, which is far more common than are purely competitive markets, sets a low bar. The Example illustrates this. There may be no single 'right' price for the cure, and an ideal price would be difficult to determine. (Given that the agent played no role in the development of the cure, a minimal price might be justified.) In any case, it may be easy to find a more reasonable price than a monopolistic rent. €1,000,001 is both morally problematic and economically dysfunctional.

Moreover, governments have been fairly successful in setting reasonable prices where there are natural monopolies. They have sought to enable producers to generate returns on investment sufficient to encourage meaningful spending on research, development, and maintenance, while providing a measure of protection to consumers. (This does not deny the potential problems with regulatory capture, problems that exist with any set of regulations that could impact profitability. Yet the potential for regulatory capture provides a general argument for insulating regulatory bodies from financial and political pressure and from outside agendas, and not a specific argument against price controls.)

Intervention$_4$ – Nationalization

This sounds like a nuclear option of dubious wisdom. For one thing, nationalized industries, not driven by profit maximization, are not subject to market disciplines. For another, they appear to be subject to a range of inefficiencies associated with regulatory capture.

However, the results do not match the cautionary warnings of orthodox economic theory. Most countries have nationalized their postal services, with few signs of gross inefficiency, corruption, or overt political meddling. Many have nationalized their health insurance, also with few signs of gross inefficiency, corruption, or meddling.

In the U.S., Medicare's overhead expenses (salaries, marketing expenses, profits) are widely estimated at about 2% of revenue, in contrast to estimates in the mid-teens for private insurance companies. U.S. Social Security's administrative costs are 1% of revenues, one-tenth of the average administrative cost of private plans. Électricité de France, largely owned by the French government, is one of the most efficient electric utilities.

The notion that profitability is the only performance-enhancing incentive is grossly inadequate. The contrasting notion of a civil service, a trust motivated by a sense of value creation and duty to one's fellows, has been effective for centuries. A variety of incentives, some influenced by social expectations, may provide constructive motivation.

It is plausible that where innovation is important, private competition may outperform nationalized industries. But even here, government grants to researchers in universities have been potent tools. Selective, cautious, nationalization should not be summarily rejected.

Intervention$_5$ – Reducing Disparity

In The Example, suppose government were to redistribute wealth to reduce disparity. Consider a one-time tax of 50% of wealth, the proceeds to be distributed evenly. Then 999 agents will have a net wealth of roughly €1,500 while the remaining one has a wealth of roughly €500,500. Given this distribution, an economically rational price for the cure is €1,500. All agents survive. (This resembles intervention$_2$ in that the owner of the cure still extracts monopolistic rents. But it resembles intervention$_3$ in that it is revenue neutral for the government.)

Over the past century governments have taken steps to reduce economic disparity, through progressive tax codes, social safety nets, minimum wages. In many developed countries post-intervention GINI ratios are substantially [less] lower than their pre-intervention ratios. While their motivation has typically been unrelated to constraining oligopolist rents (Scheve and Stasavage, 2016), a lower disparity (among other benefits) may help a populace cope with oligopolistic rent-capture.

Yet the issue of redistribution appears to strike a nerve among economists. It is often regarded as the primal sin of interventionists. Surprisingly, the reasons adduced to support the claim that redistribution is necessarily counterproductive are strikingly feeble.

There is a trade-off between equality and efficiency (Okun, 1975).

Arthur Okun, though he was dismayed at the high level of American inequality in the mid-1970s (it is much higher now), was nevertheless the progenitor of this claim and a number of the reasons below. Is the notion of such a trade-off plausible?

Consider an extreme case in which one agent owns all the wealth and earns all the income, while the others are all penniless and starving. This hardly makes for efficiency. It is clear that such a trade-off, while it might exist for some distributions, does not exist for all. What about less extreme distributions?

An initial perusal suggests caution. The most efficient countries, in terms of GDP per hour worked or GDP per unit of energy consumed, reflected in high standards of living, are disproportionately the most egalitarian countries. For the industrialized countries, as economic disparity has increased over the past four decades, economic and productivity growth have declined sharply.

In the U.S., the previous peak in economic disparity was in the late 1920s, preceding the Great Depression. The subsequent Great Compression (of economic disparity), a result of the policies of the New Deal (including a highly progressive tax code), heralded a decades-long improvement in economic growth.

The flaws in the following reasons serve to explain why there is less of a trade-off between equality and efficiency and more a trade-on, why (at least at present levels) increased equality is associated with greater efficiency.

Greater economic disparity provides greater incentive to be productive.

Might there be a diminishing marginal utility to wealth? Does €1 billion provide much more incentive to be productive than does €1 million? Would splitting the €1 billion into 1,000 packets of €1 million each provide a greater total incentive?

In a national vein, are Americans more, rather than less, hard-working today than they were in the 1940s and 1950s (a generation still referred to as 'The Greatest Generation'), when tax rates were more progressive (with highest marginal rates above 90%), when economic disparity was less and economic mobility was greater? Might economic mobility itself (which plausibly increases as disparity decreases), a more realistic opportunity to improve one's relative economic status, but also a more realistic opportunity to see it decline, enhance incentive to be productive?

More recently, policies adopted by a number of countries in the late 1900s appeared designed to increase economic disparity by lowering taxes on the wealthy at the expense of social safety nets. Economic orthodoxy might expect the heightened disparity to increase incentive and so to increase economic growth. Yet per-capita growth in the industrialized countries has declined by two-thirds since the mid-1900s. If there is a trade-off between equality and efficiency, if greater disparity increases incentive to be productive, how could one explain this?

It is immoral to deprive people of what they have earned from their hard work.

Earnings depend not only on diligent work, but also on an infrastructure that has largely been a product of government investment. Those who earn more have typically benefitted more from that infrastructure. (Thanks to television advertising revenues, the average annual salary in the National Basketball Association is $4.5 million, two orders of magnitude greater than the average salaries several generations ago. In a similar vein, would Warren Buffet, Bill Gates, Steve Jobs, Ralph Lauren, George Lucas, Steven Spielberg, Mark Zuckerberg, many hedge fund managers have become billionaires had they lived in the nineteenth century?) Is it immoral that the recipients of this infrastructure pay more [back] to sustain and improve that infrastructure?

Concern about distribution detracts from attention to overall growth and productivity. It is that improvement in productivity that has allowed so many people to improve their lives.

Severe poverty reduces potential productivity. It leads to inadequate education, insufficient skills, reduced opportunity, deadened aspirations. Redistribution may increase economic mobility and may open paths to greater productivity for many more people, increasing total productivity. As the pool of people with significant productive skills widens, the competitive challenge increases for those who have such skills to utilize them effectively.

Moreover, a broad distribution of wealth and income may also enhance a sense of community and of fairness, which reinforces civic virtues. These, in turn, reduce frictional forces in the economy. This may explain why, since Aristotle, countries with relatively large middle classes have been the most productive.

Redistribution makes people dependent on government, depriving them of dignity and a desire to be productive, making them less industrious.

We are all dependent on government, but we take for granted the infrastructure that government provides without sacrificing our dignity. Moreover, those countries with the strongest social safety nets (countries of northern and western Europe) are widely regarded as among the most industrious. Their sturdy and generous social safety nets have deprived them of neither their dignity nor their industry.

At one level, the threat of privation, even endangering survival, may provide a powerful incentive to be industrious. But at a more advanced level, as the threat of severe privation recedes, industry may depend on opportunity, expectations, a desire to create value.

It was laissez faire, free market policies, that propelled the Industrial Revolution and the economic growth of Great Britain and the United States.

Laissez faire may have developed around the time of the Industrial Revolution. But Great Britain was predominantly mercantilist at that time. And the United States, influenced by Alexander Hamilton, was dirigiste, with high tariffs, government support for manufacturing, transportation (canals, railroads, highways), and the financial sector. Later, the Homestead Acts, with a primarily political (anti-slavery) motivation, had major long-term salubrious economic effects, fostering middle class farming, as opposed to sharecropping.

Through the Second World War the U.S. had among the highest tariffs in the world. Since then, its government supported new industries, from pharmaceuticals to semiconductors, directly with research grants or indirectly by funding and providing the major market for competing mid-sized companies. It set a potent economic example.

Those countries that subsequently staged economic miracles, from Bismarck's Germany through post-war Japan to the East Asian tigers and the Peoples' Republic of China, have followed this example, similarly dirigiste. They rejected pure laissez faire, imposing tariffs, funding infrastructure, targeting promising growth sectors. In contrast, those countries that have abided by the laissez faire Washington Consensus have lagged.

It is the wealthy who provide the savings and investment, the new technologies that drive growth.

Government has been the source of much successful investment in new technology, from jet aviation to pharmaceuticals, from semiconductors to the Internet. Moreover, it is middle class spending that typically drives the market for new products and provides the incentive for investment in new technology.

The Ubiquity of Natural Monopolies and Collusion: In Defense of Intervention

The assumptions of economic rationality and perfect, or even symmetric, knowledge have been successfully challenged elsewhere. The point of The Example, and of this paper, is to show, with stark simplicity, that even if these assumptions were valid, the conclusion of free market optimality would be problematic.

In particular, the additional presupposition of perfect competition is necessary to sustain the optimality of unfettered free markets. This presupposition, often implicit and taken for granted, is not a good approximation. An economy of many equal agents in perfect competition is metastable. A sufficient fluctuation will trigger an emergent dynamic that favors size. That dynamic leads to oligopoly and monopoly. Collusion, as opposed to competition, economically rational but difficult in a competitive environment, becomes increasingly easy.

Under these conditions, a neutral supervening regulatory power, one without its own economic agenda to pursue, can add, rather than subtracting, value. The interventions discussed above (as well as others) may not, even in combination, provide a uniquely optimal set of prescriptions for a world of oligopolies. But, as illustrated by The Example, access to intervention represents a clear improvement over orthodox non-intervention. The world to which the orthodox theory applies does not exist because it is inherently unstable in 'standard' conditions.

This is important. Orthodox economic theory seeks to prove the optimality of non-intervention under conditions that it regards as at least good approximations. In some cases, financial markets often cited as a prominent example, the assumptions of economic rationality and symmetric information appear to fit decently. Orthodox theory may be somewhat satisfactory in this realm (though it fails to account for bubbles and collapses, and even has difficulty with the regularity of business cycles).

In contrast, the assumption of pure competition does not come close to reality. Consider the claim, perhaps surprising, that nearly all economic sectors are natural monopolies. Suresh Naidu puts this succinctly: '*In a world of generic increasing returns, high fixed costs, and low marginal costs, monopoly will be pervasive.*' (Boushey et al. ed., p. 109) However, monopoly may pervade even beyond this range, to sectors with relatively low fixed costs and high marginal costs. The orthodox theory has a minimal range of application.

A variety of examples support this claim, noting that the standard defense of natural monopolies is that breaking them up leads to unnecessary duplication of capital expenditures and labor, increasing costs and lowering net efficiency.

In a market for transportation, consider two alternatives: (i) 10 railways with parallel tracks serving several communities; (ii) a single railway with sufficient track and rolling stock to serve the communities. (i) involves an unnecessary duplication of capital expense and labor, higher prices to the detriment of the purchasing communities. The single railway (high fixed costs) has a natural monopoly. This is widely acknowledged.

In a market for electricity, consider two alternatives: (i) 10 utilities with parallel power grids serving the community; (ii) a single utility with a power grid sufficient to serve the community. (i) involves the unnecessary duplication of capital expense and labor, higher prices to the detriment of the purchasing community. The single utility (high fixed costs) has a natural monopoly. This, too, is widely recognized.

Yet there are other examples, not so widely recognized as natural monopolies. In a market for comestibles (low-to-moderate fixed costs) consider (i) 10 mom-and-pop small grocery stores, each carrying the same goods; (ii) a single supermarket with more buying power, less capital and labor expense per unit of revenue, which may provide a wider variety of goods and lower prices. Again, (i) involves unnecessary duplication of expenses and labor resulting in higher prices and poorer selection, a detriment to the community. The supermarket appears to have a natural monopoly and to pose an existential threat to the mom-and-pop stores.

One might extend this to a chain of supermarkets. The chain has greater buying power, more flexible logistics, access to less expensive capital. It may even market its own brands. The mom-and-pop stores, and even the supermarkets, will either go out of business or be consolidated into the chain of supermarkets, the natural monopoly.

Even in so apparently mundane a business as farming (low-to-moderate fixed costs), a single large farm (ii), able to afford mechanization, able to afford micro-weather and soil moisture forecasts, better able to sustain bad harvests or lower prices, will be more efficient than a small farm (i) worked by hand and without access to such information. Independent small farmers, unable to compete, tend to consolidate (or be consolidated) into giant agribusiness. Again, the agribusiness is a natural monopoly.

Microsoft, Apple, Samsung, Intel, Google, Facebook, J.P. Morgan - Chase have low-to-moderate fixed costs in relation to their revenues. Yet they are clearly oligopolies.

In each of these cases, there is a natural tendency for (i) to consolidate toward (ii). (i) is meta-stable. In 'normal' times, characterized by small fluctuations of economically relevant variables, this state of affairs (i) can persist. But a single large fluctuation, a recession or bad harvest, causing the bankruptcy of one or two firms (or farms) and/or their acquisition by a competitor, may spark a consolidation that bestows a significant size advantage on some of the acquiring survivors. That size advantage provides an impetus to further consolidation, which further increases the advantage to size, a positive feedback mechanism leading to yet further consolidation.

The natural end point of such consolidations is monopoly. It is the advantages that accrue to size that make most economic sectors natural monopolies. Moreover, it is a challenge to find sectors in which size fails to confer significant advantages. It is the same challenge to find sectors that are not dominated by natural monopolies. And it is the same challenge to find sectors in which there is perfect competition.

Standard economic theory, to the extent that it is based on the assumption of perfect competition, fails to come close to reality. Monopolies and oligopolies are natural and dominate modern economies. Rent-seeking is economically rational. And in an oligopolistic or monopolistic economy, competition does not pose an obstacle of consequence to rent-capture.

The dynamics of an economy of oligopolies/monopolies are very different from those of an economy of perfect competition. Micro-economic considerations may be irrelevant, or at least insufficient to draw significant conclusions. Prescriptions may be radically different. Even the questions may be different.

Why does consolidation often stop short of monopoly, with apparently stable oligopolies? Is there a point of diminishing returns to advantages of size?

What are the economic implications of oligopoly for employment in small towns and rural communities? What are economic and political ramifications of urban-rural divides?

Oligopolies have large-scale buying power, even in the labor market. With this buying power comes an ability to transfer wealth from labor to capital. How can this be moderated?

Given the role of competition in creating and increasing value, can such competition be systemically incentivized within oligopolies? Are there alternative value-enhancing incentives that can be incorporated into oligopolies?

Oligopolies may be threatened primarily by new disruptive technologies and may have a natural incentive to stifle them if they cannot acquire them. Are there measures that can be taken to protect promising disruptive technologies in the face of such incentive?

From a practical perspective, economic theory would gain relevance by seeking to understand the dynamics underlying oligopolistic collusion and to assess the effectiveness and cost of various interventions. (Some of the concepts featured in contemporary economics, game theory in particular, may prove relevant to oligopolistic economics. Separately, an advantage to size, a driving force in monopolization, triggers a positive feedback loop. Positive feedback loops, which are common in fields other than economics, suggest the fecundity of macro-economic analyses in terms of nonlinearity.) Prescriptions may be more wisely governed by common sense and considerations of other social sciences.

A Political Correlate

It may be instructive to regard the standard micro-economic picture – independent autonomous economically rational agents with symmetric information competing on a flat playing field – as an economic correlate to political democracy. In democracy, independent autonomous voters make (hopefully) politically rational choices on the basis of symmetric information. (This parallel has been used to justify non-intervention in free markets, the economic correlate to democracy.)

The parallel is not entirely positive, as the political breakdowns from this picture bear an eerie resemblance to economic breakdowns. Corresponding to a failure of economic rationality, the erosion of political rationality is increasingly common. Just as scruple-deficient traders manipulate financial markets, scruple-deficient politicians manipulate potential voters to inflame and harness animal spirits, even at the cost of a potentially destabilizing polarization of society.

> *'"Darrell [Issa] pursued headlines," Bardella told me. "He wanted visibility and publicity." ... And yet, those Obama-era investigations made an indelible mark on politics. Congressional Republicans may never have been able to prove that Obama's IRS*

unfairly singled out Tea Party groups for scrutiny; or that Eric Holder tried to hide the facts about a failed Justice Department investigation into gunrunning along the Mexican border… or that Hillary Clinton, as secretary of state, tried to over up a bungled response to the Benghazi attack – but you might not know this if you spend much time on conservative media or the Trump rally circuit… But Issa and his successors, by virtue of right-wing media megaphones like Fox News (which was in its infancy during the Clinton years) and Breitbart, were able to conduct investigations as a kind of post-truth theater, where proving charges was less important than making them, loudly and repeatedly.' (Zingerle, p. 48)

In like manner, just as symmetry of information fails in economics, it is increasingly easy to destroy symmetry of information in politics. The use of technology and the practice of morally-challenged media to warp and even create false stories is widespread. Different groups are incited to act on the basis of pre-selected information, 'alternative facts,' fed to them. Fears, prejudices, resentments, whether or not they are founded, are easily stoked. This has been set out in meticulous detail, with an interesting asymmetry showing most of the fake information (in the U.S.) to be generated by the far right, Fox News in particular (Benkler, Faris, and Roberts, 2018).

A different, but important, similarity between the economic and the political, often acknowledged but overlooked, is an advantage to size. In politics, size connotes acceptance, influencing herd behavior in the same way that rising prices in a financial market can lead to bubbles. Size provides greater resources for political advertising, for the recruitment of candidates, for the funding of partisan think tanks, for the ability to fashion policies to benefit loyalists and/or punish adversaries.

Where economics and politics intertwine, the advantage to size – both economic and political – is magnified. Money buys political influence and typically uses that, not to create new technologies that would increase the size of the pie, but – with less risk and a better return on investment – to extract wealth from the rest of society.

In the U.S. (more than the E.U.) wealth buys media, funds political campaigns, and employs lobbyists to influence legislative votes. It uses its influence to craft legislation that favors wealth in general (an effectively regressive tax system (Bogenschneider, 2015), but also the special interests of the particular sectors that provide the funding (the U.S. Medicare Prescription Drug Improvement and Modernization Act of 2003).

There is carefully documented evidence that legislators are attentive to the interests of the wealthiest, but insensitive to the interests of the rest (Gilens, 2014). Despite a superficially democratic structure, despite lip service paid to democracy, this fits the definition of 'plutocracy,' government acting to promote the interests of the wealthy, even at the expense of a large majority of citizens. It may be stealth plutocracy, but still plutocracy.

The mortal danger to modern democracy comes not from without, but from within. Some is due directly to economic theory.

Functional, as opposed to titular, democracy is incompatible with the general assumptions and principles that drive contemporary economics.

A society of economically rational symmetrically-informed agents competing with each other, agents with comparable political status, is meta-stable. A sufficient fluctuation in either the economy or the polity may trigger a dynamic that favors size. The society – spontaneously – devolves into an oligopolistic plutocracy, wealth purchasing political influence and using that influence, not to create new wealth, but to extract more wealth from the rest of society.

The U.S. has devolved disconcertingly far in this direction, under the dubious cover that it offers a flat playing field to all.

It may be important to identify and nurture countervailing factors at both the individual and structural levels, factors that may oppose this devolution. At the individual level, it is important to reject the notion that if every agent acts with economic rationality in its own best interests then competition will ensure an economic maximum. Competition is too easily replaced with consolidation, and then in the name of economic rationality, with oligopolistic collusion.

Even if this were not the case, unless agents prioritize integrity over profit, frictional forces will grind any economy to a halt. (Economists implicitly assume integrity, but have often explicitly denied the relevance, or even objectivity, of morality.)

It is not just individual integrity that must be prioritized. It is also necessary for agents to realize that their own well-being is strongly dependent on the success of their society, on the foundation that has been laid by others (often government). There are few winners in deeply impoverished societies.

It is essential to attend to those societal features that are requisite to the success of societies. Societies that have been run by and (especially) for elites have not performed well over the long term. (This does not entail that *only* democracies, which have their own intrinsic deficiencies, especially in sacrificing long term prospects to benefit voters in the short term, can be successful. Singapore, under Lee Kwan Yew, showed that an autocracy can be pluralistic, inclusive and fair, and can make wise strategic economic decisions that benefit most of its people.)

This casts a light on structural issues. Of particular concern is the natural tendency for a positive feedback loop to indefinitely concentrate both political and economic power. Such a concentration leads to reduced economic mobility, an ossification of classes, and an extractive society.

This constitutes a threat to both the polity and the economy, spawning the greatest danger most democracies face today: a devolution into either (i) plutocracy in which wealth buys political power and uses it to extract more wealth, or (ii) autocracy in which political power seizes wealth (or merely pressures it) and uses it to buttress its position. (At some level it becomes difficult to distinguish between the two.)

It is easy to underestimate this danger because such a devolution would be natural, the effect of dynamics favoring economic and political size and power. Many believe that what is natural is necessarily good.

To the contrary, it is common that extremes that are end points of positive feedback cycles, from hysteresis in electronics to 'J' curves in population ecology to cytokine cascades in medicine, are pernicious despite their being natural. Dynamics that favor size cause feedback loops in which size and power differentials increase without bounds. The extreme of economic disparity is similarly pernicious and can be fatal to democracies.

It may be vital to thwart the natural, to implement institutional buffers that constrain spontaneous positive feedback cycles. These buffers could include effectively progressive tax systems that fund social safety nets (and [to] that reduce the advantages to size that lead to oligopoly), the inhibition of the purchase of political power and of the use of political power for enrichment, and general government interventions to protect and enhance individual autonomy. Political vigilance may be necessary to moderate extreme concentrations of wealth and power, extremes that pose a grave threat to open pluralistic society and shared prosperity.

References

Acemoglu, D. and Robinson, J. (2012). *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*, Crown Business: New York.

Benkler, Y., Faris, R., Roberts, H. (2018). *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*, Oxford University Press: Oxford.

Bogenschneider, B. (2015). Income Inequality and Regressive Taxation in the United States, *Interdisciplinary Journal of Economics and Business Law* **4** (3): 8-28.

Cohen, S. and DeLong, B. (2016). *Concrete Economics: The Hamilton Approach to Economic Growth and Policy*, Harvard Business Review Press: Boston.

Gilens, M. (2014). *Affluence and Influence: Economic Inequality and Political Power in America*, Princeton University Press: Princeton.

Hayek, F. (1948). *Individualism and Economic Order*, Routledge Press: London.

Naidu, S. (2017). 'A Political Economy Take on W/Y,' in *After Piketty: The Agenda for Economics and Inequality*, ed. Boushey, DeLong and Steinbaum, Harvard University Press: Cambridge, p. 109).

Okun, A. (1975). *Equality and Efficiency: The Big Trade-off*, Brookings: Washington D.C.

Scheve, K. and Stasavage, D. (2016). *Taxing the Rich: A History of Fiscal Fairness in the United States and Europe*, Princeton University Press: Princeton.

Zingerle, J. (2018). 'Trump's Inquisitors,' *The New York Times Magazine*, December 23, 2018; p. 38f.

Interdisciplinary Journal of Economics and Business Law, 2019, Vol. 8,
Issue 3, pp. 137-144

Comment

Food for thought: Greggs' vegan sausage roll – the power of corporate outreach and brand

Dawn Carr[1]

The global market for vegan food has seen explosive growth that shows no signs of abating. The online food-delivery platform Just Eat reported a 987% increase in demand for vegetarian and vegan food in 2017 (Just Eat, 2018).

The appetite for vegan options is being driven by young people striving to reduce the amount of harm their food choices cause to others, the planet, and their own health. One in five young people think everyone will be vegan by 2030 (Williamson, 2018). The widespread and growing availability of an abundance of tasty plant-based foods is making it easier than ever to give vegan eating a try.

Restaurants and other food enterprises are recognising – and responding to – the demand for vegan food. In leveraging brand identities, PETA's corporate outreach campaigns are playing a significant role in boosting this trend. Our work urging UK bakery chain Greggs to introduce a vegan sausage roll presents a useful case study revealing the benefits to established brands of garnering a new following among millennial and Generation Z consumers.

[1] People for the Ethical Treatment of Animals Foundation, Society Building, 8 All Saints Street, London N1 9RL. Email: DawnC@peta.org.uk

The rise of vegan eating

The number of people opting to eat vegan is growing rapidly: a recent survey found that there are 3.5 million vegans in the UK, up from 540,000 in 2016 (Petter, 2018). In response to such increased demand, many new vegan enterprises are opening, but non-vegan businesses are also increasingly catering to this market by incorporating vegan options into their brand.

Research shows that ethical supply chains and corporate accountability are important to people – particularly millennials, who have major purchasing power. An Opinium poll commissioned by PETA in 2017 found that 76% of British 18- to 34-year-olds are interested in trying more vegan foods. These socially conscious consumers are significantly influencing cultural and social trends, and businesses are responding to the demand for ethical, sustainable, healthier choices.

A case in point is the release of Gregg's vegan sausage roll, which was a result of years of urging from PETA's Corporate Affairs Division for the UK's largest bakery chain to add this item to its menu. The launch exceeded all expectations – customers were literally queueing up for this tasty treat. The vegan version mirrors some of the meat-based sausage roll's features, including 96 layers of light and crisp puff pastry, but instead of being made with animal flesh, it has a delicious vegan meat filling produced specially for Greggs by Quorn, incorporating both brand names and reputations.

A lucrative opportunity

Greggs credited the success of the vegan sausage roll with supercharging its annual sales figures – which topped £1 billion for the first time (Insider.co.uk, 2019). The company experienced an 11% rise in like-for-like sales in the 19 weeks following the launch (Monaghan, 2019).

When asked about sales figures earlier in the year, Greggs Chief Executive Roger Whiteside said, "We're not giving out numbers – it's in the hundreds of thousands – but it is the fastest selling new product we have launched in six years. It's literally flying off the shelves" (Ford, 2019). But as you'll see below, Greggs almost missed this opportunity.

The importance of corporate outreach

The trend towards offering more vegan options is very much driven by demand, but we are also seeing that the wider availability and variety of vegan foods generates curiosity among meat-eaters who may be motivated to reduce their meat, egg, and dairy consumption for health, environmental, or animal welfare reasons. Such consumers – often referred to as "flexitarians" – are a key market for plant-based foods. For example, 93% of people who buy US vegan meat brand Beyond Meat's burger also purchase meat (Durbin, 2019).

Billions of non-human animals are forced to endure miserable lives and violent deaths each year for their flesh, milk, and eggs. PETA's goal is to reduce animal suffering, so we work with companies to push them to introduce cruelty-free vegan food. Our corporate outreach complements and builds on the success of our decades-long efforts to influence individual behavioural change by helping to increase the availability of tasty vegan foods, thereby making it easier and more socially acceptable for people to try them while continuing to use a brand they know.

Choosing the right targets

As a charity, PETA's resources are limited, so we focus our efforts primarily on large high-street chains, since they offer the potential to make vegan options available to people in every corner of the UK. Greggs has almost 2,000 stores across the country, serving millions of customers each week, and its low prices mean its food is affordable to people of all income levels. Focusing our corporate outreach on such brands enables us to help bring vegan food offerings to the largest number of people possible.

The push for 'grab-and-go veganism'

The goal of our corporate work is not to make eating out easier for vegans – it's to make vegan choices more accessible and appealing to meat-eaters. The key to capturing the public imagination with a vegan launch is to offer tempting, satisfying options that appeal to vegans and meat-eaters alike – and plant-based versions of brands' most iconic dishes have proved time and again to elicit the most excitement from consumers.

Greggs' vegan sausage roll holds its own next to the meaty version, and many meat-eaters have said they prefer it. It even offers more protein than its flesh-based counterpart, answering the perennial question: where do vegans get their protein? *Vice* magazine described its impact best:

> *Anti-vegans should be mad. This is no Veggie Pret opening, nor a Tesco vegan ready meal range; this is mainstream veganism in action on par with Wetherspoons introducing their vegan menu. This is the British equivalent of American McDonald's getting that vegan meat burger. This is easy and bland and delicious; this is grab-and-go veganism. It has arrived.* (Ewens, 2019)

Much of the work PETA does is behind the scenes for reasons of confidentiality, but we also work publicly with many brands. For example, we helped Italian-American restaurant chain Frankie & Benny's add exciting vegan options to its offering. We reviewed its menus and made brand-appropriate suggestions for foods that would appeal to vegans and flexitarians alike. In January 2019, the company launched one of the most extensive vegan menus of any UK chain, including vegan takes on chicken parmigiana, mac 'n' cheese, burgers, hot dogs, nuggets, garlic-mozzarella pizza bread, BBQ chicken pizza, and more.

We also supported Papa John's in developing new menu items that would make it the first national pizza-delivery chain in the UK to offer vegan cheese pizza, and just as Greggs struggled to keep up with demand for the vegan sausage roll in the days following its release, Papa John's sold out of non-dairy cheese within two days of its launch. The pizza chain's marketing director expressed confidence that tapping into the vegan market would greatly benefit its brand and acknowledged PETA's role in making this possible: "We worked closely with PETA who helped us develop the recipes and find the best vegan products, so we expect that the new additions will be a huge hit" (Chiorando, 2019).

Our corporate campaign strategy

Working behind the scenes

Corporate outreach work by PETA and our international affiliates is behind many of the shifts we are seeing globally. Our pragmatic, helpful, and strategic approach to working with corporate decision-makers has proved highly successful.

Rather than making an adversarial first approach, PETA representatives meet with companies in their boardrooms and – by recognising and acknowledging their objectives and brand identity – work to help them develop and promote winning foods that will give them access to the booming vegan market.

Knowing the market

No one goes to Greggs for carrot sticks and hummus. In 2018, the chain released a vegan Mexican Bean Wrap, and although it proved popular – and won a PETA Vegan Food Award – unlike the vegan sausage roll, it didn't generate excitement. That's not because it isn't delicious – it is – but because it's not what the brand is known for.

PETA is constantly monitoring launches of new vegan products and menu items across the food sector – noting which products do best – so we're often in a position to provide companies with market insights that they wouldn't otherwise have. At one stage in our work with Greggs, the company proposed scrapping its plans for a vegan sausage roll and working instead on a vegetarian one. We urged it to reconsider, as this would have been a huge missed opportunity for its customers and the brand. It was at this point that to highlight the demand, PETA launched a friendly public petition calling for a *vegan* sausage roll from Greggs, which was signed by more than 20,000 people eager to see it available.

Media interest in vegan food – and in particular, vegan versions of iconic dishes – meant the release of the vegan sausage roll grabbed headlines and dominated social media. Even television host Piers Morgan mocking the menu addition proved not to be harmful to the brand. There can be no doubt that a vegetarian sausage roll would have failed to generate the enthusiastic response the vegan version received.

Providing opportunities for SMEs

PETA's work with large corporations inevitably has a ripple effect, as the more that chain restaurants and shops reap the rewards of launching vegan options, the more that smaller businesses see the potential benefits to their brand of offering plant-based foods. To help them capitalise on this trend, we published our "Guide to Introducing Vegan Options", a resource that takes independent businesses and small chains through the process of "veganising" their most popular products and menu items, highlighting successful high-street case studies such as Greggs' vegan sausage roll and Wagamama's Vegatsu (vegan katsu) curry.

Our corporate work can also benefit small producers of vegan speciality foods, as we often connect them with the larger businesses we work with. Although we commonly have to keep our role in forging these links confidential, this can provide these suppliers with lucrative opportunities to grow their businesses and brands.

The future is vegan

The increased sales and surge of media attention that Greggs has enjoyed in the wake of the launch of its vegan sausage roll has made many other big players in the food industry more open to introducing ambitious vegan options. And since much of the news coverage of the sausage roll launch cited PETA's involvement, many other companies are eager to draw on our expertise in order to ensure their plant-based offering has a similar impact.

When fast-food behemoth KFC announced that it was planning to launch a meat-free chicken option, we wasted no time in urging the chain to make it vegan and launching a petition, through which 12,600 supporters added their voices. We emphasised that no one chooses a food because there is a little bit of milk or egg in it but that there certainly are people who would *not* choose it for that reason. In June, KFC launched a four-week trial of a vegan version of the brand's signature dish, fried chicken – the company is calling it the Imposter Burger. KFC used its familiar Colonel Sanders branding on the packaging, alongside a new spin on its well-known slogan: "Finger Lickin' Vegan". PETA helped promote this addition to KFC's menu, and it has proved to be an unmitigated success – the initial batch of Imposter Burgers sold out in just four days.

For years, PETA and our affiliates have engaged in traditional campaigning – including protests, boycotts, and celebrity ads – urging KFC to improve its animal welfare practices. Such campaigns played a vital role in raising awareness of cruelty to animals in the food industry and in encouraging consumers to vote with their wallets to help end this abuse. Now, companies like KFC are recognising the value to their brand and bottom line of offering sustainability-conscious young consumers exciting options that are better for the environment, animals, and human health. While the companies' goal may be financial reward, the benefit to animals is the same.

Conclusion

As concerns about health, climate change, and animal suffering continue to gain traction, the number of people looking to make sustainable, compassionate food choices will, in turn, continue to grow. PETA's corporate campaigns are helping to ensure that this demand is matched by the availability of plant-based foods, enabling consumers to choose a vegan option over an animal-based one, no matter where they are or how much they spend on food. By helping businesses introduce vegan options in line with their existing brand identity, we spark widespread interest in vegan eating, thereby reducing the consumption of animal-derived foods and sparing animals immense suffering in the food industry. In addition, those with deep environmental concerns and an interest in maintaining good health can also embrace the vegan option while continuing to choose their preferred brands.

References

Chiorando, M. (2019). Papa John's Pizza launches vegan cheese nationwide in the UK. *Plant Based News*. Available from https://www.plantbasednews.org/post/papa-johns-pizza-vegan-cheese-uk.

Durbin, D. (2019). Beyond Meat fattens up as shares more than double in IPO. *APNews.com*. Available from https://www.apnews.com/a5a1a1add5bb4844b87ebe8beaa2f0c6.

Ewens, H. (2019). The only Greggs vegan sausage roll review you need. *Vice*. Available from https://www.vice.com/en_uk/article/59v88b/the-only-greggs-vegan-sausage-roll-review-you-need.

Ford, C. (2019). Greggs' vegan sausage roll to be rolled out to ALL stores after its sellout launch. *ChronicleLive.co.uk*. Available from https://www.chroniclelive.co.uk/business/business-news/greggs-vegan-sausage-roll-rolled-15650825.

Insider.co.uk. (2019). Greggs sales top £1 billion for the first time as vegan sausage roll storm brings in more customers. *Insider.co.uk*. Available from https://www.insider.co.uk/news/greggs-sales-top-1-billion-14099373.

Just Eat. (2018). Plant-based diet 2018. Available from https://www.just-eat.ie/blog/plant-based-diet-2018/.

Monaghan, A. (2019). Greggs' vegan sausage rolls fuel profit boom. *The Guardian*. Available from https://www.theguardian.com/business/2019/may/14/greggs-vegan-sausage-rolls-fuel-profit-boom.

Petter, O. (2018). Number of vegans in UK soars to 3.5 million, survey finds. *Independent*. Available from https://www.independent.co.uk/life-style/food-and-drink/vegans-uk-rise-popularity-plant-based-diets-veganism-figures-survey-compare-the-market-a8286471.html.

Williamson, H. (2018). One in five young adults think we will all be vegans by 2030. *Metro*. Available from https://metro.co.uk/2018/09/12/one-in-five-young-adults-think-we-will-all-be-vegans-by-2030-7934759/.

Interdisciplinary Journal of Economics and Business Law, 2019, Vol. 8, Issue 3, pp. 145-152

Book Review

Sino-US Trade War: A New Challenge to Globalisation

Ed. Surender Mor

Vista International Publishing House: Delhi; 2018, 105 p.

ISBN 978-93-83905-19-5

This is a group of essays clustered, at different distances, around the general theme of the United States – People's Republic of China trade war and its impact on India.

The trade war itself, for many economists, illustrates the failure of politicians to understand economic issues. Did the P.R.C. regularly steal intellectual property? Yes. (Did the U.S., in its economic infancy, regularly steal intellectual property? Yes.) Does the P.R.C. use tariffs and government subsidies to gain 'illicit' advantage? Yes. (Did the U.S. for centuries use tariffs and government subsidies to gain 'illicit' advantage? Yes.)

In light of this history, is the present initiation of tariffs against the P.R.C. (and others) a judicious and nuanced response to complex trade and geo-political issues? Or can it be better understood as a politically-driven campaign promise, an expensive economic counterpart to a cut-rate copy of the Great Wall of China?

To some extent, the trade war has developed as one might have expected. In early 2018 the U.S. announced tariffs on a broad range of Chinese goods, followed by tariffs on all aluminum and steel imports. The Chinese, predictably, followed with retaliatory tariffs and quickly devalued their renminbi by about 10%, recouping much of their disadvantage on the tariffed goods and gaining significant advantage on other goods.

The Chinese devaluation was not a surprise. Neither was the subsequent dramatic two-thirds decline in the Baltic Dry Index, a measure of shipping rates.

Yet there have been surprises. The negative U.S. balance of trade did not decrease, but increased 10% from the first 11 months of 2017.

'*Indeed, even after the imposition of various kinds of tariff barriers the US trade balance with China has widened during the first 8 months of the year 2018.*' (p. 8)

The U.S. balance of goods with China hit a record deficit in October 2018, improving in November. A decelerating U.S. economy could lead to further improvement in coming months, but it may be difficult to link the improvement to the tariffs.

The tariffs were designed to help domestic aluminum producers, but despite a brief blip in the aluminum price, it is now lower than when the tariffs were declared.

'*The BSE Metals Index fell over 13 per cent from March after the news of trade wars came in.*' (p. 16)

If a major trade war can be kindled by the sparks of politically-incited-and-enhanced resentments, then it may be reasonable for countries that could be collateral casualties (or beneficiaries) of such a trade war to consider how to position themselves to minimize risk and maximize opportunity.

This book, a collection of papers deliberated on at a symposium, 'Sino-US Trade War: A Potential Threat to World Peace and Prosperity' in August 2018 in Sonipat, India, addresses a number of these issues.

In the first chapter, 'U.S. China Trade War: Genesis and Implications,' Professor Mor provides a valuable introduction to the book. He furnishes a plausible rationale for the trade dispute, identifying the end of 2001, when China joined the World Trade Organization, as a critical turning point affecting both China and the U.S.

After that point, the U.S. trade deficit with China widened dramatically. U.S. economic growth decreased significantly, while Chinese growth increased. (The PRC growth increase is apparent if one takes the 1995-2001 time-frame as a baseline, but is less so if one extends the baseline to 1994-2001 and inverts a bit if one also uses purchasing power parity. Moreover, Chinese economic growth has been slowing since 2008, despite its increased trade surplus.)

From the U.S. perspective, the trade imbalance led to a deindustrialization, with a concomitant loss of relatively high-paying jobs for men without college education – a major part of the Trump voting base. (Ironically, because trade flows and capital flows tend to balance each other, the Trump administration's tax bill, significantly widening the U.S. federal deficit, has contributed to the increase in the trade deficit. Still, it is easier to blame others for one's plight. And that has proven politically effective, if economically risky.)

Professor Mor expresses enthusiastic blanket approval for free trade:

'*Benefits of trade for trading partners are by now established in theory and practice. The economists have been advocating free trade among nations as first-best policy since classical era…*

International Trade has been regarded as the engine of growth and prosperity of nations all over the world and trade among nations has been carried out since time immemorial, and these efforts led to the industrial revolution in Europe and colonial rules in third world countries. The economists from classical (Adam Smith, Ricardo, Mill and so on) to the modern era (Heckscher-Ohlin, Bhagwati, Krugman and so on) have been advocating free trade among nations as first-best policy. The spirit of enhancing world trade continues…' The process of globalisation intends to connect all nations of the world and the fruits of such connectivity must trickle down to every corner of the world irrespective of their size and stage of development.' (p. 1, 3)

(Paul Krugman, cited as an advocate of free trade, has been clearly impressed by considerations of Ricardian comparative advantage. But he has subsequently qualified his position sufficiently to resemble a retraction.

'But it's also true that much of the elite defense of globalization is basically dishonest: false claims of inevitability, scare tactics (<u>protectionism causes depressions</u>!), vastly exaggerated claims for the benefits of trade liberalization and the costs of protection, hand-waving away the large distributional effects that are what standard models actually predict. I hope, by the way, that I haven't done any of that; I think I've always been clear that the gains from globalization aren't all that ... and I think I've never assumed away the income distribution effects.

Furthermore, as Mark Kleiman sagely observes, the conventional case for trade liberalization relies on the assertion that the government could redistribute income to ensure *that everyone wins – but we now have an ideology utterly opposed to such redistribution in full control of one party, and with blocking power against anything but a minor move in that direction by the other.*

So the elite case for ever-freer trade is largely a scam.' (Krugman, P. 'A Protectionist Moment,' *The New York Times*, March 9, 2016))

One can be more circumspect about the guaranteed benefit of free trade and still regard the U.S.-China trade dispute as a negative, a risk to the global economy.

Along these lines, the generally-acknowledged failure of the Washington Consensus supports a more cautious perspective that is also reflected in this book.

Ramnath Srinath provides such a guarded assessment in 'America first: Is Trump fueling De-Globalization?'

'*For developing countries, lowering all barriers to the tides of the global economy may end up drowning much of local production. Raising barriers that are too high may be counterproductive, if not futile. Countries that find the golden middle, like Chile [China?] and Singapore, tend to thrive*.' (p. 18-9)

(It may be relevant that both China and Singapore, following other success stories, built their economies on robust industrial policies. The mode of intervention may be more important than the amount.)

Srinath stresses an important point that has often been overlooked:

'*Over the last two decades, overseas development assistance from the rich to poor countries has totalled $50-80 billion per year. In the same period, every year, $500-800 billion of illegal funds have been sent from the poor to rich countries. That is, for every one dollar of aid money over the table, the West gets back $10 under the table and, for good measure, lectures the rest on corruption.*' (p. 17)

He also calls important attention to the growth in illicit trade, often in unsavory areas such as drug and human trafficking.

Jaswinder Singh Brar ('Factors in the Eruption and Persistence of the Sino-US Trade War') provides an assessment of different aspects of free trade.

'*Free trade by countries in general has been preferred in selected sectors as it always acts as the weapon of the strong. Consequently, trade has been* *used as strategic policy instrument not only for protection of domestic markets but also to contain the adversarial powers… The attempts to manage trade wars becomes fragile when trade wars have been the result of maintaining global supremacy, containment of emerging powers, realization of strategic goals, promotion of exports and providing protection to desperate domestic industries.*' (p. 22, 23)

This provides a background to his analysis of the current Sino-U.S. trade dispute. (Ironically, he notes that the U.S. imports 90% of its aluminum, making tariffs in this sector primarily a tax on consumers.) He concludes, reasonably:

'*It is important to note that international posturing of countries plays crucial role in deciding the domestic political worth and thereby standing of prominent leaders. Most often it becomes difficult to trace back from the declared global positions which entail strong domestic political costs. This also seems to be the case right now as domestic polity of both the countries revolves around strong personalities in their own right.*' (p. 30)

He warns, citing examples from history, that a full-blown trade war would have serious negative consequences for both the U.S. and China, and also for the global economy.

Pradeep S. Chauhan, in 'Critical Analysis of Sino-US Trade War: Potential Impacts on Rest of the World and India,' comments that China has effectively targeted retaliatory tariffs to sectors geared to Trump's base. By far the largest importer of U.S. soybeans, it imposed a hefty tariff on those soybeans. And from May into July, soybean (and soybean meal) prices declined by 20% (though a swine flu epidemic may have exacerbated the impact of this tariff). The E.U. similarly targeted Harley Davidson motorcycles, Kentucky bourbon, and orange juice.

He observes that in the short run, India could benefit as a producer of soybeans, but adds that in the long run, trade wars lead to higher inflation and lower growth. His primary concern is higher U.S. inflation leading to higher interest rates, which in turn would inflate the U.S. dollar, with negative ramifications for developing countries.

(The U.S., however, is facing economic headwinds of increasing intensity, primarily from excessive public and private sector debt and enormous public deficits amounting to 5%

of GDP. Monetary tightening may prove an unattractive option.)

Pawan Kumar Sharma, in 'Sino-US Trade War: A Potential Threat to World Peace and Prosperity: Why should India worry?' calls attention to the trade spat that the U.S. had with Japan in the late 1980s.

He asks an important question: '*If it is the rise of China, as many describe, then why has the US slapped tariff measures against almost all of its traditional allies? The U.S. is pursuing throwback approaches with the very partners it seeks to enlist against Chinese rule breaking'*. (p. 47) But he quickly segues into an analysis of the effect of the Sino – U.S. trade dispute on India.

One concern is the value of the Indian rupee. '*In the last one month or so, the value of the rupee has dropped to an all-time low, when on some occasions it was hovering around 70s against the US dollar. This coincided with Donald Trump's threat of imposing a fresh round of tariffs on exports worth $200 billion. This trend can be traced to the weakening of the US dollar, which automatically creates a negative impact on the trade deficit of India, causing a chain reaction of sorts.*' (p. 48) (However, the U.S. dollar index rose sharply after a major round of tariffs was announced in April, and it remains above where it was a year ago.)

An independent concern is the potential domestic inflationary effect of the retaliatory tariffs imposed by India. On the positive side, U.S. tariffs imposed on China may make Indian goods more competitive.

'Trade Openness and Economic Growth: Case of India and China' by Anju Rani is a change of pace. She measures the openness to trade of India and China since 1991 and compares that openness with the economic growth of the two countries.

Her results are surprising, that openness to trade, at least for these two countries over the time span she considers, is not relevant to economic growth. '*Results also show that there is no causality between economic growth and trade openness for both economies and are not supporting trade led growth hypothesis.*' (p. 53)

This is a startling conclusion, and it supports extending her study to other countries and longer time spans.

'China-U.S. Trade War: Key Issues' by Geetu Gupta is largely a review of the literature concerning the economic impact of tariffs. It provides a clear discussion of potential responses by China, including (an issue often overlooked) restricting U.S. access to rare earth metals, vital in a number of high-technology applications.

He advocates as an alternative to protectionism, submitting complaints to the World Trade Organization. (This may be difficult, as the Trump administration has raised questions as to whether it would be bound by WTO rulings.)

'Sino-US Trade war and Implications for Global Peace: Some Random Reflections' by R.S. Yadav adds a broader perspective, regarding this trade spat from the perspective of world peace. The author notes:

'*At the outset one has to be clear that peace is not merely a phenomenon of absence of war, rather it is also to do away with the basic determinants which directly or indirectly create conditionality of war. Present rift or trade war between the USA and China is not merely resulting of their economic differences, rather it is based on the deep-rooted cleavages between them due to the issues of geopolitical and geoeconomic supremacy. Hence, current tensions are symptomatic of their longstanding differences. Problem is not purely of economic nature related to trade, investments, intellectual property rights etc., rather it is more concerned with their power position in global politics and their desire for establishing hegemony or dominance in this de-ideological world.* (p. 77)

With this diagnosis of the root problem as mainly geopolitical, Yadav notes that neither capitalism nor communism nor globalization has successfully addressed the problems of global humanity. The imposition of tariffs is not likely to resolve the complexities of fundamentally different approaches.

He stresses the need for '*an inclusive, integrated and multilateral economic model based on the principle of sustainable development for a durable peace in the world.* (p. 80)

In a short article, 'Sino-US Trade War: A Potential Threat to World Peace and Prosperity,' Shri Bhagwan Dahiya calls attention to history. He characterizes Trump as a mercantilist who measures the fairness of the global trading system in terms of the U.S. balance of trade.

Even though he implicitly rejects mercantilism as a viable economic platform, he does not see this trade war as posing a serious threat. '*Imposition of tariff by US and China on each other goods shall positively reshape the global economy. There shall not be any threat to world peace and prosperity*.' (p. 83)

The final chapter of the book, 'Sino-US Trade Wars: Turbulence in the Global Economy,' by Geethanjali Nataraj, provides a somewhat more caustic assessment of Trump administration policies:

'*The Trump administration has been impulsive in provoking trade conflicts. There are two aspects to Trump's trade-war policy. One is the action against the EU, Canada, and Mexico (driven by right wing nationalist politics); the other is the stance toward China. Trump thinks that a protectionist approach to trade will increase jobs in the US and minimize the trade deficit, but he is mistaken. A trade war might not only trigger a recession in the global economy but could even plunge the U.S. economy into recession…*

But there is no doubt that the US is likely to become more and more isolated and will be forced to rethink its strategy… It's important that the multilateral system is put back on track and the developing world will need to put consistent pressure on the US not to abandon the multilateral trading system which is the only fair forum for settling disputes among nations.' (p. 85, 87)

The article faults China for its artificially depressed currency, an attempt to make its exports more competitive, and for its low rate of domestic consumption. It also calls attention to the failure of Western countries, especially the U.S., to adequately redistribute the fruits of global trade. This failure has driven a mounting backlash against globalization.

In sum, the essays in this book are intelligent and relevant to the problems now facing the global economy. The book is a valuable addition to what is likely to be a rapidly-growing literature.

Kenneth S. Friedman

Regis University,
Denver, Colorado