Western Undergraduate Economics Review

Western SocialScience

ISSN 1705 - 6098

Annual 2013

Western Undergraduate
Economics Review

2013

The *Western Undergraduate Economics Review* is an annual publication containing papers written by undergraduate students in Economics at Western. First published in 2002, the *Review* reflects the academic distinction and creativity of the Economics Department at Western. By showcasing some of the finest work of our students it bestows on them a lasting honour and a sense of pride. Moreover, publication in the *Review* is highly beneficial to the students as they continue their studies or pursue other activities after graduation. For many, it is their first publication, and the experience of becoming a published author is a highlight of their undergraduate career. The *Review* is a collaborative effort of the students, faculty, and staff of the Economics Department. All papers submitted to the *Review* are essays written for courses taken in the Department. Some are by students in the early stages of their Economics studies, while others are papers written by senior students for the Department's unique thesis course, Economics 4400. Selections are made by the edition editors, in consultation with a faculty advisor, based on creativity, academic merit, and the written quality of the article.

http://economics.uwo.ca/undergraduate/undergraduate_economics_review.html

ISSN 1705-6098

Contents

Editors' Comments

It gives us great pleasure to present the 2013 edition of the *Western Undergraduate Economics Review*.

The articles, originally written as essays in designated Economics courses, have been selected from papers submitted by undergraduate students. The essay requirement varies from a relatively short piece in a one semester course to a full-fledged thesis in the final year capstone course, Economics 4400, and that is reflected in the choice of topics as well as the length and depth of the various contributions.

The ongoing financial turmoil in many countries, which almost became a global crisis some years ago, continues to interest economists. Not surprisingly, three of the four papers here deal with some aspects of banking, including the role of central banks, and financial markets. Morgan MacInnes considers the case for central bank independence, Jie Ren attempts to explain the high profitability of China's major state-owned banks, and Dai Li deals with liquidity risk and asset prices. Zachary Nash examines how salary caps and shared revenue systems affect competitive balance in professional sports leagues, another timely topic given the recent history of collective bargaining in hockey and other sports.

We'd like to thank all the students who shared their work with us. Thanks are also due to Jennifer Hope for her expert help and persistence, and to our faculty advisor, Professor Kul Bhatia.

Marc Corbeil
Anna Zhu
London, Ontario
July 2013

Acknowledgments

We also extend our sincere thanks to the Social Science Student Donation Fund for its ongoing financial support of the *Western Undergraduate Economics Review*.

The Case for Central Bank Independence

Morgan MacInnes

It has been theorised that insulating a country's monetary authority from political influence has a beneficial impact on the nation's economic health. Empirical evidence shows that the degree of independence a nation's monetary authority enjoys correlates with lower and more stable levels of inflation. It does not appear, as some critics claim, that this price stability is achieved at the expense of real economic output. On the contrary, there is evidence to suggest that central bank independence is associated with stronger levels of economic growth. Therefore, governments would do well to uphold the independence of central banks.

Central bank independence refers to the degree to which a state's monetary authority is insulated from political influence. The logic behind central bank independence is that subjecting the monetary authority to influence from elected politicians, and thereby indirectly to public opinion, would be detrimental to the country's economy. In democratic states, the regular election cycle impedes long term economic planning. The gains in output generated by expansionary monetary policy manifest themselves before the resultant rise in inflation. This encourages politicians in power to implement expansionary policies in the periods before elections: the rise in output will help win the favour of the public, and any resulting rise in inflation will not be felt until after the election takes place. This deviation from optimal long term goals in favour of short term political expediency is referred to as time-inconsistent policy (Laidler and Robson 2004). Compared to a scenario in which decision makers have no incentive to eschew long term planning, such short term thinking will result in a higher average rate of inflation, with growth remaining the same or lower over the long run.

Additionally, politicians will seek to broaden their support by appealing to the various interest groups which compose the electorate. While the public in general may be averse to inflation, it is unlikely to be the foremost concern of any particular group. Farmers and fishers will likely be more concerned about subsidies for their industries than about price stability, and students will likely put greater emphasis on investments in education. In order to win the support of such groups, politicians will pledge to increase government spending in their areas of interest. Because increasing taxation would be politically unpopular, deficits incurred in such a manner are often financed through the central bank, resulting in monetary expansion and higher inflation (Laidler and Robson 2004). As such, even though the public does value low inflation, it will not be the deciding factor for most individual voters, the cumulative effect of which will be a higher rate of inflation than the populace actually desires.

The solution is to remove the reins of monetary policy from the hands of elected politicians and delegate the task to a separate body with a mandate to maintain price stability. Free from the constraints imposed by public opinion, an independent central bank would be able to stabilise inflation at a lower level than would be possible if it were beholden to an elected government.

Empirical evidence supports this line of reasoning. Alesina and Summers (1993) use an index to rank several OECD countries by degree of central bank independence, then compare these rankings with average national inflation rates between 1955 and 1988.[1] The degree of independence of a central bank is measured in terms of political and economic independence. Economic independence refers to the conditions under which the central bank is required to finance government deficit, and the political independence of the banks is based on such factors as the government's ability to appoint members of the governing board, government representation on the governing board, and to what degree monetary decisions require government approval. The study finds a "near perfect" negative correlation between average rates of inflation and the degree of central bank independence. The countries with the two most independent central banks, Switzerland and Germany, enjoyed the lowest average inflation among the countries observed (approximately three percent), while the three least independent central banks in Italy, Spain and New Zealand, corresponded to the three highest average inflation rates, all being over seven percent. The study also shows a strong negative relationship between central bank independence and variability in the inflation rate, which is itself undesirable in that it creates market uncertainty.

The case for central bank independence is further supported by the findings of a recent study by Jacome and Vazquez (2008), which looks at central bank independence and inflation rates in developing economies in Latin America and the Caribbean between 1985 and 2002. [2] In the 1990s, many countries in this region implemented reforms to increase the independence of their central banks. Inflation rates fell across the region from an average of approximately fifty percent in 1985 (excluding cases of hyper-inflation) to around seven percent in 2002.[3] The study finds that countries in the region experienced an average inflation rate of 49.53 percent during their pre-reform periods, which dropped dramatically to an average of 11.53 percent during the countries' post-reform periods. This evidence strongly suggests inflation can be greatly reduced through strengthening central bank independence.

The effects of central bank independence on the real economy are also worth discussing. Some critics allege increased central bank independence brings about lower inflation at the expense of real economic output (which would entail higher unemployment), the

---

[1] The countries included were Australia, Belgium, Canada, Denmark, France, Germany, Italy, Japan, the Netherlands, New Zealand, Norway, Spain, Sweden, Switzerland, the United Kingdom, and the United States of America.

[2] The study employs four different measures of central bank independence: the CWN index, the CWNE index, the index developed by Grilli et al. (1991), and an index based on the turnover rate of central bank governors.

[3] It should be noted that the study concludes that this drop in inflation rates was due only in part to increased central bank independence, being a product of various macroeconomic reforms.

reasoning being that the more independent the central bank is, the less it is willing to engage in active policy intervention to counter cyclical downturns in the economy (Alesina and Summers 1993). Empirical evidence points to the contrary however, as demonstrated by the work of De Long and Summers (1992, 13-16), who compare the level of real GDP per worker between 1955 and 1990 in several OECD countries and compare growth rates to central bank independence.[4] While a simple comparison between growth in output per worker and central bank independence reveals a slightly negative relationship, this fails to take convergence effects into account. The countries in the study with more independent central banks also tended to have higher output per worker levels in the year the study began. Because of diminishing returns in output from capital, the countries with less capital per worker at the beginning of the observed period would be expected to experience faster rates of "catch up" growth than those countries where capital per worker was already relatively high. Adjusting for convergence effects by holding the level of initial output per worker constant, the relationship between growth in output and central bank independence is positive. One possible explanation for this positive relationship is that, since protecting a central bank from political influence makes for more predictable monetary policy, strengthening central bank independence reduces the risk premia in real interest rates and so positively affects the real economy (Alesina and Summers 1993, 152). It is also likely that markets simply work more efficiently in the low inflation environment nurtured by independent central banks, high inflation environments containing greater uncertainty and price distortions.

Thus, the conclusion reached is that strengthening the independence of central banks is the advisable course of action. If monetary policy is left in the hands of elected politicians, who are beholden to public opinion, short term political expedience will be pursued to the detriment of long term economic health. Empirical evidence supports this reasoning: countries whose central banks enjoy greater degrees of independence experience lower and more stable inflation rates, and, rather than accomplishing this at the expense of real economic output, may actually achieve greater rates of growth than they otherwise would.

[4] The set of countries is the same as in Alesina and Summers (1993).

References:

Alesina, Alberto, and Lawrence H. Summers. 1993. "Central Bank Independence and Macroeconomic Performance: Some Comparative Evidence." *Journal of Money, Credit and Banking* 25 (2): 151-162. http://www.jstor.org.proxy2.lib.uwo.ca/stable/pdfplus/2077833.pdf?acceptTC=true

De Long, Bradford, and Lawrence H. Summers. 1992. "Macroeconomic Policy and Long-Run Growth." *Federal Reserve Bank of Kansas City Economic Review* 77 (4): 5-30. http://www.kansascityfed.org/PUBLICAT/ECONREV/EconRevArchive/1992/4Q92long.pdf

Grilli, Vittorio, Donato Masciandaro, Guido Tabellini, Edmond Malinvaud, and Marco Pagano. 1991. "Political and Monetary Institutions and Public Financial Policies in the Industrial Countries." *Economic Policy* 6 (13): 342-392. http://www.statistica.unimib.it/utenti/dallapellegrina/grilli_masciandaro_tabellini_91.pdf

Jacome, Luis, and Francisco Vazquez. 2008. "Is There any Link between Legal Central Bank Independence and Inflation? Evidence from Latin America and the Caribbean." *European Journal of Political Economy* 24 (4): 788-801.

Laidler, David, and William Robson. 2004. *Two Percent Target: The Context, Theory and Practice of Canadian Monetary Policy since 1991*.Toronto: C.D. Howe institute.

Explaining the High Profitability of China's Major State-Owned Banks

Jie Ren

I. Introduction

After the outbreak of the global financial crisis in 2007, China's banking sector emerged as one of few winners, with its major state-run banks posting record profits, besting their peers in the developed economies in terms of market capitalization, and even topping the Fortune 500 list. In fact, the Big Five[1] state-run banks were so profitable that the Chinese Prime Minister, Wen Jiabao, openly accused them of "making profits far too easily".[2]

The extraordinary profitability in itself, however, is not a bad thing. If it is the result of improved governance and advanced risk management due to decades of reforms in the banking sector, it may be just a reflection of the increased competitiveness and efficiency of the banks as financial intermediation. If, on the other hand, the outsized profits are due to other factors, such as "repressed" interest rate policy, significant entry barriers and unfair competition, this high profitability may have totally different meaning and policy implications. Thus it is imperative to have a clear understanding of the real source and nature of the exceptional profitability in China's banking sector.

This paper seeks to understand this phenomenon by studying the relationship between the profitability of major Chinese state-run commercial banks, especially the Big Five banks,[3] and various external and internal factors. Section II provides a brief overview of China's banking sector, including its historical development, main players and major reform policies. It also seeks to identify typical characteristics of the industry by comparing its recent performance to that of international peers in both developing and developed economies. In Section III, which focuses on bank performance in developing countries, I review some studies on the determinants of bank performance in Malaysia, Thailand and Tunisia. The purpose of this section is to provide some methodological background and establish the proper context on which to base and better evaluate the performance of Chinese banks. Section IV identifies the major factors that explain the high profitability of Chinese state-owned commercial banks (SOCBs) and discusses their relevance for future policymaking. Each factor is examined using both qualitative and quantitative analysis. In the last section, I summarize the major findings of my analysis

[1] The Big Five Banks are the Bank of China (BOC), the Agriculture Bank of China (ABC), the Construction Bank of China (CBC), and the Industrial and Commercial Bank of China (ICBC), the so-called ''Big Four", plus the Bank of Communication.

[2] Barboza, D., 'Wen calls China banks too powerful', The New York Times, April 3, 2012 http://www.nytimes.com/2012/04/04/business/global/chinas-big-banks-too-powerful-premier-says.html?_r=1

[3] As of 2010, the Big Five banks account for 60% of total commercial banking assets (Walter and Howie 2011).

and their policy implications, and I provide policy suggestions that may lead to a more competitive and efficient banking system.

II. The Chinese Banking System

Historical Development

China's banking system has come a long way. Until 1978, there was only one bank, the Peoples' Bank of China (PBoC), which handled virtually all banking activities as a department of the Ministry of Finance (MOF). It suffered from numerous problems: there was virtually no professional staff in the bank, it was organized along the lines of the administrative system, interest rates were fixed, and its lending decisions were dictated by the MOF (Walter and Howie 2011).

Reform in the banking sector began with the Big Four banks being removed from the PBoC during the period from 1979 to1984. In 1983, the PBoC was designated as the central bank of China. At that time, however, the central bank did not play an important role, since the local Party committee, rather than the central government, controlled the key management of the banks. This arrangement soon led to a lending spree that resulted in inflation and corruption in 1989, which caused Beijing to abandon this Soviet banking model in favor of the American one (Walter and Howie 2011). From 1992 to 2005, under the leadership of Jiang-Zhu, the pace of financial reform in China accelerated. In 1990, two stock exchanges were set up in Shenzhen and Shanghai to facilitate the financial intermediation process, and three specialized ‘‘policy” banks were established in 1994 in an effort to reduce the commercial banks' burden with respect to financing state-directed trade and development projects (Zhang 2007).

The rapid development of the financial industry, however, caused a number of problems. The banking sector soon went through a major lending and nonperforming loans cycle as aggressive directed lending to industry led to massive nonperforming loans. The real estate investment craze in Hainan province also went out of control during the early 1990s, causing concerns about the sustainability and stability of the financial system. As a result, the Party established a broad reform agenda in 1993, recognizing the need to allow banks to operate on a commercial footing. Since then, steps have been taken to gradually implement the reform agenda. Major state-owned banks were recapitalized, bad assets were expunged and moved to Asset Management Companies (AMC), bank supervision was revamped, and foreign strategic investors were introduced (Feyzioğlu 2009).

As of 2010, there were about 3,769 financial entities in China with 196,000 outlets and nearly three million employees. Total financial assets reached ¥128 trillion RMB or US $19.4 trillion, making the Chinese financial industry one of the largest in the world.

Pre- and Post-Crisis Performance

To measure a bank's profitability, researchers typically use two accounting metrics: return on equity (ROE) and return on asset (ROA). ROA reflects the profit earned per dollar of assets and is therefore a measure of management's ability to utilize the bank's financial resources to generate profits. ROE, on the other hand, represents the profit earned on every dollar invested in the firm's equity (Sufian and Habibullah 2009). Compared to ROE, ROA may be a better measure of a bank's profitability, since it is not affected by the capital structure of the bank, while ROE may be subject to distortion caused by high leverage. To measure other aspects of a bank's performance, analysts also use indicators such as growth of total assets, cost income ratio and market capitalization.

Based on the standard financial indicators discussed above, Chinese banks were doing extraordinarily well in both the pre- and post-crisis periods. According to Feyzioğlu (2009), the financial crisis that originated in 2007 did not have a noticeable impact on the Chinese banking sector, with the ROA of major Chinese banking institutions reaching 1.1 percent, much higher than banks in developed countries. Although the non-performing loan (NPL) ratio was higher than that of developed countries, it was lower than that of other developing countries such as Russia and Brazil. (See Table 1)

Table 1. Pre-Crisis Performance

Table 1. Financial Performance of Largest Banks 1/
(In percent, 2007)

	Canada	France	Germany	Italy	Japan	UK	US	HK SAR	Brazil	China	India	Russia
Return on average assets	1.0	0.3	0.5	0.7	0.4	0.6	0.7	1.6	2.5	1.1	1.0	2.2
Net interest margin	1.8	0.3	0.8	2.4	1.0	1.0	3.0	2.0	10.7	3.0	2.7	4.3
Capital asset ratio	12.1	9.9	13.1	9.7	19.0	11.0	11.2	12.8	16.0	13.4	13.5	27.2
NPL ratio	0.5	2.6	3.5	5.3	1.7	1.7	0.8	0.5	8.3	4.2	3.0	8.8
Interest earned as a share of income	72.0	73.8	79.6	78.6	42.2	79.0	76.6	83.5	55.0	89.7	78.7	65.0
Loans/Deposits	57.9	45.3	45.0	104.4	52.7	52.7	78.2	60.8	64.6	56.1	69.4	98.5
Interest expenditure as share of total	60.0	47.8	68.7	54.4	45.5	65.0	53.6	68.5	49.6	50.7	67.7	45.8
Labor cost as share of total	22.1	14.2	17.2	25.7	n.a.	21.4	22.5	13.2	21.3	15.5	14.2	17.3
Overhead cost as share of total	17.9	9.6	14.1	19.9	n.a.	17.7	25.1	18.4	32.3	30.0	18.1	36.9

1/ Covers largest five banks, except Brazil, which covers four banks, and China, which includes ICBC, CCB, BoC, and BoCOM.

Source: Feyzioğlu, 2009, p. 6

In fact, profitability strengthened even more in 2008. According to *The Banker* ranking, the Big Five state-owned banks became globally dominant in terms of their size and profitability from 2008 to 2010, and contributed one fifth of global banking profits in 2010.[4] Most notably, ICBC was the most profitable bank in the world for three consecutive years (See Table 2).

[4] *The Banker*, July 2011, p. 143

Table 2. Post-Crisis Performance

Table 1
Global top 10 banks by pre-tax profit 2007-2010[3]

2007				2008			
	Bank	**Country**	**US$ million**		**Bank**	**Country**	**US$ million**
1	Bank of America	USA	31,973	1	**ICBC**	**China**	**21,260**
2	Citigroup	USA	29,639	2	**CCB**	**China**	**17,520**
3	HSBC Holdings	UK	22,086	3	Banco Santander	Spain	15,825
4	JP Morgen Chase	USA	19,886	4	**BOC**	**China**	**12,620**
5	Royal Bank of Scotland	UK	18,033	5	BBVA	Spain	9,640
6	Crédit Agricole	France	14,060	6	HSBC Holdings	UK	9,307
7	Barclays Bank	UK	14,009	7	Barclays Bank	UK	8,859
8	BNP Paribas	France	13,921	8	**ABC**	**China**	**7,659**
9	Mitsubishi UFJ Financial Group	Japan	12,824	9	UniCredit	Italy	6,952
10	Wells Fargo	USA	12,745	10	Royal Bank of Canada	Canada	6,077
2009				**2010**			
	Bank	**Country**	**US$ million**		**Bank**	**Country**	**US$ million**
1	**ICBC**	**China**	**24,494**	1	**ICBC**	**China**	**32,528**
2	**CCB**	**China**	**20,316**	2	**CCB**	**China**	**26,448**
3	Goldman Sachs	USA	19,826	3	JPMorgan Chase & Co	USA	24,859
4	Barclays	UK	18,869	4	**BOC**	**China**	21,463
5	Wells Fargo & Co	USA	17,606	5	HSBC Holdings	UK	19,037
6	Banco Santander	Spain	16,951	6	Wells Fargo & Co	USA	18,700
7	**BOC**	**China**	**16,319**	7	**ABC**	**China**	**18,230**
8	JPMorgan Chase & Co	USA	16,143	8	BNP Paribas	France	17,406
9	BNP Paribas	France	12,222	9	Banco Santander	Spain	16,079
10	Itaú Unibanco Holding SA	Brazil	11,521	10	Goldman Sachs	USA	12,892

Source: The Banker

Source: Löchel and Li, 2011, p. 1.

Löchel and Li (2011) conduct a more comprehensive analysis, comparing the Big Five Chinese SOCBs with the twenty largest international banks according to total assets for the period 2003-2009. (See a summary of the key statistics in Table 3) They find several interesting characteristics of the large Chinese state-own banks. First, the Big Five banks, with an average ROA of 0.81% and an average ROE of 12.91%, have been consistently more profitable than their international counterparts, whose average ROA and ROE are 0.41% and 8.17%, respectively. Second, the share of bad loans in the Chinese banks (8.11%) is significantly higher than the international average of 3.01%. When the bad loans were removed to Asset Management Companies (AMC), however, the bad loan ratio decreased dramatically from 17.6% in 2003 to 1.86% in 2009. Third, corporate lending makes up the majority (81.03%) of the loan portfolios of Chinese banks, whereas their international peers have much more balanced portfolios, with only 37.6% of total loans being in the corporate sector. Fourth, the Chinese Big Five banks have an impressive cost advantage. Their average cost income ratio is about 42.29%, 40 percent lower than their international competitors. Last, these banks benefit from a high asset growth rate of 18.36%, while the average growth rate for the international banks is 9.16%.

Table 3. Comparison of the Big Five SOCBs and International Peer Banks

Table 3
Key financial indicators of top twenty banks (average value from 2003 to 2009)

Bank name	Size	Profitability		Lending business		Loan portfolio			Cost control		Business diversification	Growth	Capital strength		
	Total assets (million USD)	ROAA (%)	ROAE (%)	Net interest margin (%)	Impaired loans/ gross loans (%)	Residential mortage loans/ gross loans (%)	Other customer retail loans/ gross loans (%)	Corporate and commercia l loans/ gross loans (%)	Cost income ratio (%)	Personnel expenses/ total assets (%)	Non-interest income/ gross revenue (%)	Growth of total assets (%)	Tier 1 ratio (%)	Total regulatory capital ratio (%)	Equity/ total assets (%)
"Big five" Chinese banks															
1 ICBC	1,037,671	0.85	4.51	2.70	8.63	14.45	3.75	81.80	36.59	0.50	11.93	17.32	9.47	11.35	0.89
2 CCB	798,040	1.11	29.08	3.00	3.09	15.98	5.42	78.60	38.40	0.62	11.99	19.07	9.34	11.43	5.87
3 ABC	748,382	0.39	-0.81	2.23	19.64	10.31	5.75	83.93	53.61	0.68	19.71	17.33	7.85	9.75	0.50
4 BOC	769,076	0.91	14.28	2.39	5.43	19.06	5.10	75.84	40.58	0.54	20.00	14.00	9.29	11.29	6.43
5 BoCom	258,870	0.78	17.49	2.71	3.77	9.92	5.10	84.98	42.27	0.42	12.29	24.07	8.36	11.29	4.91
Average	***722,408***	***0.81***	***12.91***	***2.61***	***8.11***	***13.94***	***5.03***	***81.03***	***42.29***	***0.55***	***15.18***	***18.36***	***8.86***	***11.02***	***3.72***
International peers															
1 Deutsche Bank	2,532,690	0.22	10.25	0.63	1.75	21.49	n.a.	21.41	89.48	0.67	48.75	1.67	9.95	12.55	2.03
2 BNP Paribas	2,345,667	0.42	11.77	0.85	4.37	n.a.	n.a.	n.a.	62.68	0.64	55.86	13.43	8.04	11.36	3.02
3 Barclays	2,102,636	0.48	18.32	0.88	2.68	n.a.	47.70	n.a.	60.39	0.67	55.08	18.90	8.48	12.55	2.78
4 Royal Bank of Scotland	1,867,245	0.43	7.43	1.26	2.22	n.a.	n.a.	n.a.	65.00	0.61	45.20	23.10	9.07	14.10	4.37
5 Société Générale	1,377,227	0.35	10.30	0.67	4.26	22.49	20.81	35.30	67.77	0.83	70.54	5.20	8.22	11.10	3.08
6 ING Bank	1,267,358	0.29	11.03	1.18	1.54	52.16	4.29	37.91	65.84	0.57	27.96	1.93	7.86	10.98	2.72
7 UniCredit	1,262,059	0.50	9.08	1.66	6.52	n.a.	n.a.	n.a.	61.54	0.86	39.12	3.96	7.18	10.74	5.74
8 Banco Santander	1,227,051	0.91	15.24	1.88	1.62	n.a.	56.21	26.43	48.93	0.70	41.33	11.04	8.18	12.92	5.52
9 Bank of America	1,131,085	1.11	12.87	2.91	1.43	n.a.	8.89	19.85	54.45	1.13	44.12	15.19	8.77	11.38	9.01
10 RBS Holdings (former ABN AMRO)	1,073,879	0.20	8.66	0.83	1.85	0.71	35.49	41.88	97.06	0.80	111.13	-5.50	11.80	14.95	2.87
11 HSBC Bank	984,563	0.63	13.68	1.13	1.65	41.35	13.58	48.10	59.95	0.79	57.42	23.46	7.88	11.60	4.23
12 Bank of Scotland	954,861	-0.28	-11.71	1.18	4.80	52.85	4.37	41.48	49.85	0.35	31.06	16.18	6.62	10.82	2.82
13 Citibank	949,554	0.73	9.20	3.45	2.34	n.a.	27.12	24.27	64.81	1.29	35.00	13.93	9.38	13.37	7.96
14 Credit Agricole	846,972	0.09	2.15	-0.05	2.36	n.a.	n.a.	n.a.	109.29	0.38	101.82	12.42	9.22	10.48	1.87
15 BNP Paribas Fortis-Fortis Bank	843,574	-0.36	-6.77	0.89	2.68	25.08	3.49	41.62	80.98	0.54	39.28	-14.73	9.40	13.88	3.22
16 Intesa Sanpaolo	803,770	0.77	8.31	2.16	5.60	n.a.	n.a.	71.79	62.82	1.03	34.78	4.10	7.33	10.33	9.04
17 Commerzbank	768,572	-0.02	-0.35	0.79	4.61	17.86	n.a.	n.a.	79.51	0.54	44.83	11.57	8.14	12.67	2.36
18 Lloyds TSB Bank	673,396	0.55	12.51	1.79	3.57	48.50	13.83	34.92	61.27	0.80	44.25	15.36	9.25	11.55	3.37
19 Natixis	661,328	0.06	1.04	0.43	2.36	n.a.	n.a.	n.a.	106.54	0.56	43.06	1.48	8.72	10.84	3.48
20 Banco Bilbao Vizcaya Argentaria	619,262	1.10	20.41	2.37	1.93	n.a.	38.80	44.93	46.65	0.91	41.20	10.52	7.88	12.17	5.00
Average	***1,214,637***	***0.41***	***8.17***	***1.34***	***3.01***	***31.39***	***22.88***	***37.68***	***69.74***	***0.73***	***50.59***	***9.16***	***8.57***	***12.02***	***4.22***

Source: Bankscope

Source: Löchel and Li, 2011, p. 15

III. Banking Performance in Developing Countries

An extensive amount of literature examines the performance of the banking sector in the developed countries, but few studies have looked at the determinants of bank performance in developing economies. This section reviews briefly some of these studies in order to provide some background information on banking reform experiences and to show how bank performance was evaluated in other emerging markets. Guru, Staunton, and Balashanmugam (2002) investigate the determinants of bank profitability in Malaysia by focusing on a sample of 17 commercial banks during the period of 1986-1995. They divide the potential determinants into two categories, namely internal factors, such as liquidity and expense management, and external factors, such as ownership and firm size. They find that expense control contributes the most to high bank profitability, while a high interest ratio was associated with low bank profitability.

Chantapong (2005) studies the performance of domestic and foreign banks in Thailand from 1995 to 2000. The results indicate that foreign bank profitability is higher than the average profitability of domestic banks, although the gap between foreign and domestic bank profitability has closed in the post-crisis period, suggesting that the financial restructuring program has yielded some positive results.

Ben Naceur and Goaied (2008) examine the impact of bank characteristics, financial structure, and macroeconomic conditions on Tunisian banks' net interest margin and profitability during the period from 1980 to 2000. They find that banks with a relatively large amount of capital and higher overhead expenses tend to enjoy a higher level of net interest margin and profitability, while a bank's size is negatively related to its profitability. They also find that stock market development has a positive impact on banks' profitability during the period under study. In addition, their findings suggest that private banks are relatively more profitable than their state-owned counterparts.

IV. Explaining High Profitability of Chinese State-Owned Banks

As discussed above, China's banking sector, especially the Big Five state-run banks, has been highly profitable despite considerable inefficiency within the banking system. Understanding the sources of such high profitability is crucial, since their characteristics have significant implications for the direction of future banking reform policies. The unusually large profits enjoyed by Chinese banks can be explained by a set of distinct but inherently coherent factors: financial repression, market structure, and personnel costs advantage. Each of these factors will be discussed in detail this section.

Financial Repression

Financial repression is a term first used by McKinnon (1973) to refer to a set of policies typically used in developing countries that regulate interest rates, set high reserve requirements on bank deposits and direct the allocation of resources in the economy. A more precise definition is given in Reinhart (2012, 38):

> Financial repression includes directed lending to the government by captive domestic audiences (such as pension funds or domestic banks), explicit or implicit caps on interest rates, regulation of cross-border capital movements, and (generally) a tighter connection between government and banks, either explicitly through public ownership of some of the banks or through heavy 'moral suasion'. …Financial repression is also sometimes associated with relatively high reserve requirements (or liquidity requirements), securities transaction taxes, prohibition of gold purchases (as in the United States from 1933 to 1974), or the placement of significant amounts of government debt that is nonmarketable. A large presence of state-owned or state intervened banks is also common in financially "repressed" economies.

China's financial policies fit this description well. For example, the deposit and lending rates in China are partially controlled by the central bank: PBoC currently sets a mandatory depositing rate cap of 3.5% and a lending rate floor of 6.56%, essentially guaranteeing a net interest margin of 3.06% for the banks, which is significantly higher than G7 countries (Löchel and Li 2011). As Lardy (2008) points out, very low deposit rates and lending rates have resulted in an implicit tax on net lenders. Since households are major net savers in China, the redistribution has been, to some extent, from households to corporations, but even more, to the state. According to his study, one of the most significant gains for the state has been that the cost of sterilization has been kept relatively low, thus allowing for a significantly undervalued RMB during most of the past decade.

Löchel and Li (2011) also reach the conclusion that the Big Five Chinese banks' outperformance of their international counterparts in asset return is caused, to a large extent, by the high interest rate margin realized "in the current environment of guaranteed margin system and isolation from the competition on the international financial markets due to foreign capital control" (Löchel and Li 2011, 20). One may argue that the "windfall" profits in the banking sector are the indirect consequence of the government's deliberate intention to keep RMB undervalued. On the other hand, evidence suggests that there has been gradual interest rate liberalization since 1996. For instance, the interbank lending rates and interbank repo rates were liberalized in 1997, and deposit rates were partially relaxed for large amounts of local currency in 2000 (Löchel and Li 2011). In addition, the Shanghai Interbank Offered Rate (Shibor) was set up in 2007, a notable step towards a market-oriented interest rate system.

The interest rate reform is closely related to the loosening of foreign capital control in China. As the external pressure for a higher valuation of RMB grows and the Chinese economy gradually adjusts its structural imbalances, capital account controls may eventually be eliminated, offering greater room for further interest rate liberalization. This, however, may not be good news for the large banks, since the current protective environment does not provide them with enough incentives to develop internal competitive advantages. If the interest rate were to be liberalized, they may find themselves unable to compete with other foreign or joint stock commercial banks.

Market Structure

By simulating a stressed scenario in which the Big Five's average margin is reduced from 2.62% to the international level of 1.24%, Löchel and Li (2011) find that their ROA would decrease from 0.81% to 0.41%, but still be on the same level as the international peers, suggesting that high margin advantage is not the sole source of the high profitability of the Big Five Banks. Another factor may be the market structure of the banking industry, which has become increasingly complex over the years.

The structure of the banking industry can be analyzed using a number of different techniques. As Table 3 shows, state-run banks in China still constitute the dominant force in the banking system by owning more than half of total assets. A more sophisticated approach is to look at the Herfindahl-Hirschman Index (HHI), which sums the squares of the market shares of the firms in the market, ranging from 0 to 1, and thus serves as a measure of the level of market concentration Feyzioğlu (2009). From international experience, an index above 0.18 suggests that the market is highly concentrated. According to Feyzioğlu's study, the adjusted HHI for China's banking industry, which includes banks that compete in similar markets such as state-owned commercial banks and joint stock commercial banks, is 0.11, indicating a fairly high concentration level.

In addition, there has been no entry or exit among the large or medium size banks in the 1999-2009 period. Moreover, despite the introduction of foreign banks decades ago, their share of the market has remained around 2%, reflecting a difficulty in expanding their presence in China. The existence of significant entry barriers to the banking industry is obvious. In fact, Walter and Howie (2011) argue that the level of market concentration is much higher than the HHI index suggests. According to them, despite the different names, locations and categorizations of Chinese banks, most of them have significant state ownership, and all Chinese banks are used basically as utilities providing unlimited capital to state-owned enterprises, or rather Party-owned enterprises, for the purpose of improving and strengthening "the economy inside the system (tizhinei jingji 体制内经济)", which they believe has been the goal of "every reform effort undertaken by the Party since 1978" (Walter and Howie 2011, 8). This commonality among the banks creates incentive for them to maintain the status quo and compete against non-state-owned entrants as a group, rather than against each other for a greater share of the profits.

Table 4. Banking System Overview

Table 6. The Banking System, 2008
(In percent of total)

	Assets	Loans	Deposits
State commercial banks	48.4	43.2	53.3
Joint stock commercial banks	18.4	19.4	17.5
Policy banks	9.0	15.6	0.9
City commercial banks	6.5	7.0	7.5
Urban credit cooperatives	0.1	0.1	0.2
Rural credit cooperatives	7.8	6.0	9.7
Foreign banks	2.1	2.5	0.9
Finance companies	1.5	1.9	1.7
Postal saving bank	3.4	1.4	4.9

Source: Monetary survey.

Source: Feyzioğlu, 2009, p. 22

Several studies provide further evidence of the lack of competition and efficiency among China's state-run banks. Berger, Hasan and Zhou (2007) analyze the efficiency of Chinese banks over 1994–2003, and find that state-owned banks such as the Big Five are by far the least efficient, foreign banks are most efficient, and minority foreign ownership is associated with significantly improved efficiency. Fu and Heffernan (2007) investigate the relationship between market structure and performance in China's banking system from 1985 to 2002, a period when this sector was subject to gradual but notable reform, and suggest that, on average, most banks were operating below scale efficient levels and that the reforms had little impact on the structure of China's banking sector, while the ‘‘joint stock" banks became relatively more efficient. In addition, after studying the Big Four state commercial banks during the period 1994–2001 in China, Ho (2012) finds no clear evidence that the pricing of banking services has become more competitive after the reform.

It seems clear that lowering the entry barrier and opening up the banking industry to private and foreign capital can increase the level of competition, the efficiency of banks in allocating scarce financial resources, and the general level of innovation and profitability among banks. However, this would certainly hurt the vested interests built around major state-owned banks, whose power cannot be underestimated. For instance, although Wen repeatedly and openly has called for bank reform, no meaningful action has been taken so far.[5]

Personnel Cost Advantage

To understand the high profitability of Chinese banks, Löchel and Li (2011) compare the financial data of Chinese banks with a large sample of international peers from Asian, Europe and North America for the period of 2003-2009, and find that Chinese banks are very good at controlling costs. According to their study, despite low efficiency, the top Chinese banks enjoy a cost income ratio of 42.29%, which is 40% lower than the international average, and a personnel expense to total assets ratio of 0.55%, 30% lower compared to 0.73% for the international peer banks. For instance, as of 2010, the largest Chinese bank, ICBC, had 397,339 employees with total personnel expenses of US $10,515 million, compared to Deutsche Bank with 102,062 employees costing US $16,931 million; the average wage at Deutsche Bank is thus more than six times that of ICBC (Löchel and Sottocornola 2011). The favorable lower labour cost, however, is not the result of better operational efficiency, as is evident in Feyzioğlu (2009) and Fu and Heffernan (2007).

Löchel and Li (2011) further demonstrate the importance of lower labour cost to the profitability of Chinese banks by testing the Big Five banks' profitability in a stressed scenario Their analysis shows that, assuming a net interest margin of international average level, an increase of personnel costs ratio by 30% would reduce the banks' asset return dramatically from 0.81% to 0.34%, which is far below the international peer average of 0.41%. Given that China's population is ageing rapidly and that its government aims to increases both minimum and average wages significantly in the next five to ten years, the assumption of a 30% increase in labour costs is not unreasonable. The results of the study, consequently, cast doubt on the sustainability of the high profitability of SOCBs in the long term.

[5] Barboza, D. ‘Wen calls China banks too powerful', The New York Times, April 3, 2012 <http://www.nytimes.com/2012/04/04/business/global/chinas-big-banks-too-powerful-premier-says.html?_r=1>

V. Conclusion

After three decades of banking reforms, the Chinese state-run banks have become the dominant financial force in one of the world's largest economies. They have also become much more efficient and profitable than before. Their high profitability however, is rooted in guaranteed high net interest margin, lower personnel cost advantage and a oligopolistic market structure with strong protective restrictions – factors determined externally by government policy and the so called "demographic dividend".

As indicated in the 12th Five Year Plan, gradual interest rate liberalization is likely to continue and the average wage is expected to double in the next ten years. In addition, the rapidly ageing population in China may cause a labour shortage in the not-so-distant future and further increase the labour costs for the banks. Given these challenges, whether the major SOCBs can sustain their current level of profitability remains questionable, since the current regulatory environment creates little incentive for them to improve their efficiency and competitiveness. Moreover, the inherently political nature of the state-owned banks may also prevent them from becoming truly market-oriented public companies.

To help improve the financial intermediation of the current banking system, policy makers can consider lifting the ceiling on deposit rates. Doing this could facilitate the movement of deposits from large to smaller banks, which are more efficient at utilizing these deposits. It might also help lower the level of market concentration and encourage competition among the banks. Unless the large banks develop their own internal competitive advantages, the high profitability they enjoy now is not likely to last in the long run.

References:

Berger, A., I. Hasan, and M. Zhou. 2009. "Bank Ownership and Efficiency in China: What will Happen in the World's Largest Nation?" *Journal of Banking & Finance* 33 (1): 113-30.

Ben Naceur, S., and M. Goaied. 2008. "The Determinants of Commercial Bank Interest Margin and Profitability: Evidence from Tunisia." *Frontiers in Finance and Economics* 5 (1): 106–130.

Chantapong, S. 2005. "Comparative Study of Domestic and Foreign Bank Performance in Thailand: The Regression Analysis." *Economic Change and Restructuring* 38 (1): 63–83.

Feyzioğlu, T. 2009. "Does Good Financial Performance Mean Good Financial Intermediation in China?" IMF Working Paper 09/170, International Monetary Fund, Washington DC.

Fu, X., and S. Heffernan. 2009. "The Effects of Reform on China's Bank Structure and Performance." *Journal of Banking & Finance* 33 (1): 39-52.

Guru, B. K., J. Staunton, and B. Balashanmugam. 2002. “Determinants of Commercial Bank Profitability in Malaysia.” Working Paper, Multimedia University.

Ho, Chun-Yu. 2012. "Market Structure, Welfare, and Banking Reform in China." *Journal of Comparative Economics* 40 (2): 291-313.

Lardy, N. 2008. “Financial Repression in China.” Policy Brief PB08-8, Peterson Institute of International Economics, Washington, DC.

Löchel, H., and H. X. Li. 2011."Understanding the High Profitability of Chinese Banks." Frankfurt School of Finance and Management Working Paper No. 77, Frankfurt, Germany.

Löchel, H., and M. Scottocornola. 2011. “Evolving Banking Business Models: The Case of Bank of China and Deutsche Bank.” EU-CHINA BMT Working Paper Series No. 14.

McKinnon, R. I., 1973, *Money and Capital in Economic Development*, Washington, DC: The Brookings Institution.

Reinhart, C. 2012. "The Return of Financial Repression." *Financial Stability Review* 16: 37-48.

Sufian, F., and M. S. Habibullah. 2009. “Bank Specific and Macroeconomic Determinants of Bank Profitability: Empirical Evidence from The China Banking Sector.” *Frontiers of Economics in China* 4 (2): 274-291.

Walter, C. E., and F. J. T. Howie. 2011. *Red Capitalism: The Fragile Financial Foundation of China's Extraordinary Rise*. Singapore: Wiley.

How Does Liquidity Risk Disturb Asset Prices? A General Equilibrium Approach

Dai Li

Winner of the Mark K. Inman Senior Essay Prize, 2012

Abstract

Liquidity risk was conspicuous in the recent financial market turbulence. This paper presents a liquidity risk model in which two financial institutions trade an illiquid risky asset. The model develops explicit liquidity demand and supply curves along with analytical solutions, and it inherently generates two types of general equilibrium – liquid and illiquid. Liquidity risk manifests in the illiquid equilibrium to depress the asset price to deviate from the fundamental value. In turn, the model shows that riskier assets have thinner liquidity supply and heavier liquidity demand. The model is able to analyze precautionary hoarding, runs on financial institutions, and loss spiral. The model suggests that hoarding liquidity in turmoil is an effective way for a financial institution to earn profit and also maintain a solid financial condition. Bank-run is an important externality of deteriorating market condition caused by hoarding. It can motivate financial institutions to hoard less liquidity. Lastly, financial institutions should be relieved from marking-to-market to prevent loss spiral, as it may lead to illiquidity and, eventually, to insolvency.

I. Introduction

During the liquidity and credit crunch in 2007 and 2008, the U.S. corporate bond index spread increased to five times its average pre-crisis level, rising from roughly 90 basis points between 2004 and 2006 to a peak of 450 basis points in 2008. Garcia and Prokopiw (2009) used a structural credit-risk model to explain the spread by two factors – credit risk and liquidity risk. They concluded that the increase in the model-implied credit risk explained only a small portion of the spread, most of which was attributed to liquidity risk.

In this paper, I construct a theoretical liquidity risk model incorporating the two characteristics of liquidity risk and their impact on asset prices. The model inherently generates two types of general equilibrium – liquid equilibrium and illiquid equilibrium. The liquid equilibrium is characterized by assets trading at fundamental values. In an illiquid equilibrium, however, asset prices deviate a great deal from the fundamentals and are very sensitive to marginal changes in market liquidity condition.

This model captures two generalized characteristics of liquidity risk implied by Garcia and Prokopiw's study (2009). The first one is that, without any change in the fundamental value of an asset, liquidity risk itself can greatly disturb prices. The second characteristic is that liquidity risk does not manifest itself in normal times; however, it depresses asset prices severely in a distressed market.

Next, I will introduce briefly the theoretical framework, the primary results, and the applications of the model on three liquidity related issues. First, the theoretical framework of the model is depicted in the following. Suppose that there are only two financial institutions (called Bank A and Bank B) in the financial market. Two banks are required to maintain their capital ratios (defined later) above a threshold with a very high probability. Such regulation poses a problem for Bank A, which has experienced an idiosyncratic shock. It needs to sell an illiquid risky asset to reduce the uncertainty of its capital ratio. Liquidation at a fire-sale price may be very costly for Bank A when the market is thin. Bank B, as the only potential buyer in the market, sees this as an opportunity to make profit via buying mispriced assets. Nevertheless, Bank B is also subject to regulations on its capital ratio limiting its ability to inject liquidity into the market to earn profit. As a result, Bank A attempts to minimize the loss by selling only what is necessary at all given prices. The set of Bank A's choices at given prices forms the liquidity demand curve. In compliance with regulations, Bank B utilizes all capital available to maximize profit. Solving Bank B’s problem produces the liquidity supply curve in this market. The general equilibrium occurs when the equilibrium trading price solves both banks’ problems and the market clears.

The primary results of this model are analytical solutions for liquidity demand and supply functions. Unlike conventional supply curves, the liquidity supply curve is downward sloping because lower prices motivate financial institutions to purchase more assets. The regulation on both banks’ capital ratios pins down the position of the demand and supply curve. The relative position of demand and supply curve determines which type of equilibrium occurs. Asset prices in an illiquid equilibrium are very sensitive to liquidity condition mainly because both demand and supply curves are downward sloping.

In terms of the existence and uniqueness of equilibriums, I show that the illiquid equilibrium is unique if it exists under the condition that both banks face required thresholds (thresholds could be different for two banks) on their capital ratio with the **same** probability. In the comparative analysis section of this paper, I will show that riskier assets have thinner liquidity supply and heavier liquidity demand in a time of stress, which means the trading price will deviate from the fundamental value more severely.

The analysis of the model application sheds light on three liquidity related issues: precautionary hoarding runs on financial institutions, and loss spirals.

In the context of my model, Bank B can conduct precautionary hoarding by setting an overly conservative target capital ratio. A higher target ratio reduces liquidity supplied at any given price, i.e., shifts the supply curve downward. My model predicts that financial institutions will hoard liquidity to enhance profit by setting the target ratio as high as possible provided that an illiquid equilibrium occurs. The equilibrium situation deteriorates in the sense that asset prices have larger swings and Bank A suffers huge losses by liquidating all of its risky assets at the lowest acceptable price. Overall, my model suggests that, when facing a desperate seller (Bank A), hoarding is an effective way for the counterparty financial institution to generate profits while maintaining a solid financial condition.

The next application on bank runs captures a noticeable adverse externality of precautionary hoarding. When investors (depositors) are informed about the trade between two banks with a price far below the previously perceived fundamental value, they may mistakenly consider the plunge in Bank A's asset price as a decline in the fundamental value. If panicked investors collectively decide to withdraw investment (deposits), they run indiscriminately on both banks. Further, the lower the trading price, the more likely investors are to run. In the environment with bank run threat, although the decline of asset price is still a profit opportunity for Bank B, it also causes higher expected bank run loss on Bank B as the bank-run probability increases. In terms of the model setup, Bank B's objective changes to maximizing expected net profit rather than the trading profit. My model shows that Bank B is willing to hoard less liquidity and to purchase assets rationally at higher prices in the case where bank run is incorporated. To sum up, if market participants are aware of externalities of declining assets prices, the market liquidity position can be moderately eased to generate higher trading prices in equilibrium.

The final application of my model is on the study of loss spiral. Suppose that after selling a portion of risky assets in an illiquid market with a fire-sale price, Bank A must mark its remaining portfolio to the fire-sale price. The resulting write-down loss would immediately bring down Bank A's capital ratio below regulation threshold again. To be compliant with the regulation, it has to sell more portfolios at even lower prices. It is expected that with reiterated costly liquidations and write-down losses, Bank A's problem would quickly evolve from illiquidity to insolvency. Based on this expectation, financial institutions should be relieved from marking-to-market regulation, at least in the time of stress, to prevent liquidity problems from being transmitted to solvency problems. This is mainly because when a market lacks liquidity, market prices observed from sporadic trades of an asset do not necessarily reflect its fundamental price. My model shows that the market price in a turbulent time may include a large "liquidity risk premium" and it greatly deviates from the fundamental value of an asset.

This paper is structured as follows. Section II is a literature review focusing mainly on two papers that are closely related to my model. Section III builds up the detailed theoretical framework of the model. Section IV first describes the model setup and defines the liquid and illiquid equilibriums. The derivation of liquidity demand and supply functions and the discussion about the equilibrium condition is also included in Section IV. Section V provides numerical examples on liquidity supply and demand curves along with a comparative statics analysis to further illustrate the feature of the model. Section VI addresses how to apply the model to study the three liquidity related issues sketched above. Section VII concludes.

II. Literature Review

In the recent financial crisis, liquidity risk was noticeable in various financial markets. In the debt market, liquidity risk increased due to three factors: falling risk capital, rising repo haircut, and increased counterparty risk (Krishnamurthy 2010). In the money market, banks or investment banks that used off-balance-sheet vehicles faced funding liquidity risk because of the mismatch between the maturity of long-term investment and short-term borrowing (Brunnermeier 2009). In unsecured interbank money markets, Eisenschmidt and Tapking (2009) find that the market spreads have been largely attributable to liquidity risk since the start of the turmoil in 2007.

Cifuentes, Ferrucci, and Shin (2005) study the liquidity-triggered financial contagions using a common illiquid asset as the channel of contagion in a banking system. My model is similar to theirs in terms of the motivation of liquidation – complying with regulatory requirements or internal regulations. Based on regulatory provisions on banks' capital adequacy ratio, an idiosyncratic shock may force one bank to reduce its balance sheet by selling the common illiquid asset that is held by all banks to an external market. In Cifuentes, Ferrucci and Shin's model, the liquidity supply curve in the external market is assumed to be a downward sloping exponential function, so that the price tumbles if one bank is dumping the common illiquid asset. Thus, one bank's behavior may create downward pressure on all other banks' balance sheets, which possibly triggers a wave of liquidation by the other banks. There are two major differences between my model and theirs. First, in my model the transaction price is determined endogenously by two counterparties involved in the trade instead of an external market. Second, their model converts any marginal increase in liquidity demand into a decrease in the asset price. Conversely, in my model, the asset price is invariant to marginal change in supply or demand if the market is awash with liquidity.

My model employs the same measurement device for liquidity risk as Brunnermeier and Pedersen's model (2009). When the market price deviates from the fundamental value, the absolute value of the deviation is defined as the market illiquidity. In addition, they assume that the fundamental value follows a geometric Brownian motion with the volatility following an ARCH process. For simplicity, I assume that it follows a normal distribution, which is sufficient to demonstrate the excessive sensitivity against liquidity supply. One implication of their work is that if the fundamental volatility of an asset is high, then the asset has high market illiquidity. My model implies the same characteristic of illiquid assets.

My liquidity risk model differs from most of the existing literature. Most theoretical models are characterized by similar forms of liquidity shocks – mismatch between stochastic liquidity demand of depositors or consumers and the timing for illiquid investments to pay off (Allen and Gale 2000). Allen, Carletti, and Gale (2009) argue that if banks lack hedging tools, they may hoard liquidity because they face uncertain liquidity demand from depositors, which reduces efficiency in the use of capital. The study also theorizes that the inefficiency should be removed by central banks adopting open market operations. My model suggests that hoarding may be also an effective way for financial institutions to earn profit while maintaining a solid financial condition. Gorton and Huang (2006) justify that banking systems consisting of well-diversified big banks are less prone to idiosyncratic liquidity shocks. Tirole (2011) summarizes the interrelationships among illiquidity, market freezes, fire sales, contagion, insolvency, and bailouts. In terms of empirical studies, Chen, Lesmond and Wei (2007) analyze a comprehensive set of four thousand corporate bonds covering both investment grade and speculative grade bonds, and find that liquidity is a key determinant in yield spreads. Similarly, De Jong and Driessen (2006) find that corporate bond returns have significant exposures to fluctuations in Treasury bond liquidity and equity market liquidity. However, liquidity risk is a minor concern in the credit default swap (CDS) market, which is not surprising because CDS is inherently used to addressing credit risk (Longstaff, Mithal, and Neis 2005).

III. Theoretical Framework

Based on the simple theoretical framework depicted in the introduction, my focus here is on the problems faced by two banks and on the details of the banking regulations.

III.1 Regulatory Environment

In my model, the particular motivation behind Bank A's liquidation is regulation. The regulation stipulates that Banks must maintain their capital ratio (defined later) above a **target level** with a very high **probability**. I employ such a restriction because, usually in a financial crunch, banks must adjust equity capital to keep the probability of financial distress sufficiently low (Krishnamurthy 2010). The "capital ratio" in my model is similar to the capital adequacy ratio used in actual regulation. Although it is not calculated in exactly the same way as capital adequacy ratio, it imposes similar restrictions on financial institutions' behavior. Also, I assume that both Banks are in compliance with the regulation before the idiosyncratic shock hits Bank A.

III.2 Bank A's Problem

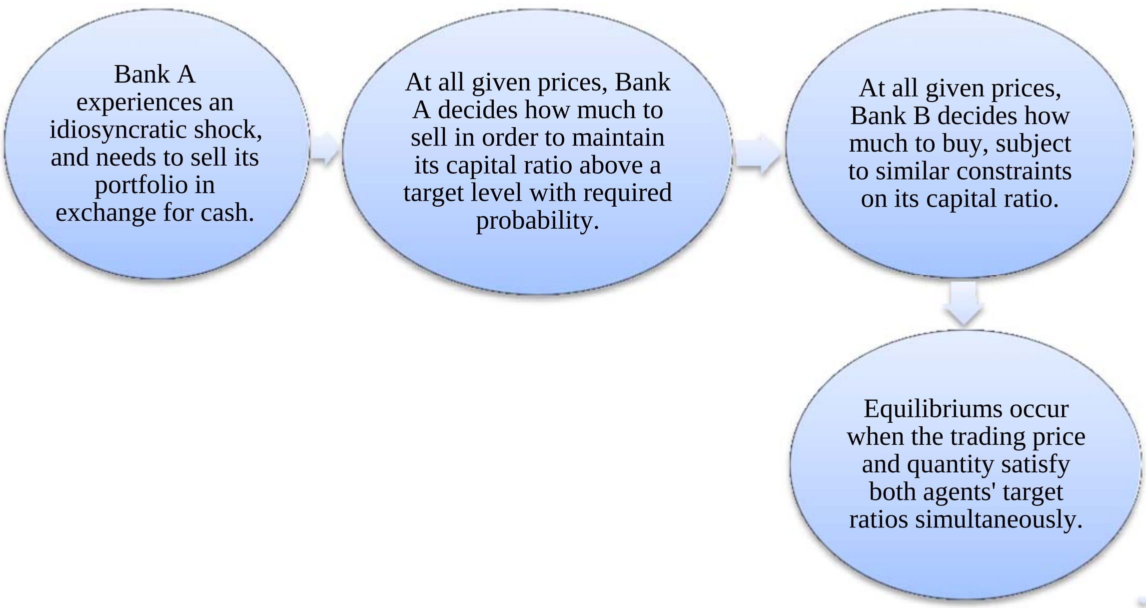
In my model, after the idiosyncratic shock, Bank A wants to lower the holding of its risky portfolio. This is simply because the probability of Bank A's capital ratio falling **below the target level exceeds the required probability**. Selling risky assets helps Bank A reduce the volatility of its capital ratio, and, in turn, the probability of violating the regulation decreases. However, liquidation at a fire-sale price (below fundamental value) is very costly. To minimize the liquidation loss, Bank A calculates the minimum amount of portfolio to sell at all given prices, which reveals the relationship between the liquidity demanded and prices.

III.3 Bank B's Problem

On the buy side of this financial market, Bank B, as the only potential buyer in this market, may want to buy Bank A's portfolio because it makes profit if the portfolio is sold below fundamental value. The regulations on Bank B's capital ratio limit its ability to inject liquidity into the market to earn profit. This is mainly because purchasing the risky asset, even at prices below fundamental value increases the volatility of Bank B's capital ratio. Therefore, Bank B faces a profit maximization problem. It calculates the maximum amount of the risky asset it can purchase at given prices under the constraint of regulations on its capital ratio.

III.4 Results

Solving Bank A's loss minimization and Bank B's profit maximization problems respectively generates liquidity demand and supply curve. The market clears when the liquidity supplied is equal to liquidity demanded at the equilibrium price. The following procedure summarizes how a financial institution may have to liquidate its portfolio in a distressed market.



The first step of modeling this trade is to develop explicitly the demand and supply curve, which in turn will directly determine the equilibrium. In the next section, I will first explain the setup of the model in detail in Section IV.1, the derivation of demand and supply function in Section IV.2, and the equilibrium conditions in Section IV.3.

IV. The Model

IV.1 Model Setup

Notation :

f_A	The fundamental value of Bank A's risky portfolio
e_{AA}	The quantity of Bank A's portfolio
c_A	Bank A's cash
OA_A	Bank A's other illiquid assets
r_A^*	Bank A's target capital ratio
r_A	Bank A's actual capital ratio
μ_A	The expected value of f_A in the next period
σ_A	The volatility of Bank A's portfolio

(Changing subscripts of the above notations to B to obtain all the corresponding notations for Bank B.)

Δ_A	The amount of portfolio sold by Bank A
Δ_B	The amount of portfolio purchased by Bank B
$\underline{p_A}$	The trading price of Bank A portfolio

IV.1.1 Definitions and Assumptions

Suppose that at time t=0, two Banks have the same capital structure. Bank A's capital ratio r_A is defined as:

$$r_A = \frac{f_A * e_{AA} + c_A}{f_A * e_{AA} + c_A + OA_A}$$

(Same definition for r_B) In the next period t=1, the fundamental value of Bank A's risky portfolio f_A follows the normal distribution of $N\ [\ \mu_A,\ \sigma_A^2]$. The fundamental value of Bank B's portfolio f_B follows the normal distribution $N\ [\ \mu_B,\ \sigma_B^2]$, and for simplicity f_B is assumed to be uncorrelated with f_A. Suppose that at time t=0 both Banks' portfolios are at their fundamental values, i.e., $f_A = \mu_A$ and $f_B = \mu_B$. Since the fundamental values of portfolios are random variables, the capital ratios are also random variables at time t=1. Based on the randomness of r_A and r_B, the regulation imposes that at time t=1 the probabilities of $r_A \geq r_A^*$ and $r_B \geq r_B^*$ must be at least $q_A > 50\%$ and $q_b > 50\%$ respectively. Put differently, the $1 - q_A$ and $1 - q_B$ percent quantile of r_A and r_B must be at least r_A^* and r_B^* respectively.

IV.1.2 Bank A's Problem

At time t=0, Bank A experiences an idiosyncratic shock that clears out all of its cash c_A. In order to control the probability of violating regulations, Bank A sells some of its risky portfolio at time t=0 to reduce the volatility of its capital ratio at time t=1. After the shock and liquidating Δ_A at $\underline{p_A}$, Bank A's capital ratio changes to:

$$r_A = \frac{f_A(e_{AA} - \Delta_A) + \Delta_A \underline{p_A}}{f_A(e_{AA} - \Delta_A) + \Delta_A \underline{p_A} + OA_A} \quad (0 \leq \Delta_A \leq e_{AA}) \qquad (1)$$

The random component in r_A is $f_A(e_{AA} - \Delta_A)$, which follows the normal distribution $N\ [(e_{AA} - \Delta_A)\mu_A, (e_{AA} - \Delta_A)^2\sigma_A^2]$. In expression (1), parameters e_{AA}, OA_A, σ_A^2, and μ_A are constants. The effects of trading price $\underline{p_A}$ and the amount of portfolio sold Δ_A on distribution of r_A are the key determinants of Bank A's decision. Bank A's liquidation decision, which involves a choice of Δ_A at a given $\underline{p_A}$, affects r_A in two ways. On one hand, selling portfolio ($\Delta_A > 0$) reduces the volatility of r_A, which helps Bank A to control its risk. On the other hand, the liquidation decreases the expected value of r_A if the portfolio is sold at a loss ($\underline{p_A} < \mu_A$). To be compliant with the regulatory requirement may be costly for Bank A. Thus, it aims to minimize the loss incurred by liquidation. Bank A's problem is summarized as follows:

$$\underset{\underline{p_A},\Delta_A}{Min}(\mu_A - \underline{p_A})\ \Delta_A$$

$$s.t\ 0 \leq \Delta_A \leq e_{AA}$$

$$Pr\left[r_A\left(\Delta_A, \underline{p_A}\right) \geq r_A^*\right] \geq q_A$$

The first constraint controls the amount portfolio sold below the total amount. The second constraint is the regulatory requirement on r_A. The solution of this optimization problem is the liquidity demand function as a relationship between Δ_A and $\underline{p_A}$. Cast in mathematical form,

$$\Delta_A\left(\underline{p_A}\right) = \Delta_A^*, for\ all\ \underline{p_A} \epsilon [\widetilde{p_A}, \mu_A]$$

where Δ_A^* is given $\underline{p_A}$ the amount of portfolio to sell such that $Pr\left[r_A\left(\Delta_A^*, \underline{p_A}\right) \geq r_A^*\right] = q_A$. $\widetilde{p_A}$ is defined as $Pr[r_A(e_{AA}, \widetilde{p_A}) \geq r_A^*] = q_A$, and it is the lowest trading price that Bank A would accept to sell its entire portfolio. The second constraint always binds in an optimal solution because selling more than the necessary amount causes greater loss. The first constraint is also binding when Bank A sells its entire portfolio at $\widetilde{p_A}$. Economically, the liquidity demand curve represents the required liquidity support of a certain asset at all given prices.

IV.1.3 Bank B's Problem

Bank B does not experience any idiosyncratic shock. After buying Δ_B amount of portfolio at $\underline{p_A}$, Bank B's capital ratio changes from

$$r_B = \frac{f_B * e_{BB} + c_B}{f_B * e_{BB} + c_B + OA_B}$$

to

$$r_B = \frac{f_B e_{BB} + \Delta_B f_A + (c_B - \Delta_B\, \underline{p_A})}{f_B e_{BB} + \Delta_B f_A + (c_B - \Delta_B\, \underline{p_A}) + OA_B}.$$

The effects of $\underline{p_A}$ and Δ_B on the distribution of r_B are the key determinants of Bank B's decision. Bank B's capital ratio r_B changes in two ways. First, r_B becomes more volatile due to the purchase of risky assets. Second, Bank B makes profit if the trading price is below the current fundamental value. That is to say, Bank B records a trading profit of $\Delta_B\left(\mu_A - \underline{p_A}\right)$ at t=0 if $\underline{p_A} < \mu_A$, and thus the expected value of $r_B$at t=1 increases accordingly. This regulation constraint on r_B limits Bank B's ability to inject liquidity and make profit in the market. Bank B's profit maximizing problem is the following:

$$\underset{\underline{p_A}, \Delta_B}{Max} (\mu_A - \underline{p_A})\, \Delta_B$$

$$s.t\ \Delta_B \geq 0$$

$$Pr\left[r_B\left(\Delta_B, \underline{p_A}\right) \geq r_B^*\right] \geq q_B$$

This optimization problem will give the liquidity supply function as a relationship between Δ_B and $\underline{p_A}$. Formally,

$$\Delta_B\left(\underline{p_A}\right) = \begin{cases} [0, \widehat{\Delta_B}] & if\ \underline{p_A} = \mu_A \\ \Delta_B^* & if\ \underline{p_A} < \mu_A \end{cases}$$

where Δ_B^* is given $\underline{p_A}$ the amount of portfolio to buy such that $Pr\left[r_B\left(\Delta_B^*, \underline{p_A}\right) \geq r_B^*\right] = q_B$. By the same token, for $\widehat{\Delta_B}$, $Pr\left[r_B\left(\widehat{\Delta_B}, \mu_A\right) \geq r_B^*\right] = q_B$. The second constraint always binds in an optimal solution if $\underline{p_A} < \mu_A$ as Bank B seeks to maximize profit using all resources available. If $\underline{p_A} = \mu_A$, Bank B's trading profit is zero for any Δ_B, so it is indifferent among buying anything between zero and Δ_B^*. As for the economic meaning, the liquidity supply function represents the market capacity of a certain asset at all given prices. For example, $\widehat{\Delta_B}$ stands for the market capacity at fundamental price.

IV.1.4 Definition of Equilibrium

Define the general equilibrium in this financial market:

The general equilibrium is a set of $\{\underline{p_A^e},\ \Delta^e\}$, such that given the trading price $\underline{p_A^e}$, Bank A chooses $\Delta_A^*(\underline{p_A^e})$ to minimize losses, and Bank B chooses $\Delta_B^*(\underline{p_A^e})$ to maximize trading profit. The market clears with $\Delta^e = \Delta_A^* = \Delta_B^*$.

Definition 1: Liquid Equilibrium. If given $\underline{p_A} = \mu_A$, $\Delta_A^*(\mu_A) = \Delta_B^*(\mu_A)$. The Liquid Equilibrium is said to occur at the point where Bank A liquidates with no loss.

Definition 2: Illiquid Equilibrium. If at some $\underline{p_A^e} \in [\widetilde{p_A}, \mu_A)$, $\Delta_A^*(\underline{p_A^e}) = \Delta_B^*(\underline{p_A^e})$. The Illiquid Equilibrium is said to occur with insufficient liquidity, where Bank A liquidates at a loss of $\left(\mu_A - \underline{p_A^e}\right) * \Delta_A(\underline{p_A^e})$.

Definition 3: No Equilibrium. If at all given $\underline{p_A} \in [\widetilde{p_A}, \mu_A]$, $\Delta_A^*\left(\underline{p_A}\right) > \Delta_B^*(\underline{p_A})$. No equilibrium exists. Market is of zero liquidity since Bank A is not able to meet regulatory requirement via liquidating.

IV.2 Derivation of the Demand and Supply Functions

The analytical solutions of Bank A's liquidity demand curve and Bank B's liquidity supply curve are the primary results of this model. The demand function determines the "required liquidity support" at all given price levels.

First, as e_{AA} is the quantity of the portfolio, r_A's expression (1) can be normalized by setting $e_{AA} = 1$.

$$\frac{f_A(1-\Delta_A)+\Delta_A \underline{p_A}}{f_A(1-\Delta_A)+\Delta_A\, \underline{p_A}+OA_A} \geq r_A^* \quad (0 \leq \Delta_A \leq e_{AA})$$

Let the component $f_A(1-\Delta_A)+\Delta_A \underline{p_A}$ be a new normal random variable f_{ANew}, with mean $\mu_A+\Delta_A\left(\underline{p_A}-\mu_A\right)$ and variance $(1-\Delta_A)^2\sigma_A^2$. As discussed earlier, when the second constraint of Bank A' problem is binding, it implies

$$Pr\left(1-\frac{OA_A}{f_{ANew}+OA_A} \leq r_A^*\right) = 1-q_A,$$

which is equivalent to

$$Pr\left(f_{ANew} \leq \frac{r_A^* OA_A}{1-r_A^*}\right) = 1-q_A\,. \qquad (2)$$

(See the Appendix I.1 for more discussion about the inequality in equation (2).)

Since f_{ANew} follows the normal distribution$N\{\mu_A+\Delta_A\left(\underline{p_A}-\mu_A\right), (1-\Delta_A)^2\sigma_A^2\}$, equation (2) is equivalent to

$$P\,r\left(z \leq \frac{\frac{r_A^* OA_A}{1-r_A^*} - \left[\mu_A+\Delta_A\left(\underline{p_A}-\mu_A\right)\right]}{(1-\Delta_A)\sigma_A}\right) = 1-q_A \qquad (3)$$

Let $K_A = \Phi^{-1}(1-q_A)$, and thus equation (4) follows equation (3),

$$\frac{\frac{r_A^* OA_A}{1-r_A^*} - \left[\mu_A+\Delta_A\left(\underline{p_A}-\mu_A\right)\right]}{(1-\Delta_A)\sigma_A} = K_A. \qquad (4)$$

$\Phi^{-1}(\cdot)$ stands for the CDF of the standard normal distribution. Note that $K_A < 0$ because $q_A > 50\%$. Rearrange the above equation to obtain the inverse demand function

$$\underline{p_A} = \frac{\frac{r_A^* OA_A}{1 - r_A^*} - \mu_A - K_A \sigma_A}{\Delta_A} + \mu_A + K_A \sigma_A.$$

Further, the demand function

$$\Delta_A = \frac{\frac{r_A^* OA_A}{1 - r_A^*} - \mu_A - K_A \sigma_A}{\underline{p_A} - \mu_A - K_A \sigma_A}$$

reveals the "required liquidity support" at all given prices between $[\widetilde{p_A}, \mu_A]$,

Following the similar procedure, I derive the inverse liquidity supply function of Bank B,

$$\underline{p_A} = \mu_A + \frac{K_B \sqrt{\sigma_B^2 + \Delta_B^2 \sigma_A^2} - (\frac{r_B^* OA_B}{1 - r_B^*} - c_B - \mu_B)}{\Delta_B}$$

$where$ $K_B = \Phi^{-1}(1 - q_B)$. Note that $K_B < 0$ because $q_B > 50\%$. (See the derivation of the inverse liquidity supply function in Appendix I.2.) Next I will prove an important proposition of this model: unlike conventional supply curves, liquidity supply curves are downward sloping. Put mathematically,

Proposition 1:

$$\frac{\partial \underline{p_A}}{\partial \Delta_B} = \frac{T_2 \sqrt{\sigma_B^2 + \Delta_B^2 \sigma_A^2} - K_B \sigma_B^2}{\Delta_B^2 \sqrt{\sigma_B^2 + \Delta_B^2 \sigma_A^2}} < 0$$

for $\Delta_B \in [\widehat{\Delta_B}, 1]$, where $T_2 = \frac{r_B^* OA_B}{1 - r_B^*} - c_B - \mu_B$.

Proof: It is straightforward that we only need to show that $T_2 \sqrt{\sigma_B^2 + \Delta_B^2 \sigma_A^2} < K_B \sigma_B^2$. Since $K_B \sigma_B^2 < 0$, so that for $\Delta_B \in (0, 1]$, the upper bound of $T_2 \sqrt{\sigma_B^2 + \Delta_B^2 \sigma_A^2}$ is $T_2 \sigma_B$ when $\Delta_B \rightarrow 0$. Therefore, proving $\frac{\partial p_A}{\partial \Delta_B} < 0$ for all $\Delta_B \in (0, 1]$ is equivalent to proving $T_2 <$

$K_B\sigma_B$, i.e., $r_B^* < \frac{c_B+\mu_B+K_B\sigma_B}{c_B+\mu_B+OA_B+K_B\sigma_B}$. Before buying additional risk assets, the $1-q_B$ quantile of r_B is $\frac{c_B+\mu_B+K_B\sigma_B}{c_B+\mu_B+OA_B+K_B\sigma_B}$. (See Appendix I.3 for the derivation of this expression.) If the regulatory requirement r_B^* is below $\frac{c_B+\mu_B+K_B\sigma_B}{c_B+\mu_B+OA_B+K_B\sigma_B}$, then Bank B is originally compliant with the regulation, and thus Bank B is eligible to take in more risky assets. In the context of my model, the regulation threshold r_B^* should be below the current $1-q_B$ quantile of r_B. Otherwise, Bank B is supposed to be lowering its holdings of risky assets to comply with regulation as well. Therefore, the condition $r_B^* < \frac{c_B+\mu_B+K_B\sigma_B}{c_B+\mu_B+OA_B+K_B\sigma_B}$ is satisfied, which guarantees $\frac{\partial \underline{p_A}}{\partial \Delta_\mathrm{B}} < 0$ for all $\Delta_\mathrm{B} \in (0,1]$. Namely, the supply function is monotonically decreasing over $\Delta_\mathrm{B} \in (0,1]$.

It may be interesting to delve into the $r_B^* > \frac{c_B+\mu_B+K_B\sigma_B}{c_B+\mu_B+OA_B+K_B\sigma_B}$ case for further study. This is mainly because, if r_B^* is only slightly above $\frac{c_B+\mu_B+K_B\sigma_B}{c_B+\mu_B+OA_B+K_B\sigma_B}$, the supply curve only increases for very small Δ_B and then becomes downward sloping again. This paper focuses on the $r_B^* < \frac{c_B+\mu_B+K_B\sigma_B}{c_B+\mu_B+OA_B+K_B\sigma_B}$ case, which has more sensible financial meaning.

▪

The liquidity supply curve depicts the relationship between the amount of portfolio that Bank B can buy Δ_B and the trading price $\underline{p_A}$. Economically speaking, liquidity supply curves are downward sloping because the larger the spread between fundamental value and trading price, the more Bank B wants to purchase for profit maximization.

IV.3 Existence and Uniqueness of Equilibriums

To study the existence condition of liquid equilibrium, I calculate the $\widehat{\Delta_A}$ such that trading price equal to the fundamental value $\underline{p_A} = \mu_A$,

$$\widehat{\Delta_A} = \frac{\mu_A + K_A\sigma_A - \frac{r_A^* OA_A}{1-r_A^*}}{K_A\sigma_A}$$

$\widehat{\Delta_\mathrm{A}}$ is the required liquidity support for fundamental price. When the market capacity for Bank A's portfolio is more than $\widehat{\Delta_A}$, the market is in liquid equilibrium. Then I calculate $\widehat{\Delta_\mathrm{B}}$ such that $\underline{p_A} = \mu_A$,

$$\widehat{\Delta_B} = \sqrt{\frac{\left(\frac{\frac{r_B^* OA_B}{1-r_B^*} - c_B - \mu_B}{K_B}\right)^2 - \sigma_A^2}{\sigma_B^2}}$$

$\widehat{\Delta_B}$ is the market capacity for Bank A's portfolio at fundamental price, which means the market is only able to absorb $\widehat{\Delta_B}$ amount of Bank A's portfolio at fundamental price. For now, simply comparing $\widehat{\Delta_B}$ and $\widehat{\Delta_A}$ reveals the existence of **Liquid Equilibrium**.

***Proposition 2*:** If $\widehat{\Delta_A} \leq \widehat{\Delta_B}$, the market has **Liquid Equilibrium**. Otherwise, the market either has **Illiquid Equilibrium** (with liquidity shortage) or **No Equilibrium** (with zero liquidity).

Put differently, in the latter two cases, Bank B either requires a significant liquidity risk premium as compensation for taking over illiquid assets or doesn't take over illiquid assets at all. Also note that increasing Bank B's target ratio reduces market capacity $\widehat{\Delta_B}$ because

$$\frac{\partial \widehat{\Delta_B}}{\partial r_B^*} = \frac{T_2\left(\frac{OA_B}{1-r_B^*} + \frac{r_B^* OA_B}{(1-r_B^*)^2}\right)}{\sigma_B K_B^2 \sqrt{\frac{(T_2)^2}{K_B^2} - \sigma_A^2}} < 0$$

where $T_2 = \frac{r_B^* OA_B}{1-r_B^*} - c_B - \mu_B < 0$. However, the required liquidity support $\widehat{\Delta_A}$ is invariant to changes in r_B^*, namely Bank A's decision is independent of the actual liquidity condition in the market.

Next, in the case of illiquid equilibriums where $\widehat{\Delta_B} < \widehat{\Delta_A}$, I show the condition for the uniqueness of equilibrium. For an **Illiquid Equilibrium**, the equation $\Delta_A = \Delta_B = \Delta^e$ at $\underline{p_A^e} \in [\widetilde{p_A}, \mu_A)$ leads to,

$$\frac{\frac{r_A^* OA_A}{1-r_A^*} - \mu_A - K_A\sigma_A}{\Delta^e} + \mu_A + K_A\sigma_A = \mu_A + \frac{K_B\sqrt{\sigma_B^2 + (\Delta^e)^2\sigma_A^2} - \left(\frac{r_B^* OA_B}{1-r_B^*} - c_B - \mu_B\right)}{\Delta^e}$$

Solving the above equation will give the solution of Δ^e and $\underline{p_A}^e$. When $K_A = K_B$, the above equation has unique solution if it exists in $\Delta^e \in \left(\widehat{\Delta_B}, 1\right]$, $\underline{p_A^e}(\Delta^e) \in [\widetilde{p_A}, \mu_A)$.

Proposition 3: If $K_A = K_B = K$, the Illiquid Equilibrium $(\underline{p_A^e}(\Delta^e), \Delta^e)$, if it exists, must be unique.

$$\Delta^e = \frac{K^2\sigma_B^2 - (T_1 + T_2)^2}{2(T_1 + T_2)K\sigma_A}$$

where $\Delta^e \in \left(\widehat{\Delta_B}, 1\right]$, $\underline{p_A^e}(\Delta^e) \in [\widetilde{p_A}, \mu_A)$.

$$T_1 = \frac{r_A^* OA_A}{1 - r_A^*} - \mu_A - K\sigma_A \text{ and } T_2 = \frac{r_B^* OA_B}{1 - r_B^*} - c_B - \mu_B.$$

Multiple illiquid equilibria may exist if $K_A \neq K_B$, which is equivalent to $q_A \neq q_B$. Since Bank A is selling assets at fire-sale prices to meet regulatory requirements, the equilibrium with lowest Δ^e and highest $\underline{p_A^e}(\Delta^e)$ at the mean time is the most favorable equilibrium for Bank A. Note that for the lowest Δ^e, $\underline{p_A^e}(\Delta^e)$ is assured to be the highest equilibrium trading prices among potential multiple equilibriums because liquidity demand function is monotonically decreasing with respect to Δ^e.

V. Numerical Examples and Comparative Statics Analysis

V.1 Numerical Examples of Liquidity Demand and Supply Curve

Before providing the simulation results of r_A, r_B, liquidity supply curve and liquidity demand curve to further illustrate the model, it is helpful to see how the distributions of r_A and r_B are affected by trading the portfolio.

For r_A, I can first simulate its distribution using expression (1) at any given $\underline{p_A}$ *and* Δ_A. Parameters used for the simulations in Figure 1 are the following: $f_A = \mu_A = f_B = \mu_B = 1,\ OA_A = OB_B = 4, c_B = 1,\ \sigma_A = \sigma_B = 0.2,\ e_{AA} = e_{BB} = 1, q_A = q_B = 99\%$.

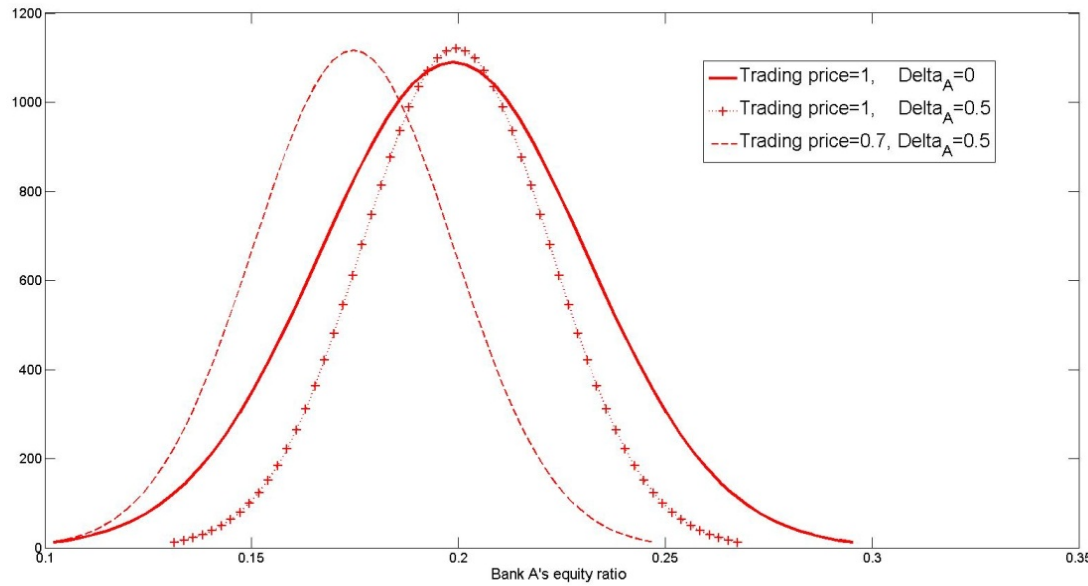


Figure 1: The distributions of Bank A's capital ratios under three conditions.

As shown in Figure 1, if Bank A liquidates half its portfolio at μ_A, then r_A's distribution will become more concentrated. (Figure 1: $\underline{p_A} = 1$ and $\Delta_A = 0.5$) If Bank A must liquidate at a loss, the distribution will still be more concentrated but shifted to the left in parallel. (Figure 1: $\underline{p_A} = 0.7$ and $\Delta_A = 0.5$)

Based on the distribution of r_A, it is easy to obtain the liquidity demand curve of Bank A at a given r_A^*. For any given $\underline{p_A}$, increase Δ_A from zero until the q_A quantile of the distribution is just equal to r_A^*. Repeating such iteration at all $\underline{p_A} \epsilon [\widetilde{p_A}, \mu_A]$ will generate the liquidity demand curve in the situation where the regulatory requirement is (r_A^*, q_A) Here are several examples of simulated liquidity demand curves based on different r_A^*s fitted with analytical solutions. (Fig. 2)

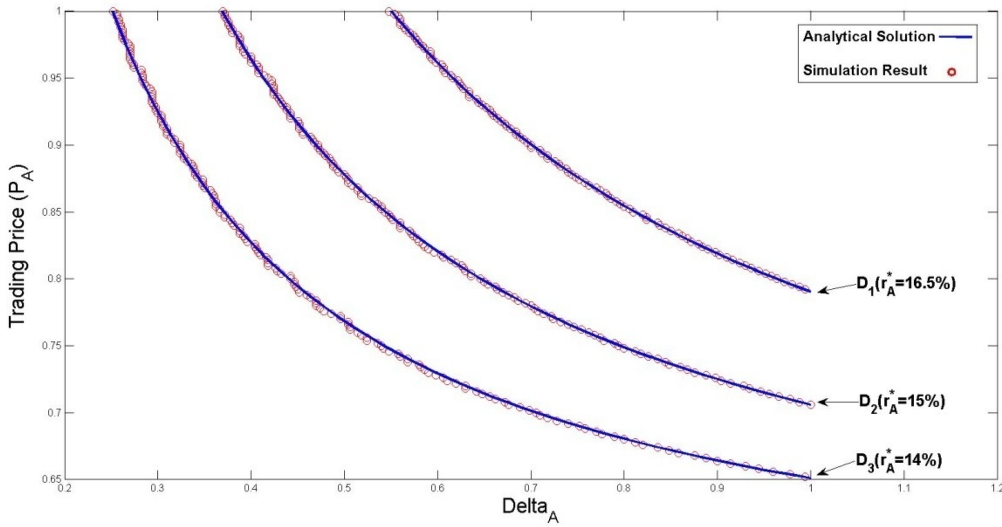


Figure 2: Bank A's liquidity demand curves with different target ratios.

It is clear that increasing r_A^* shifts demand curve rightward. Then I show the relationship between the demand curve and r_A^* more formally.

Proposition 4: $\frac{d\Delta_A}{dr_A^*} > 0$, i.e., liquidity demand curves shift upward if r_A^* increases.

Proof: From Section IV.2, the demand function is

$$\Delta_A = \frac{\frac{r_A^* OA_A}{1 - r_A^*} - \mu_A - K_A\sigma_A}{\underline{p_A} - \mu_A - K_A\sigma_A},$$

and thus

$$\frac{d\Delta_A}{dr_A^*} = \frac{\frac{OA_A}{1 - r_A^*} + \frac{r_A^* OA_A}{(1 - r_A^*)^2}}{\underline{p_A} - \mu_A - K_A\sigma_A}.$$

In order to have $\frac{d\Delta_A}{dr_A^*} > 0$, it must be true that $\underline{p_A} - \mu_A - K_A\sigma_A > 0$ for all $\underline{p_A} \in [\widetilde{p_A}, \mu_A)$, which in turn implies that $\widetilde{p_A} - \mu_A - K_A\sigma_A > 0$ must be true. With the previous result $\widetilde{p_A} = \frac{r_A^* OA_A}{1-r_A^*}$, showing $\widetilde{p_A} - \mu_A - K_A\sigma_A > 0$ is equivalent to showing $r_A^* > \frac{\mu_A+K_A\sigma_A}{\mu_A+K_A\sigma_A+OA_A}$.

Now I employ the similar method used in the proof of **Proposition 1**. As $\frac{\mu_A+K_A\sigma_A}{\mu_A+K_A\sigma_A+OA_A}$ stands for the $1-q_A$ quantile of r_A before Bank A selling assets, $r_A^* > \frac{\mu_A+K_A\sigma_A}{\mu_A+K_A\sigma_A+OA_A}$ must be true, otherwise Bank A does not have to liquidate its portfolio in the first place. Therefore, $\frac{d\Delta_A}{dr_A^*} > 0$ is true.

▪

When it comes to Bank B, if it buys a half of Bank A's portfolio at μ_A, then the distribution of r_B will be more dispersed. (Figure 3: $\underline{p_A} = 1$ and $\Delta_A = 0.5$) If $\underline{p_A} = 0.7$, Bank B makes a profit of 0.3 from each unit of portfolio bought from Bank A. In this case, the distribution will still be more dispersed but shifted to the right in parallel. (Figure 3: $\underline{p_A} = 0.7$ and $\Delta_A = 0.5$) Based on a similar simulation algorithm, several liquidity supply curves are obtained and fitted with analytical solutions in Fig. 4. For $\underline{p_A} = \mu_A$, Bank B is indifferent among $[0, \widehat{\Delta_B}]$ as its expected profit is zero for all $\Delta_B \in [0, \widehat{\Delta_B}]$. This situation is captured by the horizontal line at $\underline{p_A} = \mu_A = 1$ in Fig. 4.

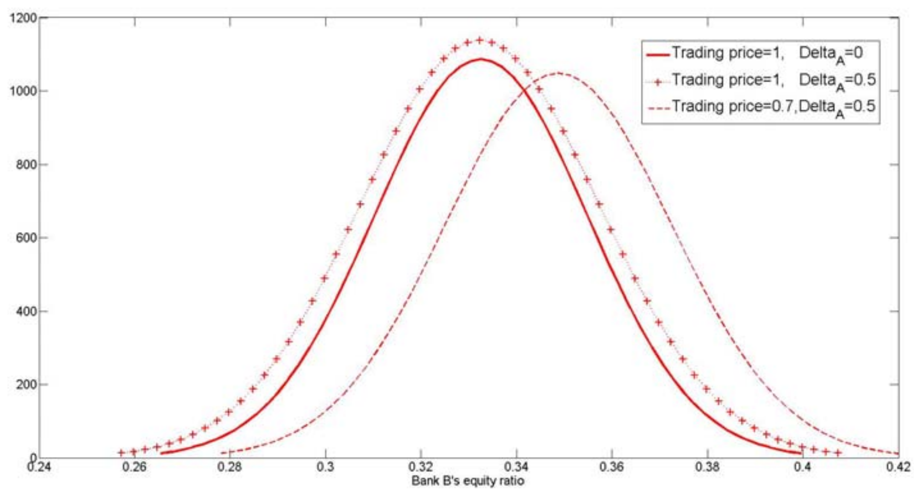


Figure 3: The distributions of Bank B's capital ratios under three conditions.

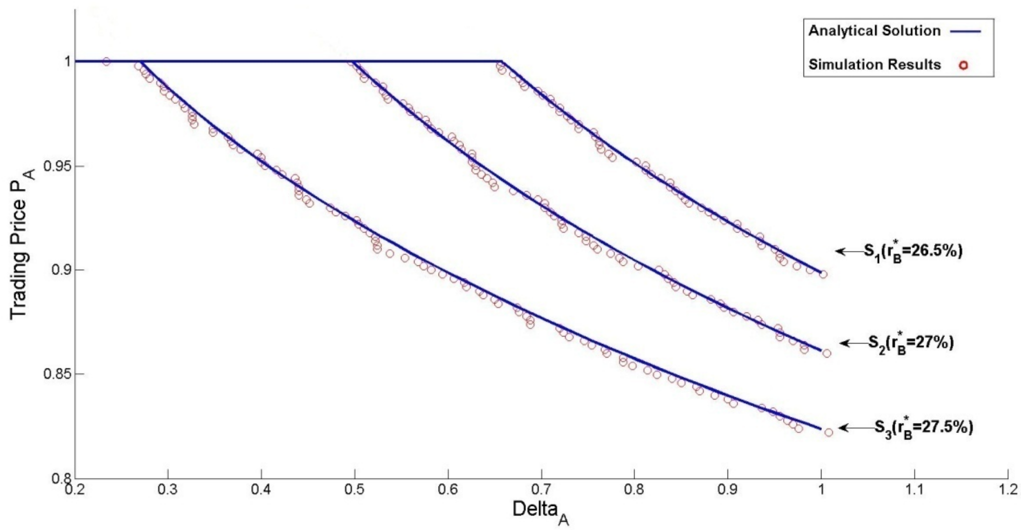


Figure 4: Bank B's liquidity supply curve with different target capital ratios.

Finally, I prove the proposition regarding the relationship between liquidity supply curves and r_B^*.

***Proposition 5*:** $\frac{d\Delta_B}{dr_B^*} < 0$, i.e., liquidity supply curves shift downward as r_B^* increases.

Proof: Define a new implicit function of Δ_B and r_B^*,

$$F(\Delta_\mathrm{B}, r_B^*) = \mu_A + \frac{K_B\sqrt{\sigma_B^2 + \Delta_B^2\sigma_A^2} - (\frac{r_B^* OA_B}{1 - r_B^*} - c_B - \mu_B)}{\Delta_\mathrm{B}} - \underline{p_A} = 0$$

Thus,

$$\frac{d\Delta_B}{dr_B^*} = -\frac{\frac{\partial F(\Delta_\mathrm{B}, r_B^*)}{\partial r_B^*}}{\frac{\partial F(\Delta_\mathrm{B}, r_B^*)}{\partial \Delta_\mathrm{B}}} = -\frac{\frac{-(\frac{OA_B}{1 - r_B^*} + \frac{r_B^* OA_B}{(1 - r_B^*)^2})}{\Delta_\mathrm{B}}}{\frac{T_2\sqrt{\sigma_B^2 + \Delta_B^2\sigma_A^2} - K_B\sigma_B^2}{\Delta_B^2\sqrt{\sigma_B^2 + \Delta_B^2\sigma_A^2}}} < 0.$$

(See the proof of $\frac{T_2\sqrt{\sigma_B^2+\Delta_B^2\sigma_A^2} - K_B\sigma_B^2}{\Delta_B^2\sqrt{\sigma_B^2+\Delta_B^2\sigma_A^2}} < 0$ in the proof of **Proposition 1**)

V.2 Numerical Examples of Different Equilibriums

After showing the property of liquidity supply and demand curve, the next two sections focus on different types of equilibria along with a comparative statics analysis regarding the riskiness of Bank A's portfolio. It is found that highly risky assets face more severe liquidity shortage in a stress market, and thus they have large swings in prices.

The demand curve $\mathbf{D_2}$ and the supply curve $\mathbf{S_2}$ in Fig. 5 generate a **Liquid Equilibrium** $\mathbf{E_1}$. When $\underline{p_A} = \mu_A = 1$, the maximum amount that Bank B is willing to take over $\widehat{\Delta_B}$ is more than what Bank A needs to liquidate, i.e. $\widehat{\Delta_A} \epsilon [0, \widehat{\Delta_B}]$. Hence, Bank A is able to liquidate around 40% of its portfolio at the current fundamental price in a liquid market.

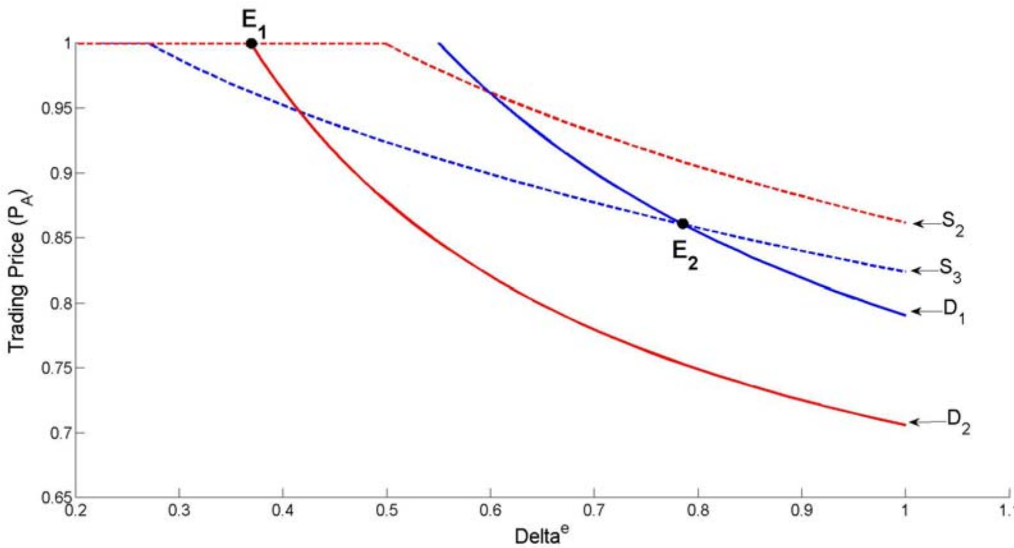


Figure 5: Use supply and demand curves to determine the equilibrium. Note that the demand and supply curves in Fig. 5 are taken directly from Fig. 2 and Fig. 4. Dashed lines are supply curves, and solid lines are demand curves.

Fixing either the demand or supply curve determines a "tipping point" that differentiates liquid market and illiquid market. For example, given a fixed $\mathbf{S}(r_B^*)$, the equilibrium price is invariant to any marginal shift in the liquidity demand curves as long as $\widehat{\Delta_A} \leq \widehat{\Delta_B}$. This property is consistent with the generalized characteristic of liquidity risk that I mentioned in the Introduction: liquidity risk does not manifest itself in normal times, so that assets are traded at fundamental prices in equilibrium. Beyond this critical point ($\widehat{\Delta_A} = \widehat{\Delta_B}$), the equilibrium price will become very sensitive to marginal changes in liquidity demand as the market may fall into an **Illiquid Equilibrium** $\widehat{\Delta_A} < \widehat{\Delta_B}$. Consider the Illiquid Equilibrium $\mathbf{E_2}$ generated by supply curve $\mathbf{S_1}$ and demand curve $\mathbf{D_3}$ in Fig. 5. The equilibrium price drops by roughly fifteen percent with only small changes in r_A^* and r_B^*.

The high sensitivity is mainly due to the similar supply curve slope and demand curves slope. The high price sensitivity confirms the other characteristic of liquidity risk mentioned in the Introduction: without any change in the fundamental value of an asset, liquidity risk itself makes asset prices greatly deviate from the fundamental value, and asset prices are very sensitive to marginal changes in market liquidity condition.

V.3 Comparative Statics Analysis

To better illustrate the characteristics of liquidity risk implied by this model, I will conduct a comparative statics analysis between two portfolios with different volatilities. For the required liquidity support, the first sensitivity analysis regards the standard deviation of f_A. The first order partial derivative of Δ_A with respect to σ_A is,

$$\frac{\partial \Delta_A}{\partial \sigma_A} = \frac{\left(\frac{r_A^* OA_A}{1 - r_A^*} - \underline{p_A}\right) K_A}{\left(\underline{p_A} - \mu_A - K_A \sigma_A\right)^2} > 0$$

for $\underline{p_A} \in (\widetilde{p_A}, \mu_A]$. The above inequality holds for all $\underline{p_A} \in (\widetilde{p_A}, \mu_A]$, because $\frac{r_A^* OA_A}{1-r_A^*} = \widetilde{p_A}$. $\frac{\partial \Delta_A}{\partial \sigma_A} > 0$ means that the required liquidity support increases with $\sigma_A$at all given prices. On the other hand, the market capacity reduces with σ_A at all given prices because $\frac{\partial \Delta_B}{\partial \sigma_A} < 0$ (See Appendix II.1 for the proof). Based on the above comparative analysis, I pose the following proposition,

***Proposition 6*:** For $\underline{p_A} \in (\widetilde{p_A}, \mu_A]$, $\Delta_A \in [0,1]$, and $\Delta_B \in [0,1]$, higher riskiness of Bank A's portfolio results in leftward shift in liquidity supply curves and rightward shift in liquidity demand curves, i.e. $\frac{\partial \Delta_B}{\partial \sigma_A} < 0$ and $\frac{\partial \Delta_A}{\partial \sigma_A} > 0$.

With the opposite effects of σ_A on Δ_A and Δ_B, it is expected that the market situation will deteriorate if σ_A increases. The following numerical example will demonstrate how the originally liquid equilibrium would become an illiquid equilibrium with the increase in the asset riskiness. I compare the market liquidity condition of highly risky portfolio $\boldsymbol{\sigma'_A} = 25\%$ (Fig. 6) with previous examples of portfolio $\boldsymbol{\sigma_A} = 20\%$ in Figure 5.

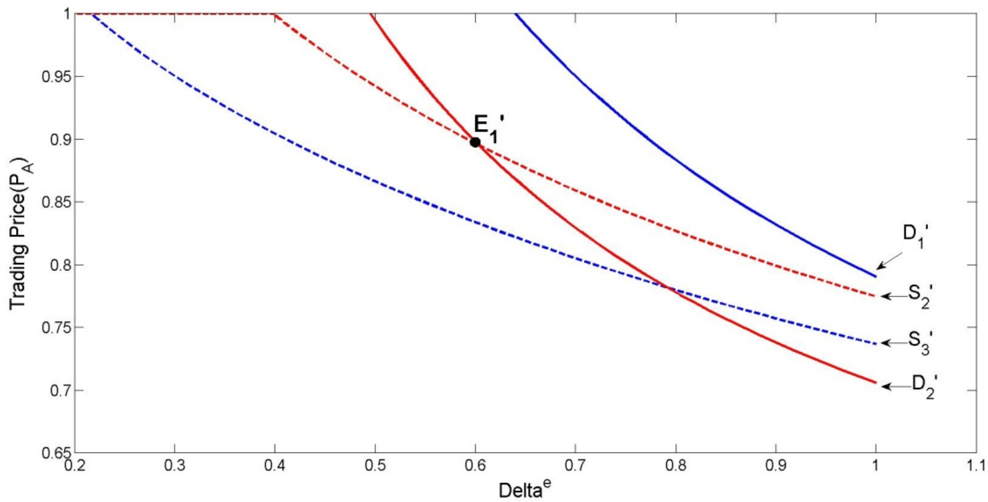


Figure 6: Demand and supply curves with portfolio's volatility equal to 25%.

As shown both analytically and numerically, the increase in riskiness shifts liquidity demand curves rightward, and it shifts liquidity supply curve leftward. Compared with the previous Illiquid Equilibrium $\mathbf{E_2}$ generated by $\mathbf{S_3}$ and $\mathbf{D_1}$ (Fig. 5), two Banks are not able to reach an equilibrium given $\sigma_A' = 25\%$ (See $\mathbf{S_3'}$ and $\mathbf{D_1'}$ in Fig. 6). The market provides zero liquidity for Bank A's highly risky portfolio because the liquidity required at any $\underline{p_A} \in (\widetilde{p_A}, \mu_A]$ is higher than market capacity. With $\mathbf{S_2}$ and $\mathbf{D_2}$ (Fig. 5) shifting to $\mathbf{S_2'}$ and $\mathbf{D_2'}$ (Fig. 6), the originally liquid equilibrium ($\mathbf{E_1}$ in Fig. 5) becomes an illiquid equilibrium ($\mathbf{E_1}'$in Fig. 6). Bank A now has to sell nearly 60% of its portfolio at 10% loss.

The intuitive explanation is as follows. First, investors tend to be reluctant to purchase highly volatile assets in a stressed market, and thus the liquidity supply for such assets is low. Second, in terms of reducing the uncertainty of capital ratio, selling highly volatile assets is more efficient relative to selling less volatile assets. Banks tend to sell risky asset first, which in turn creates higher liquidity demand. The increase in liquidity demand superimposes the decrease of liquidity supply to generate severe downward pressure on the asset price.

VI. Applications on Liquidity Related Issues

Based on the recent crisis in 2007 and 2008, Brunnermeier emphasizes three mechanisms—precautionary hoarding by individual banks, runs on financial institutions and loss spiral, through which mortgage delinquency shock is amplified to a full-blown liquidity crisis (Brunnermeier 2009). The three liquidity-related issues are closely related to my current model.

VI.1 Precautionary Hoarding

The model has an interesting prediction for the impact of precautionary hoarding on asset prices. In the context of my model, Bank B is the financial institution that can conduct precautionary hoarding by setting an overly conservative target ratio. If Bank B is afraid of potential liquidity shocks against itself, it may set its target capital ratio at a very conservative level. Thus, the market capacity reduces as $\frac{\partial \Delta_B}{\partial r_B^*} < 0$ (**Proposition 1**), which may lead Bank A to liquidate at greater losses. Besides dealing with uncertain liquidity demand, Bank B prefers hoarding to supplying sufficient liquidity for another reason. If Bank B realizes its monopolistic power in the market, it is able to make more profit by setting a high r_B^*. To reflect Bank B's decision on the target ratio, I modify Bank B's problem by allowing Bank B to freely set its target ratio above the regulatory requirement. The modified profit maximization problem is shown below.

$$\underset{r_B^*}{Max}\ Z = (\mu_A - \underline{p_A})\ \Delta^e$$

$$s.t.\ \ 0 < \Delta^e \leq 1$$

$$\Delta^e = \frac{K^2\sigma_B^2 - (T_1 + T_2)^2}{2(T_1 + T_2)K\sigma_A}$$

$$\underline{p_A} = \frac{\frac{r_A^* OA_A}{1 - r_A^*} - \ \mu_A - K\sigma_A}{\Delta^e} + \mu_A + K\sigma_A$$

$$r_B^* \geq R_{regulation}$$

where $T_1 = \frac{r_A^* OA_A}{1 - r_A^*} - \ \mu_A - K\sigma_A \ and\ T_2 = \frac{r_B^* OA_B}{1 - r_B^*} - c_B - \mu_B$

Based on the previous analytical solutions of the general equilibrium, Bank B can choose an optimal r_B^* to achieve an optimal equilibrium with maximum profit. For simplicity, I only study the case where unique illiquid equilibrium occurs under the condition $K_A = K_B = K$. The objective function Z and $\frac{dZ}{dr_B^*}$ have explicit expressions in terms of r_B^*:

$$\frac{dZ}{dr_B^*} = \frac{d(\mu_A - \underline{p_A})\ \Delta^e}{dr_B^*}$$

$$= \frac{d}{dr_B^*}(-T_1 - K\sigma_A\Delta^e)$$

$$= \frac{1}{2}[K^2\sigma_B^2 + (T_1 + T_2)^2]\left[\frac{OA_B}{1 - r_B^*} + \frac{r_B^* OA_B}{(1 - r_B^*)^2}\right] > 0\ for\ all\ r_B^* \in [0,1]$$

Therefore, Bank B will set its target ratio as high possible. It will not choose any r_B^* that makes $\underline{p_A}$ lower than $\widetilde{p_A}$, because such r_B^* leads to a zero-liquidity market in which Bank A will not trade at all. (Recall that $\widetilde{p_A} = \frac{r_A^* OA_A}{1-r_A^*}$, and it is the lowest price that Bank A would accept to sell the entire portfolio.) The relationship between Bank B's profit and its target ratio is summarized in the following proposition.

***Proposition 7*:** For $\Delta^e(r_B^*) \in [0,1]$ and $\underline{p_A}(\Delta^e) \in [\widetilde{p_A}, \mu_A]$, Bank B's profit Z increases with r_B^* because $\frac{dZ}{dr_B^*} > 0$ for all $r_B^* \in [0,1]$.

I provide a numerical example (Fig. 7) to demonstrate the effect of r_B^* on $\underline{p_A}$, Δ^e, Z and the expected return $\frac{Z}{\underline{p_A} * \Delta^e}$. I select the demand curve $\mathbf{D_1}$ (where $r_A^* = 16.5\%$ in Fig. 2), and calculate the equilibrium prices as r_B^* changes from 26% to 28% to obtain Figure 7. When $r_B^* \leq 26.85\%$, there is sufficient liquidity in the market to support trades at the fundamental value. Illiquid Equilibria occur for $26.85\% < r_B^* \leq 27.93\%$, the equilibrium trading price will fall by about twenty percent if r_B^* drops by merely one percent. For $r_B^* > 27.93\%$, no equilibrium occurs and there is zero liquidity in the market. Overall, this high price sensitivity is consistent with the observation—when market participants become concerned about liquidity issues, asset prices fall drastically. The upward sloping Z in illiquid market confirms the result $\frac{dZ}{dr_B^*} > 0$. The upward sloping expected return curve suggests that if Bank B's object is to maximize expected return instead of expected profit, the optimal r_B^* remains the same.

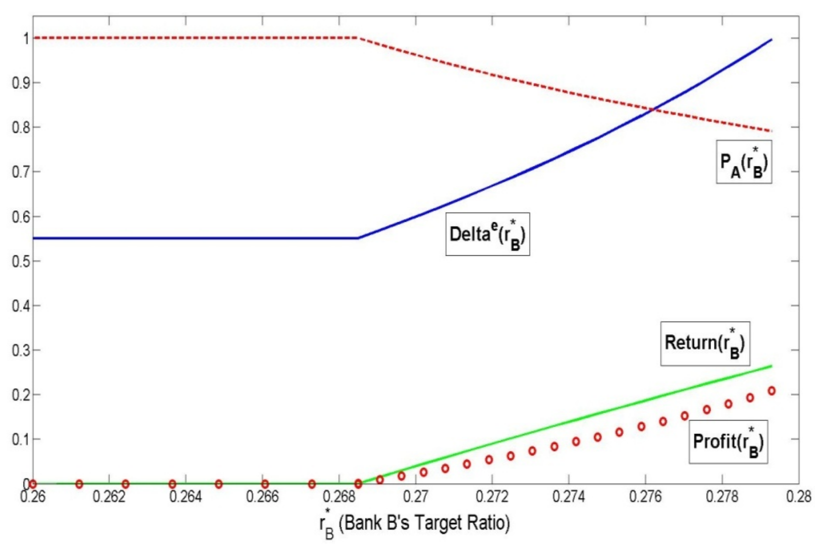


Figure 7: Given a liquidity demand curve, the equilibrium price changes with Bank B's target ratio.

To sum up, my model suggests that when facing a desperate seller (Bank A), hoarding is an effective way for the counterparty financial institution to earn profit and maintain solid financial condition at the same time. The market situation, however, deteriorates in the sense that asset prices have larger swings and the desperate seller suffers a huge loss.

VI.2 Bank Run Loss

The precautionary case does not incorporate any externality of deteriorating market situation. A noticeable adverse consequence is runs on financial institutions. If investors are panicked by the steep decline in trading price, they may trigger another amplifying mechanism—runs on financial institutions. Depositors want to withdraw funds as early as possible if they are afraid that Bank A will become insolvent in the future. There are two scenarios for depositors' behavior:

Scenario 1: Depositors are well informed, namely they know this trade with discounted price is only due to an idiosyncratic shock against Bank A. Since the quality of Bank A's portfolio is fundamentally unchanged, it is not necessary to withdraw deposits immediately. Bank A manages to save itself via liquidation.

Scenario 2: Depositors are not well informed, namely they mistakenly attribute the steeply declining asset price to a downgrade of Bank A's portfolio or a decrease in its fundamental value. Fearful investors may start to dump those holdings that are most correlated with Bank A's portfolio, which aggravates Bank A's financial condition. Or a bank run may drain the cash that Bank A has just acquired. The consequence will spread to Bank B as bank runs are often indiscriminate. Thus, it is assumed that once depositors decide to run, they run on all banks regardless of the financial condition of individual institutions. Another critical property of a bank run assumed in this scenario is that the greater the declines in asset prices, the more likely investors are to run. With an assumed constant bank run loss, the expected bank run loss increases accordingly.

Under **Scenario 2**, although the spread between the trading price and the fundamental value is still a profit opportunity for Bank B, it also causes a higher probability of a bank run. After calculating the expected trading profit and the expected bank-run loss, Bank B may be willing to supply more liquidity to purchasing assets rationally at higher prices. Current literature about bank runs does not provide an explicit quantitative relationship for the probability of bank run and the asset price. The study of such relationship is beyond the scope of this paper. To illustrate, I assume a logistic probability function for the probability of bank run and a lump-sum loss if a bank run occurs. Suppose that bank-run loss is \$0.2, and the probability of a bank run follows

$$Pr(Bankrun) = \frac{1}{1 + exp\left[-\frac{Fall\ in\ asset\ price - 0.17}{0.02}\right]}$$

Then Bank B’s problem differs from the last one only in terms of the objective function, and the new one is:

$$\underset{r_B^*}{Max}\ Z = (\mu_A - \underline{p_A})\ \Delta^e - \frac{0.2}{1 + \exp\left[-\frac{(\mu_A - \underline{p_A}) - 0.17}{0.02}\right]}$$

Compared with the no-bank-run case, the $Z(r_B^*)$ is not monotonically increasing anymore as the expected bank-run loss eventually would exceed the trading profit. Based on the results in Fig. 8, it is obvious that focusing only on trading profit is a myopic behavior. Bank B should instead a set lower target ratio to maximize the expected net profit. Bank B’s lower target ratio would ease the depressed trading price and liquidity shortage.

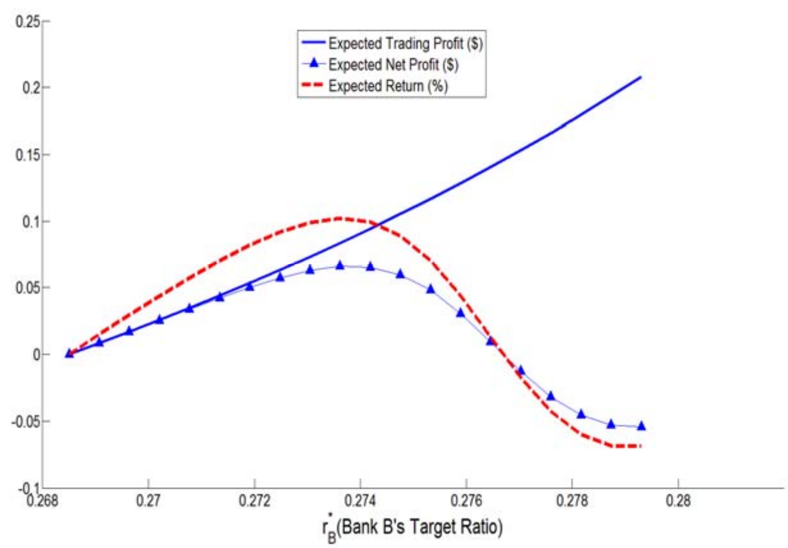


Figure 8: Bank B’s expected trading profit, expected net profit, and expected return with r_B^*.

To conclude, my model shows that financial institutions are willing to hoard less liquidity and rationally purchase assets at higher prices in the case where a bank run is incorporated. That is to say, if market participants are provided with clear information about the externalities of declining assets prices, the market liquidity position can be moderately eased to generate higher trading prices in equilibrium.

VI. 3 Loss Spiral

Following the Bank A's fire sale, a loss spiral is another mechanism that can lead to a deteriorating market situation. Marking-to-market is the mechanism that triggers loss spiral. Suppose that after liquidating $\Delta^e \in (0,1)$ in an illiquid equilibrium with a fire-sale price, Bank A has to mark its remaining portfolio to the equilibrium trading price $\underline{p_A}^e$, then the write-down loss would immediately bring down Bank A's capital ratio. Again, if Bank A has to be compliant with the regulation, it has to sell more assets at even lower prices. It is expected that with reiterated costly liquidations and write-down losses, Bank A's problem would quickly evolve from illiquidity to insolvency. The critical concern about marking-to-market is that it relies on the perception that the market price should reflect the "fair value" of assets. When a market lacks liquidity, however, sporadic trades of an asset do not necessarily reflect its fundamental price. As my model has shown, the trading price may include a large "liquidity risk premium" and it deviates greatly from fundamental value. In general, banks should be relieved from marking-to-market, at least during times of stress, to prevent liquidity problems from being transmitted to solvency problems.

VII. Conclusion

This paper studies the effect of liquidity shortage on asset prices using a general equilibrium approach. Two critical results of this liquidity risk model are as follows. First, this model generates explicit downward sloping liquidity demand and supply curves. Second, this model inherently produces a tipping point that differentiates liquid market from illiquid market. In a liquid market, the equilibrium trading price equals the fundamental value of an asset. In an illiquid market, however, the equilibrium trading price deviates from the fundamental value and becomes very sensitive to marginal changes in market liquidity position. This property is consistent with the observation that liquidity risk only manifest in turbulent markets to depress asset prices.

My model also sheds light on three liquidity related issues, including precautionary hoarding by individual banks, runs on financial institutions and the loss spiral problem. The important conclusions are: although hoarding is an effective way for financial institutions to earn profit, the externality of hoarding should be taken into consideration to achieve a better overall market condition; moreover, financial institutions should be relieved from marking-to-market, at least in the time of stress, to prevent liquidity problems from being transmitted to solvency problems.

For further studies, the model can be expanded to study financial contagion by adding more financial institutions to the market and using the common illiquid risky asset as the channel of contagion. As several other forms of liquidity supply and demand functions are simply assumed in the literature of financial contagion, it may be interesting to construct a financial contagion model based on the liquidity supply and demand function derived in this paper, and then compare the results with existing literature.

References

Allen, F., E. Carletti, and D. Gale. 2009. "Interbank Market Liquidity and Central Bank Intervention." *Journal of Monetary Economics* 56: 639-652.

Allen, Franklin, and Douglas Gale. 2000. "Financial Contagion." *Journal of Political Economy* 108 (1): 1-33.

Brunnermeier, Markus K. 2009. "Deciphering the Liquidity and Credit Crunch 2007-2008." *Journal of Economic Perspectives* 23 (1): 77-100.

Brunnermeier, Markus K., and Lasse Heje Pederson. 2009. "Market Liquidity and Funding Liquidity." *The Review of Financial Studies* 22 (6): 2201–2238.

Chen, L., D. A. Lesmond, and J. Wei. 2007. "Corporate Yield Spreads and Bond Liquidity." *The Journal of Finance* 62 (1): 119-149.

Cifuentes, R., G. Ferrucci, and H. S. 2005. "Liquidity Risk and Contagion." *Journal of the European Economics Association* 3 (2-3): 556-566.

De Jong, Frank, and Joost Driessen. 2006. "Liquidity Risk Premia in Corporate Bond Markets." Working Paper. Available at SSRN: http://ssrn.com/abstract=686681 or http://dx.doi.org/10.2139/ssrn.686681

Eisenschmidt J., and J. Tapking J. 2009. "Liquidity Risk Premia in Unsecured Interbank Money Markets." ECB Working Paper No.1025, European Central Bank, Frankfurt.

Garcia, Alejandro, and Andrei Prokopiw. 2009. "Measure of Aggregate Credit Conditions and Their Potential Use by Central Banks." Bank of Canada Discussion Paper 2009-12, Bank of Canada, Ottawa.

Gorton, Gary, and Lixin Huang. 2006. "Bank Panics and the Endogeneity of Central Banking." *Journal of Monetary Economics* 53: 1613-1629.

Krishnamurthy, Arvind. 2010. "How Debt Market Have Malfunctioned in the Crisis." *Journal of Economic Perspectives*. 24 (1): 3-28.

Longstaff, F., S. Mithal, and E. Neis. 2005. "Corporate Yield Spreads: Default Risk or Liquidity? New Evidence from the Credit Default Swap Market." *Journal of Finance* 60 (5): 2213-2253.

Tirole, Jean. 2011. "Illiquidity and All Its Friends." *Journal of Economic Literature* 49 (2): 287-325.

Appendix I.1

In order to have $Pr\left(1 - \frac{OA_A}{f_{ANew}+OA_A} \leq r_A^*\right) = 1 - q_A$ equivalent to $Pr\left(f_{ANew} \leq \frac{r_A^* OA_A}{1-r_A^*}\right) = 1 - q_A$, $f_{ANew} + OA_A$ must be positive, which is not guaranteed mathematically because f_{ANew} follows a normal distribution. I will discuss the situations that $f_{ANew} + OA_A \leq 0$ to show that $f_{ANew} + OA_A > 0$ is a reasonable assumption.

Since $f_{ANew} + OA_A$ financially stands for the total asset of Bank A, if $f_{ANew} + OA_A < 0$, Bank A should file bankruptcy because it loses all of its assets. Moreover, $f_{ANew} + OA_A = 0$ is a zero probability event, so it is impossible to occur. Lastly, if $f_{ANew} + OA_A \to 0$ from the right side, r_A approaches negative infinite, which is included in the case $1 - \frac{OA_A}{f_{ANew}+OA_A} \leq r_A^*$. Therefore, studying the model under the assumption that $f_{ANew} + OA_A > 0$ is reasonable and it does not impede the generality of this model.

Appendix I.2

Bank B's capital ratio constraint says that the capital ratio must be greater than or equal to the target ratio:

$$r_B = \frac{f_B * e_{BB} + \Delta_B f_A + (c_B - \Delta_B \underline{p_A})}{f_B * e_{BB} + \Delta_B f_A + (c_B - \Delta_B \underline{p_A}) + OA_B} \geq r_B^*$$

Set $e_{BB} = 1$ to have the normalized expression for this constraint, and define $f_B + \Delta_B f_A$ as a new normal random variable f_{BNew}, with mean of $\mu_B + \Delta_B \mu_A$ and variance $\sigma_B^2 + \Delta_B^2 \sigma_A^2$. The second constraint of Bank B's problem becomes

$$Pr\left(\frac{f_{BNew} + c_B + \Delta_B \underline{p_A}}{f_{BNew} + c_B + \Delta_B \underline{p_A} + OA_B} \geq r_B^*\right) = q_B$$

which is equivalent to

$$\Pr(f_{BNew} \leq \Delta_B \underline{p_A} - c_B + \frac{r_B^* OA_B}{1 - r_B^*}) = 1 - q_B$$

Here I use the property that Bank B's constraint on r_B must be binding in an optimal solution, as well as the argument in Appendix I.1 regarding the sign of the above inequality. Since f_{ANew} follows a normal distribution, the above equation is equivalent to

$$Pr(z \leq \frac{\Delta_B \underline{p_A} - c_B + \frac{r_B^* OA_B}{1 - r_B^*} - (\mu_B + \Delta_B \mu_A)}{\sqrt{\sigma_B^2 + \Delta_B^2 \sigma_A^2}}) = 1 - q_B$$

Let $K_B = \Phi^{-1}(1 - q_B)$, and thus

$$\frac{\Delta_B \underline{p_A} - c_B + \frac{r_B^* OA_B}{1 - r_B^*} - (\mu_B + \Delta_B \mu_A)}{\sqrt{\sigma_B^2 + \Delta_B^2 \sigma_A^2}} = K_B$$

Rearrange to obtain the liquidity supply curve of Bank B,

$$\underline{p_A} = \mu_A + \frac{K_B \sqrt{\sigma_B^2 + \Delta_B^2 \sigma_A^2} - (\frac{r_B^* OA_B}{1 - r_B^*} - c_B - \mu_B)}{\Delta_\mathrm{B}}$$

Note that

$$\frac{\partial \underline{p_A}}{\partial \Delta_\mathrm{B}} = \frac{T_2 \sqrt{\sigma_B^2 + \Delta_B^2 \sigma_A^2} - K_B \sigma_B^2}{\Delta_B^2 \sqrt{\sigma_B^2 + \Delta_B^2 \sigma_A^2}}$$

where $T_2 = \frac{r_B^* OA_B}{1 - r_B^*} - c_B - \mu_B$.

Appendix I.3

Given that $\widetilde{r_B}$ is the $1 - q_B$ percent quantile of r_B before Bank B purchases any risky portfolio, it follows that

$$Pr\left(\frac{f_B + c_B}{f_B + c_B + OA_B} \leq \widetilde{r_B}\right) = 1 - q_B.$$

With similar argument about the sign of $f_B + c_B + OA_B$ in Appendix I.1, rearrange the above expression to be

$$Pr\left(f_B \leq \frac{\widetilde{r_B}(OA_B + c_B) - c_B}{1 - \widetilde{r_B}}\right) = 1 - q_B.$$

As $f_B \sim N(\mu_B, \sigma_B)$, the above equation is equivalent to

$$P\,r\left(z \leq \frac{\frac{\widetilde{r_B}(OA_B + c_B) - c_B}{1 - \widetilde{r_B}} - \mu_B}{\sigma_B}\right) = 1 - q_B.$$

Since $K_B = \Phi^{-1}(1 - q_B)$, this implies that

$$\frac{\frac{\widetilde{r_B}(OA_B + c_B) - c_B}{1 - \widetilde{r_B}} - \mu_B}{\sigma_B} = K_B.$$

Rearrange to obtain

$$\widetilde{r_B} = \frac{c_B + \mu_B + K_B\sigma_B}{c_B + \mu_B + OA_B + K_B\sigma_B}.$$

Using the same method, one can show that $\widetilde{r_A}$, the $1 - q_A$ percent quantile of r_A before Bank A selling any risky portfolio, is

$$\widetilde{r_A} = \frac{\mu_A + K_A\sigma_A}{\mu_A + K_A\sigma_A + OA_A}.$$

Appendix II. 1

Define a new implicit function of Δ_B and σ_A,

$$F(\Delta_\mathrm{B}, \sigma_A) = \mu_A + \frac{K_B\sqrt{\sigma_B^2 + \Delta_B^2\sigma_A^2} - (\frac{r_B^* OA_B}{1 - r_B^*} - c_B - \mu_B)}{\Delta_\mathrm{B}} - \underline{p_A} = 0$$

Thus,

$$\frac{d\Delta_B}{d\sigma_A} = -\frac{\frac{\partial F(\Delta_\mathrm{B}, \sigma_A)}{\partial \sigma_A}}{\frac{\partial F(\Delta_\mathrm{B}, \sigma_A)}{\partial \Delta_\mathrm{B}}} = -\frac{\frac{K_B\sigma_A\Delta_B}{\sqrt{\sigma_B^2 + \Delta_B^2\sigma_A^2}}}{\frac{T_2\sqrt{\sigma_B^2 + \Delta_B^2\sigma_A^2} - K_B\sigma_B^2}{\Delta_B^2\sqrt{\sigma_B^2 + \Delta_B^2\sigma_A^2}}} < 0$$

where only $K_B < 0$ and $T_2 = \frac{r_B^* OA_B}{1 - r_B^*} - c_B - \mu_B < 0$.

Examining Competitve Balance and Assessing Strategies to Improve it in Sports Leagues: A Model with Profit-Maximizing Teams

Zachary Nash

Abstract

This paper outlines a simple profit-maximization model for a sports league with *n* teams which explains that talented players concentrate in large market teams. This reproduces one of the worries of many sports leagues – that varying market sizes reduce competitive balance. It provides a framework for investigating the effectiveness of salary caps and shared revenue systems in sports leagues. It finds that neither strategy is effective at increasing competitive balance. It also finds that leagues with high TV revenues as a share of total revenues will have better competitive balance.

Acknowledgements

I would like to thank Professors Pavlov, Livshits, and Conley, as well as my peers in Economics 4400, for their assistance and guidance throughout the year. Also, I thank Professor Slivinski for helping immensely in the development of the model presented.

1. Introduction

Team sports leagues are one of the few forms of legalized business cartels that we witness in a society draped in Anti-Trust regulation (El-Hodiri and Quirk 1971). Because of this protection from Anti-Trust regulation, sports leagues provide economists with a number of unique natural experiments to investigate how profit-maximizing firms interact with one another (Syzmanski 2003). Sports leagues control how many teams are allowed to operate within their respective leagues by directly controlling the number of entrants, and thus directly controlling competition within the league. For most industries, this would violate Anti-Trust regulations; however, for sports leagues this is not the case (Fort and Quirk 1995). Sports fans prefer to attend games where the outcome is uncertain. For a league to be successful, teams must be close in competition since weak teams create negative externalities for strong teams (Dietl, Grossman, and Lang 2011; Crooker and Fenn 2007).

Unlike other industries, sports leagues depend on close competition between teams to survive. Sports are entertainment – they are driven by close games, close races to make playoffs, unpredictability in the playoffs, and by opportunities for underdogs to win games (Dietl, Lang, and Rathke 2011). With perennial winners, sports leagues can lose their fan base. The dichotomy within any sports league is that while individual teams need the league to be successful through close competition, their profits are driven by winning. Because of this, sports leagues are able to bypass Anti-Trust regulation (El-Hodiri and Quirk 1971).

In most professional sports leagues persistent inequality between teams, often a result of big-market teams having higher revenues to spend on higher quality players, is a chronic issue. Sports teams aim to maximize profits, which can result in large market teams having better teams than small market teams (Fort and Quirk 1995). For this paper, the terms small- and large-market teams are used to describe variation in a team's market size. Small markets have intrinsically lower demand than large markets. League policy-makers are interested in promoting competitive balance, and a number of techniques have been introduced by authorities to create parity in leagues (Fort and Quirk 1995). The main strategies for combating competitive imbalance have been revenue sharing schemes and salary caps.

Competitive balance issues could be observed easily in baseball and hockey before 2003, where big-market teams like the New York Yankees and Toronto Maple Leafs dwarfed the average payroll in their respective leagues (Zimbalist 2002). This problem became even more apparent during the late 1990s in baseball, and in the early 2000s in hockey. In both leagues, standard deviations in team payrolls increased drastically (Wiseman and Chatterjee 2003; Zimbalist 2002). While this did not necessarily lead to anti-competitiveness in hockey (Zimbalist, 2002), it created a severe concentration of success for big-market teams in baseball (Wiseman and Chatterjee 2003).

In each of the four major North American sports leagues – National Football League (NFL), National Hockey League (NHL), National Basketball Association (NBA), and Major League Baseball (MLB) – there are approximately 30 teams. As a result of the high number of teams, there is a large variation in each team's fan base size. These variations may have a variety of sources, including size of a city, sports culture in a city, how many other professional sports teams compete for fans in a city, etc. (Dietl, Grossman, and Lang 2011). There is a plethora of exogenous variables that determine the size of a fan base; however, the important fact is simply that different cities have different sized markets. This results in individual teams having relatively different demands for their respective franchises.

Using a simple Cournot model, this paper investigates how these different demands create inequality between teams and reduce competition in sports leagues. This model can be of interest to competition authorities and league authorities because it provides new insights into the effect of revenue sharing and salary caps on competitive balance. In contrast to previous models, my analysis shows that revenue sharing and salary caps do not improve incentives for small market teams to invest in playing talent. It follows Atkinson, Stanley, and Tschirhart's (1988) profit-maximization model for a league with *n* teams. However, my model will implement increasing marginal costs (Syzmanski and Smith 1997) and a demand function based on the relative quality of a team, rather than number of wins. Intuitively, these changes do not stray far from Atkinson, Stanley, and Tschirhart (1988); these changes in assumptions, however, will lead to much different conclusions. Most literature focuses on leagues using only two teams to examine the effects of competition levels on franchise utility (Késenne 2000a; Késenne 2005; Dietl, Grossman, and Lang 2011; El-Hodiri and Quirk 1971; Vrooman 1995). These models tend to overestimate the positive or negative effect of strategies for improving competitive balance. Through the use of a profit-maximization model with *n* teams, my model will show that Rottenberg's (1956) invariance proposition in sports leagues is incorrect, as revenue sharing systems do change talent distribution in leagues. However, it also confirms Rottenberg's (1956) hypothesis, as my model shows that, while the distribution of talent changes under revenue sharing and salary cap systems, the overall effect on competitive balance is small.

The paper will proceed as follows: Section 2 outlines literature related to competitive balance in sports leagues and analysis of how effective policies are in improving it. Section 3 outlines the profit-maximization model and the implications for leagues with teams with different demand functions, i.e., small- and large-market teams. Section 4.1 outlines and examines how a revenue sharing system affects competitive balance. Similarly, section 4.2 outlines and examines how salary cap implementation affects competitive balance. Then I will show how leagues with large television audiences are more competitive than those without, which could lead to an organic change in competitive balance for leagues, bypassing the use of exogenous strategies for improving competitive balance.

2. Literature

2.1 *Empirical Literature*

2.1.1 Competitive Balance in Sports Leagues

There has been a wide array of work detailing competitive balance in sports leagues. Essentially, there are two broad groups of literature on the subject: investigations into the optimal level of competitive balance in leagues and examinations of how competitive balance has been improved in leagues. For example, Zimbalist (2002) investigates what factors influence competitive balance in each of the major North American professional sports – baseball, hockey, football, and basketball – and summarizes the relative competitiveness of each league. He identifies the optimal level of competitive balance as a combination of the distribution of fan preferences, fan population base, and fan income across cities. He finds that, generally, leagues with control over the number of teams maximize revenues when big-market teams win more often. He identifies several different measurements of competitive balance in sports leagues, which revolve mainly around standard deviations of winning percentages. Zimbalist (2002) concludes that in all major sports leagues there exists problems of competitive balance, and that each league has introduced policies to try to create higher parity rates.

Hamlen (2007) finds that big-market teams, on the margin, have a higher probability of making the playoffs than small-market teams. He uses an empirical approach to investigate the effect of relative wealth on winning percentage in the National Football

League. One prediction in his paper is that teams in smaller markets have a greater incentive to relocate to larger markets, which is evident in North American sports leagues.

Wiseman and Chatterjee (2003) examine the growing disparity among payrolls in Major League Baseball teams. They investigate the relationship between payroll and winning percentage over the time period of 1985 to 2002. They find that the increasing disparity in team payrolls is having an adverse effect on the competitive balance in baseball.

2.1.2 Examining the Effectiveness of Techniques in Creating Parity

In all four of the major North American sports leagues, there are league policies designed to create more equality between teams. These policies are not designed just to close the gap between team profits, but also to create better competition between teams through greater parity in the quality of franchises. The three main overarching policies that sports leagues implement are revenue sharing, luxury taxes, and salary cap systems. Zimbalist (2010) examines the effectiveness of salary caps on salary shares in the four main professional sports in North America. He concludes that salary caps may not be effective at reducing relative salaries, as the salary share of total league revenues is lower in Major League Baseball (MLB) than in the two leagues with stringent salary caps – the National Hockey League (NHL) and the National Football League (NFL). He also examines the effect of revenue sharing in the MLB, which has created a system of incentives for small-market teams to adopt a strategy of having lower payrolls, further increasing the gap between big and small market competitiveness levels.

Booth (2004) finds that both revenue sharing and salary caps in the Australian Football League, which he identifies as having win-maximizing teams, have helped to achieve better levels of competitive balance. Atkinson, Scott, and Tschirhart (1988) examine revenue sharing in the NFL. They conclude that revenue sharing in the NFL has desirable properties; however, the effect is negligible.

2.2 *Theoretical Literature*

Rottenberg's (1956) seminal paper on the invariance proposition claims that revenue sharing does not affect the distribution of talent among profit-maximizing clubs. Through the law of diminishing returns on player quality and the fact that teams benefit from their opponents' quality, he argues that, in a non-collusive market, player talent distribution will not be concentrated in large-market teams. While large-market teams may perform better team than small-market teams, the difference is minimal; consequently, if a revenue sharing system is implemented, it will have minimal effect on the distribution of talent in the league. According to Rottenberg (1956), the only incentive that leagues have for tampering with free agency – such as with a salary cap – is to increase profits for owners. Subsequently, El-Hodiri and Quirk (1971) provide a proof that predicts that the economic structure of professional sports leagues produces competitive imbalance, that is, large markets will have higher quality teams than small markets. While this conclusion is different from Rottenberg's, their conclusion about revenue sharing agreements confirms the invariance proposition given by Rottenberg.

Vrooman (1995) examines Rottenberg's invariance proposition, but incorporates the effects of winning and market size on cost and revenue. He finds that the degree of competitive balance in a sports league depends on the size of these effects. In equilibrium, large-market teams will attract higher quality talent and have better winning percentages than small-market teams. He examines the effects of revenue sharing and a salary cap on the equilibrium of player distribution and competitive balance. He concludes that, while salary caps are effective at creating competitive balance, it may be through the decrease of large-market teams' quality, and not through the increase of small-market teams' quality. So, while competitive balance may be increased through a salary cap, it may be due to the overall effect of decreasing the league's talent supply. Vrooman also finds that revenue sharing does not increase competitive balance.

Késenne (2005) challenges the invariance proposition; if the incentives of revenue sharing parameters are changed so that teams become win-maximizers rather than profit-maximizers, then revenue sharing improves competitive balance. Using a mixed-talent model, he concludes that a pool-revenue sharing arrangement concentrates talent in a league. However, he also finds that in some leagues poorer teams are profit-maximizers and richer teams are win-maximizers. This results in improving competitive balance with revenue sharing. Késenne (2000b) finds that if teams are profit-maximizing firms, then revenue sharing will not improve competitive balance. However, if a team is utility-maximizing, that is, it prefers winning and profits, then revenue sharing can improve competitive balance. In his examination of salary caps, Késenne (2000a) uses a two-team model to examine a sports league. His model indicates that salary caps can improve competitive balance in a league while only marginally disrupting total league revenues and team profits.

Dietl, Grossman, and Lang (2011) provide a convincing argument for utility-maximizing teams with small-market teams' utility based on profit-maximization and large-market teams' utility based on profits and wins. They find that revenue sharing does not necessarily reduce incentives for teams to invest in playing talent. They emphasize the importance a mixed-utility function based on wins and profits for teams, and they point out how their approach differs from previous literature in this regard. However, their approach uses a contest model with only two teams. This does not capture the free-riding effect of having revenues shared between $n > 2$ teams. Similarly, Syzmanski (2004) introduces a Cournot game between two teams. His findings indicate that revenue sharing decreases competitive balance. Through the introduction of a league with $n > 2$ teams, my model will show that revenue sharing will not increase incentives to invest in playing talent in contrast to Dietl, Grossman, and Lang (2011), and will not necessarily improve competitive balance (Szymanski, 2004).

Atkinson, Scott, and Tschirhart (1988) employ a profit-maximization model for a league with n teams. They assume that team revenues are positively correlated with winning and that marginal costs are constant. Under this model, they find that if owners behave as profit-maximizers, then equal revenue sharing maximizes league revenues by optimally distributing talent among teams. My model will augment their model to show that these conclusions are false: revenue sharing and salary caps have little effect on competition levels, and potentially can have negative effects on competitive balance.

3. A Model of Pricing and Franchise Quality in a Sports League

3.1 *A League under no Regulations*

This model reproduces a sports league that has two types of franchises: small and large market. The small-market franchise is specified as follows:

The team spends money on inputs (stadium, players, coaches, etc.), which results in having a team with quality q. Assume that the cost of attaining quality level q is $c(q) = \gamma q^2$, which means the cost of quality is increasing at a rate $2\gamma q$. Szymanski and Smith (1997) indicate that quality costs are highly correlated with player talent, which implies that as teams spend more on players, the talent of the team increases and the team is relatively better. As Lewis, Sexton, and Lock (2007) demonstrate through empirical analysis, increasing player salaries leads to increased ability or quality. $\frac{\partial c(q)}{\partial q} = 2\gamma q > 0$ indicates that diminishing returns on quality leads to increasing marginal costs for teams. Revenue initially comes only from ticket sales, and is given by $p_s d_s$, where d_s is the demand for tickets to a small-market team's games, and p_s is the price of a ticket.

The demand function for a small-market team's tickets is derived as follows: there is a unit measure of potential game attendees (fans) in the team's area whose willingness to pay for tickets is given as v_i. Assume that the payoff to fan *i* of attending a game is:

$$u_i = \left(\frac{q_s}{Q}\right) v_i - p_s$$

where q_s is the quality for the small-market team, and Q is the average quality of teams in the league. The idea here is that if the quality of the team in a city is below average, the payoff to attending its games will diminish, whereas, if the quality is above average, the fan's payoff is increased. This fan utility function follows Atkinson, Scott, and Tschirhart (1988) and Szymanski (2004), who underline the fact that fans prefer winning teams to losing teams. What changes in this utility function is that perceived team success is based on relative talent level. While real 'fanatics' do exist, this payoff function reproduces the notion that a team's relative quality has an impact on attendance, at the margin. Also, note that

$$\frac{\partial\left(\frac{q_s}{Q}\right)}{\partial q_s} = \frac{Q - q_s\left(\frac{\partial Q}{\partial q_s}\right)}{Q^2} = \frac{Q - \frac{q_s}{n}}{Q^2} > 0$$

because $Q = \frac{1}{n}\sum q_j$ means that $\frac{\partial Q}{\partial q_s} = \frac{1}{n}$. This means that improving a team's quality always increases the utility from attending games. Also, assume that the v_i of potential fans are distributed uniformly over the interval [0,1], so the set of fans who buy tickets are those *i* for whom $u_i > 0$, or $v_i > \frac{Qp_s}{q_s}$. This does mean that if other teams get better and yours does not, attendance will be hurt unless you lower the ticket price.

So, the demand for tickets for a small-market team is given by:

$$d_s(p_s,\ q_s, Q) = \begin{cases} 1 - \dfrac{Qp_s}{q_s} & \text{if } q_s > Qp_s \\ 0 \text{ otherwise} \end{cases}$$

The profits of a small-market team are then given by:

$$\pi_s(p_s,\ q_s, Q) = p_s\left[1 - \frac{Qp_s}{q_s}\right] - \gamma q_s^2$$

This model assumes that there are no costs associated with d_s, that is, it is not more costly to have more people come to games. While this assumption is initially false, higher attendance generates extra revenue from beer, food, and merchandise, so we can assume that p_s is the net addition to revenue the team gets from each costumer who buys a ticket. The team can alter p_s by altering the prices of tickets, food, or beer, and it is this composite price that the fans use to decide whether or not to attend.

A big-market team differs from the above in only one way: the demand for its tickets by any one fan is the same as for a small-market team, but there are λ times as many fans in the large market, where $\lambda > 1$, so that demand for a big-market team is:

$$d_l(p_l,\ q_l, Q) = \lambda\left(1 - \frac{Qp_l}{q_l}\right)$$

by the same reasoning. Thus, the profits of a big-market team are:

$$\pi_l(p_l,\ q_l, Q) = p_l\lambda\left[1 - \frac{Qp_l}{q_l}\right] - \gamma q_l^2$$

This model assumes that the costs of quality are the same in both markets, which seems reasonable since the biggest cost in producing high quality is player salaries. Késenne (2004) indicates that under a perfectly competitive labour market, teams are wage takers, so that quality costs are the same across the league. Either type of team then chooses its *p* and q to maximize its profits. The two first-order conditions for the small-market team's profit maximization problem [$\max_{p_s q_s}\pi_s(p_s,\ q_s)$] are as follows:

$$\frac{\partial \pi_s}{\partial p_s} = 1 - \frac{2Qp_s}{q_s} = 0$$

which clearly implies that $q = 2Qp$ is the profit-maximizing relationship between q and p. The first-order condition for q is slightly more complicated, since Q is a function of each team's q. Q' will be substituted for $\partial Q/\partial q_s$ – the derivative of average team quality with respect to this particular team's quality. This results in the following first-order condition for *q*:

$$\frac{\partial \pi_s}{\partial q_s} = -p_s^2 \left[\frac{Q' q_s - Q}{q_s^2}\right] - 2\gamma q_s = 0$$

which can be simplified to

$$-\left(\frac{p_s}{q_s}\right)^2 [Q' q_s - Q] = 2\gamma q_s$$

$$\frac{\partial^2 \pi_s}{\partial {q_s}^2} > 0$$

For every team, $Q' = 1/n$, so the expression in brackets is just

$$= \frac{q_s}{n} - Q$$

$$= -\left[Q - \frac{q_s}{n}\right]$$

$$= -\frac{1}{n}\sum_{j \neq s} q_j \equiv -Q_{-s}$$

That is, Q_{-s} is just the average quality of the league if the team in question (s, in this case) had a quality of 0. So, this first-order-condition can be written as:

$$\left(\frac{p_s}{q_s}\right)^2 Q_{-s} = 2\gamma q_s$$

If we use the first first-order condition to substitute in $1/2Q$ for p/q, we get

$$\frac{Q_{-s}}{4Q^2} = 2\gamma q_s$$

which implies that

$$q_s = \left(\frac{Q_{-s}}{8\gamma Q^2}\right)$$

so that

$$p_s = \frac{q_s}{2Q} = \frac{Q_{-s}}{16\gamma Q^3}$$

These are not ‘closed-form’ expressions, since q_s appears in Q. The Nash equilibrium values of *all* teams’ q and p depend on Q, which in turn depends on the q’s of other teams. But, we can still use the relationships above; in equilibrium, ticket sales are $d_s(p_s, q_s, Q) = \frac{1}{2}$ for the small-market team, because all $v_i > \frac{Qp_s}{q_s} = \frac{Q}{2Q} = \frac{1}{2}$ buy tickets.

The same exercise for a large-market team results in one first-order condition, which implies$q_l = 2Qp_l$, whereas the other first-order condition now implies

$$\lambda\left(\frac{p_l}{q_l}\right)^2 Q_{-l} = 2\gamma q_l$$

so that we get

$$q_l = \frac{\lambda Q_{-l}}{8\gamma Q^2}, p_l = \frac{\lambda Q_{-l}}{16\gamma Q^2}, \text{and } d_l = \frac{\lambda}{2} > d_s$$

As noted, these are not ‘closed-form’ expressions for the equilibrium values of *p* and *q*, stated entirely in terms of exogenous parameters; however, they allow us to answer several questions regarding this model’s predictions about prices, quality, profits, and attendance.

3.2 *Is it true that the l team charges higher prices and has a higher quality team?*

Suppose the answer to the *q* part of the question is no, so that $q_l \leq q_s$. This would imply that

$$\frac{\lambda Q_{-l}}{8\gamma Q^2} \leq \frac{Q_{-s}}{8\gamma Q^2}, \text{so that}$$

$$\lambda Q_{-l} \leq Q_{-s}, \text{and since } \lambda > 1, \text{this implies}$$

$$Q_{-l} < Q_{-s}$$

but the definitions of the Q_{-l} mean that this can only be true if $q_s > q_l$, which is a contradiction of the hypothesis that the opposite is true, so the hypothesis must be false. Thus, the model predicts $q_l > q_s$, as we would expect. This in turn means that

$$p_l = \frac{q_l}{2Q} > \frac{q_s}{2Q} = p_s$$

and the large-market team also charges higher prices, and, since only $v_i > \frac{Qp_l}{q_l} = 1/2$ buy tickets in the large market, the model predicts $d_l = \lambda/2 > 1/2$. Even though ticket prices are higher, the large market team draws more fans to its higher quality franchise.

3.3 *Is it true that l teams earn higher profits than s teams?*

Suppose the l firm chose exactly the same p and q as the s firm. Then its costs would be the same as the s firm's, but its revenues would be higher, since it would be $p_s\,\lambda/2$. This means that even this naïve choice of p_l and q_l would give it higher profits than the small-market firm, so it would only choose the higher p and q that has been shown, which shows that l firms are more profitable than s firms. This can also be shown directly through calculation:

$$\begin{aligned}\pi_s^* &= \frac{p_s}{2} - \gamma q_s^2 \\ &= \frac{q_s}{4Q} - \gamma q_s^2 \\ &= q_s\left[\frac{1}{4Q} - \frac{\gamma Q_{-s}}{q_s \gamma Q^2}\right] \\ &= \frac{q_s}{Q}\left[\frac{1}{4} - \frac{Q_{-s}}{q_s Q}\right]\end{aligned}$$

which has to be positive since $Q_{-s} < Q$ and

$$\begin{aligned}\pi_l^* &= \frac{\lambda p_l}{2} - \gamma q_l^2 \\ &= \frac{\lambda q_l}{4Q} - \gamma q_l^2 \\ &= q_l\left[\frac{\lambda}{4Q} - \frac{\gamma \lambda Q_{-l}}{q_l \gamma Q^2}\right] \\ &= \frac{\lambda q_l}{Q}\left[\frac{1}{4} - \frac{Q_{-l}}{q_l Q}\right]\end{aligned}$$

Since $\lambda q_l > q_s$ and $Q_{-l} < Q_{-s}$, it follows that $\pi_l^* > \pi_s^*$.

Thus, the model reproduces a worry of any sports league: large-market teams have higher profits and better teams than small-market teams. This finding is consistent with empirical analysis of sports leagues (Hamlen 2007; Wiseman and Chatterjee 2003) and is supported by Vrooman's (1995) theoretical model of a sports league. This model fails to capture the increased utility fans associate with close competition, as outlined by Szymanski (2004) and Crooker and Fenn (2007). For future research, a change in the demand structure may be required to take this properly into account.

4. Strategies for Improving Competitive Balance

4.1 *League under Revenue Sharing*

To show the effects of a revenue sharing system on a sports league, I will introduce a tax defined as *t*, which is applied equally to each team in the league. These tax revenues are then pooled and distributed equally among all *n* teams in the league. The revenues for small-market teams in the original model are defined as

$$R_s = \frac{p_s}{2} = \frac{q_s}{4Q}$$

and similarly, the revenues for a large-market team are defined as,

$$R_l = \frac{p_l}{2} = \frac{q_l}{4Q}$$

then the average team revenue must be

$$\bar{R} = \frac{\bar{p}}{2Q} = \frac{\bar{q}}{4Q} = \frac{Q}{4Q} = \frac{1}{4}$$

The new profit function for a small-market team in a league with revenue sharing will be:

$$\pi_s = (1-t)p_s\left[1 - \frac{Qp_s}{q_s}\right] - \gamma q_s^2 + t\bar{R}$$

$$\pi_s = (1-t)p_s\left[1 - \frac{Qp_s}{q_s}\right] - \gamma q_s^2 + \frac{t}{4}$$

$$\frac{\partial \pi_s}{\partial p_s} = (1-t)\left[1 - \frac{2Qp_s}{q_s}\right] = 0$$

which still simplifies to $q = 2Qp$, as in the original model. Similarly,

$$\frac{\partial \pi_s}{\partial q_s} = -(1-t)p_s^2\left[\frac{Q' q_s - Q}{q_s^2}\right] - 2\gamma q_s = 0$$

which results in

$$q_s^* = (1-t)\left(\frac{Q_{-s}}{8\gamma Q^2}\right) = (1-t)q_s$$

where $q_s = \left(\frac{Q_{-s}}{8\gamma Q^2}\right)$. This result clearly shows that small-market teams decrease their quality under a revenue sharing system.

This results in small market-teams' profits being

$$\pi_s = (1-t)\frac{q_s^*}{4Q} - \gamma q_s^{*2} + \frac{t}{4}$$

substituting q_s^* into π_s gives

$$\pi_s = (1-t)\frac{(1-t)q_s}{4Q} - \gamma(1-t)^2 q_s^2 + \frac{t}{4}$$

$$\pi_s = (1-t)^2\left[\frac{q_s}{4Q} - \gamma q_s^2\right] + \frac{t}{4}$$

$$\pi_s = (1-t)^2\pi_s^* + \frac{t}{4}$$

where $\pi_s^* = \frac{q_s}{4Q} - \gamma q_s^2$, which is the profit under the original model for small-market teams. To show the effects of *t* on the profitability of a small market,

$$\frac{\partial \pi_s}{\partial t} = -2(1-t)\pi_s^* + \frac{1}{4}$$

$$= 2t\pi_s^* - 2\pi_s^* + \frac{1}{4}$$

$$= 2\pi_s^*(t-1) + \frac{1}{4} = 0$$

where $\pi_s^* > 0$ and $0 < t < 1$, which implies that $(t-1) < 0$, resulting in $2\pi_s^*(t-1) < 0$. Therefore $\frac{\partial \pi_s}{\partial t} < 0$ if $|2\pi_s^*(t-1)| > \frac{1}{4}$ and $\frac{\partial \pi_s}{\partial t} > 0$ if $|2\pi_s^*(t-1)| < \frac{1}{4}$. While it is clear what the effect of *t* is on q_s, the effect on profits is not as clear. However, it is clear that as the original model profits π_s^* increase (decrease), then the likelihood of the effect of *t* on profits is negative (positive). Intuitively, this result makes sense: small-market teams with smaller profits benefit more from a revenue sharing system, or at least are not as negatively impacted, while larger market teams with high profits are negatively impacted, or at least not as positively impacted, from a revenue sharing system.

The results from this model show how revenue sharing systems can impact competitive balance negatively for sports leagues: small-market teams are induced to spend less on players, making themselves less competitive. While small-market teams may be more profitable through revenue sharing systems, this is not immediately clear from the model. If league authorities are concerned with competitive balance, then revenue sharing systems do not induce small-market teams to spend more on players, and is therefore an ineffective mechanism for making a league more competitively balanced. In contrast to Atkinson, Scott, and Tschirhart (1988), the implementation of increasing marginal costs and fan preferences based on relative quality rather than winning leads to revenue sharing's being ineffective at increasing competitive balance. This follows Rottenberg's (1956) invariance proposition that revenue sharing will not change talent distribution in a sports league.

4.2 *League under a Salary Cap*

To show the effect of a salary cap on sports leagues, I will introduce a ceiling on salary expenditures (which amounts to a ceiling on *q*). This will be defined as $\bar{q}$, where $\bar{q} < q_l$. The model will stay the same, but with *l* teams only being allowed to spend up to $\bar{q}$. Because $\bar{q} < q_l$, *l* teams will spend as much as they can to maximize profits. Thus, $q_l = \bar{q}$, and following the same steps as before shows that $p_l = \frac{\bar{q}}{2Q} < \frac{q_l}{2Q}$.

For a small-market team, the effect of a salary cap implementation would be as follows:

The effect of q_l on $Q_{-s} = \frac{1}{n}\sum_{j\neq s} q_j$ is central to this argument, and is $\frac{\partial Q_{-s}}{\partial q_l} = \frac{1}{n}$, which is the same as $\frac{\partial Q}{\partial q_l} = \frac{1}{n}$. Calculating the effect of a change in q_l on q_s is

$$\frac{\partial q_s}{\partial q_l} = \frac{8\gamma Q^2\left(\frac{\partial Q_{-s}}{\partial q_l}\right) - Q_{-s}\left(\frac{\partial Q}{\partial q_l}\right)}{(8\gamma Q^2)^2} = \frac{8\gamma Q^2\left(\frac{1}{n}\right) - Q_{-s}\left(\frac{1}{n}\right)}{64\gamma^2 Q^4} = \frac{8\gamma Q^2 - Q_{-s}}{64n\gamma^2 Q^4} > 0$$

which implies that an increase in q_l will have a (slightly) positive impact on q_s. Under a salary cap system, $q_l = \bar{q}$ will have the effect of reducing q_s.

The model shows that the relative decrease in q_l is much larger than the decrease in q_s. Thus, we should observe closer competition in a league with a salary cap system in place than in a league without one. However, this model points out that a salary cap may not be an effective way to create higher parity in leagues. Ideally, a salary cap should have decreased q_l, which it did, and also increased q_s, which it failed to achieve under this model. This may explain why salary cap systems have not been as effective at creating parity as league policy makers might have originally anticipated (Zimbalist 2010).

4.3 *League with High Shared Television Revenues*

In 2008, the NFL made an estimated $7.6 billion in total revenues (Fisher 2010). In 2011, the NFL made $4 billion in national television contracts, constituting approximately half of the total league revenues (Bloomberg 2011). Zimbalist (2002) found the NFL to be the most competitive of the four major North American sports leagues. Of the four major North American sports leagues, the NFL has the highest percentage of total revenues from television contracts (Forbes 2011). This TV revenue is divided equally among franchises, and represents a shift in importance away from gate revenues to TV revenues in the incentive structure for sports leagues and their various franchises. The shift from gate revenues to TV revenues in the NFL may explain why it has such high levels of competitive balance.

For this analysis, I will augment the original model slightly: there will be *n* teams, but *only* two different types of teams – small-market teams and large-market teams. There will be *m* number of small-market teams and *n-m* number of large-market teams in the league. The new profit function for a small-market team will be denoted as

$$\pi_s(p_s,\ q_s, Q) = p_s\left[1 - \frac{Qp_s}{q_s}\right] - \ \gamma q_s^2 + \frac{1}{n}TV$$

$$\pi_s(p_s,\ q_s, Q) = p_s\left[1 - \frac{Qp_s}{q_s}\right] - \ \gamma q_s^2 + \frac{1}{n}[\delta[Var(Q)] + c]$$

where $\delta < 0$ and $c > 0$. Also, $c > \delta[Var(q_s)]$ so that $\delta[Var(q_s)] + c > 0$. This model has shared television revenue that all teams benefit from; as the variance of the quality of teams increases, television revenue decreases. The variance of the quality of teams in the league is calculated as

$$E(Q) = \frac{1}{n}\left[\sum_{s=1}^{m} q_s + \sum_{l=m+1}^{n} q_l\right]$$

$$Var(Q) = E(Q^2) - E(Q)^2$$

$$Var(Q) = \frac{1}{n}\left[\sum_{s=1}^{m} q_s^2 + \sum_{l=m+1}^{n} q_l^2\right] - E(Q)^2$$

$$\frac{\partial}{\partial q_i}Var(Q) = \frac{2q_i}{n} - \frac{2}{n}E(Q) = \frac{2}{n}[q_i - E(Q)]$$

If i is large, then $\frac{\partial}{\partial q_i}Var(Q) > 0$ and if i is small, then $\frac{\partial}{\partial q_i}Var(Q) < 0$. If large-market teams increase their quality, the variance of the league quality increases. Conversely, if a small-market team increases its quality, the variance of the league decreases. This will induce small-market teams to spend more on quality players, and large-market teams to spend less on quality players, thus narrowing the quality gap between small-l and large-market teams, and achieving a better competitive balance. This model shows that for leagues like the NFL, where a major portion of revenue is from national television contracts, greater parity among teams may occur, resulting in better competitive balance.

5. Conclusion

This paper addresses a number of issues facing professional sports leagues using a simple profit-maximization model based on fan utility increasing with a relative increase in team quality. This model incorporates the effect of market size in determining the quality of different teams, and the distribution of talent across a league. Recreating a sports league where teams are profit-maximizers, it has illustrated that large-market teams have higher levels of talent and are more profitable than small-market teams – recreating one of the concerns of sports league policymakers: competitive imbalance. The model allows for insight into the effect of strategies used by sports leagues to increase competitive balance: revenue sharing and salary caps. It also provides insight into why leagues with high TV revenues may have better competitive balance than those with low TV revenues.

There have been many investigations into the effectiveness of revenue sharing and salary caps in increasing competitive balance. Some conclude that these strategies do not change competitive balance, e.g., Rottenberg (1956), El-Hodiri and Quirk (1971); some find that they improve competitive balance, e.g., Vrooman (1995), Késenne (2005), Atkinson, Scott, and Tschirhart 1(988); some find that they reduce competitive balance, e.g., Dietl, Grossman, and Lang (2011). The model used in this paper shows that revenue sharing is ineffective at increasing competitive balance and may, in fact, reduce incentives to invest in talent. It also finds that salary caps may impact competitive balance positively, but may not have a significant overall effect. In investigating the impact of high TV revenues on a league, this model finds that leagues with high TV revenues may have better competitive balance than leagues that rely primarily on gate revenue.

References

Atkinson, Scott E., Linda R. Stanley, and John Tschirhart. 1988. "Revenue Sharing as an Incentive in an Agency Problem: An Example from the National Football League." *The RAND Journal of Economics* 19 (1): 27-43.

Booth, Ross. "Labour Market Intervention, Revenue Sharing and Competitive Balance in the Australian Football League, 1897-2002." In *International Sports Economics Comparisons*, edited by Rodney Fort and John Fizel, Chapter 18. Greenwich, CT: Praeger Publishers, 2004.

Crooker, John R., and Aju J. Fenn. 2007. "Sports Leagues and Parity: When League Parity Generates Fan Enthusiasm." *Journal of Sports Economics* 8 (2): 139-164.

Dietl, Helmut M., Martin Grossman, and Markus Lang. 2011. "Competitive Balance and Revenue Sharing in Sports Leagues with Utility-Maximizing Teams." *Journal of Sports Economics* 12 (3): 284-308.

Dietl, Helmut M., Markus Lang, and Alexander Rathke. 2011. "The Combined Effect of Salary Restrictions and Revenue Sharing in Sports Leagues." *Economic Inquiry* 49 (2): 447-463.

El-Hodiri, Mohamed, and James Quirk. 1971. "An Economic Model of a Professional Sports League." *Journal of Political Economy* 7 9(6): 1302-1319.

Fort, Rodney, and James Quirk. 1995. "Cross-Subsidization, Incentives, and Outcomes in Professional Team Sports Leagues." *Journal of Economic Literature* 33 (3): 1265-   1299.

Fisher, Jake. 2010. “The NFL’s Current Business Model and the Potential 2011 Lockout.” Department of Economics, Harvard University, Cambridge, MA.

Hamlen, William. 2007. “Deviations from Equity and Parity in the National Football League.” *Journal of Sports Economics* 8 (6): 596-615.

Késenne, Stefan. 2000a. “The Impact of Salary Caps in Professional Team Sports.” *Scottish Journal of Political Economy* 47 (4): 422-430.

Késenne, Stefan. 2000b. “Revenue Sharing and Competitive Balance in Professional Team Sports.” *Journal of Sports Economics* 1 (1): 56-65.

Késenne, Stefan. 2004. “Competitive Balance and Revenue Sharing: When Rich Clubs Have Poor Teams.” *Journal of Sports Economics* 5 (2): 206-212.

Késenne, Stefan. 2005. “Revenue Sharing and Competitive Balance: Does the Invariance Proposition Hold?” *Journal of Sports Economics* 6 (1): 98-106.

Lewis, Herbert, Thomas Sexton, and Kathleen Lock. 2007. “Player Salaries, Organizational Efficiency, and Competitiveness on Major League Baseball.” *Journal of Sports Economics* 8 (3): 266-294.

Rottenberg, Simon. 1956. “The Baseball Players’ Labor Market.” *Journal of Political Economy*, 64 (3): 242-258.

Szymanski, Stefan. 2004. “Professional Team Sports Are Only a Game: The Walrasian Fixed-Supply Conjecture Model, Contest-Nash Equilibrium, and the Invariance Principle.” *Journal of Sports Economics* 5 (2): 111-126.

Szymanski, Stefan, and Ron Smith. 1997. “The English Football Industry: Profit, Performance and Industrial Structure.” *International Review of Applied Economics* 11 (1): 135-153.

Vrooman, John. 1995. “A General Theory of Professional Sports Leagues.” *Southern Economic Journal* 61 (4): 971-990.

Wiseman, Frederick, and Sanjit Chatterjee. 2003. “Team Payroll and Team Performance in Major League Baseball: 1985-2002.” *Economics Bulletin* 1(2): 1-10.

Zimbalist, Andrew. 2002. “Competitive Balance in Sports Leagues: An Introduction.” *Journal of Sports Economics* 3 (2): 111-121.

Zimbalist, Andrew. 2010. “Reflections on Salary Shares and Salary Caps.” *Journal of Sports Economics* 11 (1): 17-28.