FEDERAL RESERVE BANK of ATLANTA

Business Cycles: A Role for Imperfect Competition in the Banking System

Federico S. Mandelman

Working Paper 2006-21
November 2006

WORKING PAPER SERIES

FEDERAL RESERVE BANK *of* ATLANTA WORKING PAPER SERIES

Business Cycles: A Role for Imperfect Competition in the Banking System

Federico S. Mandelman

Working Paper 2006-21
November 2006

Abstract: This paper studies the cyclical pattern of ex post markups in the banking system using balance-sheet data for a large set of countries. Markups are strongly countercyclical even after controlling for financial development, banking concentration, operational costs, inflation, and simultaneity or reverse causation. The countercyclical pattern is explained by the procyclical entry of foreign banks, which occurs mostly at the wholesale level and signals the intention to spread to the retail level. My hypothesis is that wholesale entry triggers incumbents' limit-pricing strategies, which are aimed at deterring entry into retail niches and which, in turn, dampen bank markups. In the second part of the paper, I develop a general equilibrium model that accounts for these features of the data. I find that this monopolistic behavior in the intermediary financial sector increases the volatility of real variables and amplifies the business cycle. I interpret this bank-supply channel as an extension of the credit channel pioneered by Bernanke and Blinder (1988).

JEL classification: C23, E32, G21, L12, O16

Key words: countercyclical bank markups, limit pricing, business cycles, panel data, generalized method of moments

---

The author gratefully acknowledges Fabio Ghironi, Peter Ireland, and Fabio Schiantarelli for their invaluable guidance. He also thanks Ingela Alger, Stephen Bond, Luisa Lambertini, Hideo Konishi, Todd Prono, Julio Rotemberg, Richard Tresch, and colleagues at the New York Fed, the Atlanta Fed, the R@BC workshop at Boston College, Universidad Nacional de la Plata, Universidad del CEMA, the University of Torcuado Di Tella, Universidad de San Andrés, the Ninth World Congress of the Econometric Society, and the central bank of Argentina. The views expressed here are the author's and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Any remaining errors are the author's responsibility.

Please address questions regarding content to Federico Mandelman, Research Economist and Assistant Policy Adviser, Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree Street, N.E., Atlanta, GA 30309, 404-498-8785, federico.mandelman@atl.frb.org.

Business Cycles: A Role for Imperfect Competition in the Banking System.

1 Introduction

A variety of theoretical models and ample empirical evidence support the existence of countercyclical markups in real goods markets.[1] In turn, these countercyclical markups constitute an important internal propagation mechanism in business cycle models. This paper is concerned with financial markets, for which practically no work on the cyclicality of the markups exists. The questions this paper addresses are:

1- Is there a countercyclical pattern in financial markups?

2- If yes, what generates this pattern?

3- What are the implications of these countercyclical markups for the real economy?

To test the first assertion, I use bank data across 124 countries for the years 1990-2000. I use dynamic panel techniques to confront the potential bias induced by simultaneity or reverse causation, and examine whether the exogenous component of GDP growth negatively affects bank markups. Since past work shows that long-run economic growth is a good predictor of financial development that enhances competition and thus erodes markups, I control for a three-year average of financial development to isolate the business cycle component.

To assess the strength of an independent link between the markups and the business cycle, I use various conditioning variables that include a proxy for concentration, overhead costs (operative and administrative costs), inflation volatility, and changes in real interest rates. The results remain robust to any of these factors. The countercyclical behavior of the markups vanishes, however, when controlling for the entry of foreign banks, which happens to be highly procyclical.

I rely on this last result to support my hypothesis regarding the second question. In the last decade, banks have expanded internationally by establishing foreign subsidiaries and branches or by taking over established banks.

[1] For instance, see Pigou (1927), Keynes (1939), Rotemberg and Saloner (1986), Rotemberg and Woodford (1992), Chevallier and Sharfstein (1996), Galeotti and Schiantarelli (1998) among many others.

It is well-documented that foreign entry initially occurs at the wholesale level, but with the final intention to spread to the retail niches.[2] Therefore, we can predict that the threat of foreign banks encroaching on retail markets may induce greater efficiency of the established banks at the retail level. The penetration into the retail sector is obstructed, however, by the need to incur large sunk entry costs (for instance, large advertisement expenditures or the construction of a network of branches and ATMs required to accommodate small transactions). This implies that banks need to enter at a minimum-efficient-scale (MES) to justify the sunk costs incurred. In turn, it follows that right after entering they must capture a large enough fraction of the market to make the constructed network profitable. This is particularly difficult in the banking industry since the markets are highly segmented into regional or sectorial niches (Petersen and Rajan, 1994). In this scenario, the size of the market constitutes a barrier to entry. If the relevant financial market is small or underdeveloped there is space for only few incumbents working at an efficient scale. Thus, during recessions, the banks in the local financial system are able to exert their monopolistic power by charging high markups. In contrast, boom periods lead to an expansion of the financial system that attracts potential competitors who can operate at an efficient scale. In this situation, contestable markets force incumbents to charge markups well below short-run profit maximizing levels to avoid entry. As shown in Bain (1956), pricing decisions strongly influence firms contemplating entry and justify limit-pricing strategies that are counter to short run profit maximization. With this idea in mind, I extend the empirical analysis to show that the competitive pressure of entry is short-lived. The aim is to show that foreign entry affects markups by triggering pricing strategies among incumbents rather than by transforming an existent monopolistic market structure.

Regarding the third question, the cyclicality of the markups may help to explain evidence that suggests an important role for financial development in the magnitude of the business cycle [3]. If bank markups are countercyclical, then there is a bank-supply channel that extends the credit channel to reinforce the same vicious circle: Credit is more expensive during recessions, and firms and households postpone investment and work decisions and make the recession deeper. But while the standard version of the credit channel relies

[2] See Claessens et al (2001).

[3] For a Survey, see Gertler and Hubbard (1988).

on an external finance problem that induces banks to charge a premium to cover the increasing expected bankruptcy costs during recessions, the bank-supply channel is solely the result of imperfect competition in the banking system. This channel may be particularly relevant in developing countries, in which bank credit remains the primary source of funds for entrepreneurs. The last section of this paper develops a dynamic stochastic general equilibrium model that is designed to highlight the macro implications of the limit pricing scheme that constitutes the bank-supply channel. The microfoundations of the banking system embedded in the general equilibrium setup account for several features of the data.

The paper is organized as follows. In section 2, I discuss the methodology to measure bank markups and provide a literature review. In section 3, I present the empirical results. In section 4, I introduce the theoretical model. Concluding remarks are in section 5.

2 Markups Measurement and Literature Review

The first step in answering the three questions posed is to find a proper measure for markups in the banking industry data. A simple approach is to consider the ex-ante (posted) spreads or the difference between lending and deposits rate, as a proxy for financial markups. The difficulty here is that the spreads include a risk premium to cover expected borrowers' bankruptcy costs that increase during recessions and cause the spread to be countercyclical.

However, we expect that, in the long run, aggregate bank income obtained from such risk premium charges match banks' loan default costs. Therefore, I use annual bank balance sheet ex-post data that accounts for defaulted loans to proxy for net markups. In particular, I use net interest margins (NIM), equal to bank's total interest income minus interest expense over total assets after subtracting defaulted loans. Other issues remain though. Some of the loan contracts are settled for even longer periods of time and at predetermined rates. It may be the case that during recessions, riskier entrepreneurs facing liquidity constraints are more prone to demand credit. Higher margins would reflect the premium obtained for new riskier loans that would not necessarily fall in default during the year in consideration. Nonetheless, defaults

are much more likely to occur in recessions. Following this line of argument, if a particular bank liked to enter into long-term contracts at predetermined rates, one would expect that the default frequency for all its loans would significantly increase during recessions, driving down its net interest income and offsetting any positive effect from the new loans. To sum up, evidence of increasing margins in bad times would provide support to my idea of countercyclical markups. As explained in Demirguc-Kunt and Huizinga (1998), bank interest margins can be seen as an indicator of the pure inefficiency of the banking system.

Practically all the existent literature is focused on ex-ante spreads. Related to this paper, Hannan and Berger (1991) find that after a monetary contraction ex-ante spreads tend to increase more in regional U.S. markets in which the banking industry is more concentrated.[4] In addition, Angellini and Cetorelli (2003) consider the growth of GDP as an additional control variable in the estimation of Lerner indexes for the Italian banking industry, finding a negative association. However, they do not settle the issues of causality and endogeneity.

Although the cyclicality of the markups does not receive particular attention, there is an enormous literature on bank structure and efficiency. This literature contains ambiguous results. In a survey Rhoades (1977) expresses "disbelief and frustration" in the overall inability to link concentration and efficiency. New surveys and studies reach the same conclusions.[5] Contradictory results must be preceded by contradictory theories. The intuition of the Structure-Conduct-Performance (S-C-P) hypothesis is straightforward: A more concentrated market lowers the costs of collusion and fosters tacit or explicit collusion on the part of the banks. In contrast, the Efficient-Structure (E-S) hypothesis predicts efficiency gains from market consolidation. Firms possessing a comparative advantage in production become large and, as a natural consequence, the market becomes more concentrated. Such cost differences may be due to differences in technological or managerial skills. The effect is amplified because of large economies of scale existent in the bank industry. They are derived from risk diversification, lower average adminis-

[4] See also Edwards and Vegh (1997) and Olivero (2004) for additional references.

[5] More contradictive results are found in Berger and Hannan (1989); Neumark and Sharpe (1992); Smirlock (1985); Grady and Kyle (1979). For a new survey, see Bank for International Settlements 2001. In a worldwide analysis Demirguc-Kunt et al (2003) found, at the same time, high net interest margins associated with both small banks and banks with a large market share.

trative costs (Demsetz 1973), and the efficient use of large sunk costs like the construction of large networks of branches and ATM's (Cerasi et al, 1997). Additionally, Gilligan et al (1984) provide evidence that banking is characterized by economies of scope from joint production of financial services. Finally, in the absence of restrictions on entry, excessive inefficient profits are precluded (Baumol, 1982). These conclusions led to a new literature aimed at finding evidence of efficiency gains resulting from mergers and acquisitions. The results are, once again, ambiguous and inconclusive.[6]

Results regarding bank entry deregulation and efficiency are instead mostly unambiguous and conclusive. Several surveys hold that new legislations that remove substantial entry barriers and expose national banking markets to potential new entrants produce pro-competitive effects and reduce margins. Besides, banks lose market power following financial liberalization even in if the banking industry remains highly concentrated.[7]

A more interesting result about deregulation is related to the timing of the efficiency gains. In a study of over 80 countries, Claessens et al (2001) find that foreign bank entry is significantly associated with a reduction in domestic bank profitability. However, the impact of foreign bank entry on local bank competition is "..felt immediately upon entry decision is taken rather after they have gained substantial market share." Angelini and Cetorelli (2003) find that net interest margins declined sharply immediately after a banking reform was made effective in Italy. Similarly, Shaffer (1993) analyzes the impact of the Bank Act Revisions in Canada and finds evidence of an "unexpected supercompetitive state" right after entry deregulation, with negative bank markups observed. The author concludes that such atypical outcome ".. is not consistent with long run equilibrium behavior under known static or dynamic models of profit maximization; and it may simply reflect a temporary disequilibrium...(which) may warrant further study." These last three studies resemble my hypothesis of limit pricing.

To summarize, the evidence fits well with my limit-pricing hypothesis: a) Bank spreads are more countercyclical in concentrated markets; b) When bank systems are exposed to potential competition, efficiency gains are im-

[6] See for instance, Focarelli et al (1999), Petersen and Rajan (1994).

[7] For a survey see Vives (1991) and Demirguc-Kunt (1998). Also Spiller and Favaro (1984) focus on the pro-competitive impact of the relaxation of entry restrictions in the Uruguayan industry, concluding that collusive strategic interactions across banks significantly decrease after the regulatory reform. Ribon and Yosha (1999) reach similar conclusions for the Israeli banking industry.

mediately observed and occur long before any change in the market structure is registered c) There exists an ambiguous and contradicting relationship between concentration and efficiency. That is, my hypothesis predicts that incumbents experience periods of monopolistic markups (i.e. recessions) followed by periods in which the efficiency gains from consolidation and exploitation of economies of scale and scope prevail (i.e. booms).

3 Evidence on Markups and Foreign Entry

I construct an unbalanced panel from several data sources. The resulting sample covers 124 countries during the years 1990-2000. Bank structure information is taken from Scope Database provided by IBCA, which contains data for 137 countries. To ensure reasonable coverage, only countries with at least three banks in a given year are included. Coverage by IBCA is comprehensive, accounting for roughly 90% of the assets of banks in each country. Each country has its own data template that allows for differences in account conventions. However, these are converted to a format which is a globally standardized template derived from the country-specific templates. In the regressions, I control for unobserved time-invariant country-specific effects to account for the minor differences in the valuation of assets that necessarily remain.

Measures of the activity of financial intermediaries are taken from the Levine-Loayza-Beck Data Set. Macroeconomic data comes from the Penn World Table 6.1 (PWT 6.1.). Data on real interest rates is taken from the World Development Indicators 2002. Variable definitions and a few descriptive statistics are provided in the statistical appendix. The degree of financial development captured by the ratio of *Private Credit/ GDP* is significantly larger in developed countries than in developing ones. On average, margins are 571 bp (5.71%) for developing countries and 268 bp (2.68%) for developed ones. Besides, these are much more volatile in developing economies. Poorer countries also have a relatively high degree of concentration and foreign penetration.

Econometric Methodology The aggregation of time series would obscure underlying microeconomic dynamics, whereas panel data techniques

allow for the investigation of heterogeneity in adjustment dynamics between different types of countries.

The estimation procedure needs to tackle some important issues. As mentioned, I must allow for the presence of unobserved country-specific effects that are correlated with the regressors. Besides, most of the explanatory variables in the specifications to be used (e.g. GDP growth rates, private credit, etc.) are determined jointly with the dependent variable (i.e. net interest margins). Therefore, I must also allow and control for joint endogeneity. To confront these issues I use a GMM in differences specification that controls for endogeneity by using internal instruments (i.e. instruments based on lagged values of the explanatory variables).

However, other concerns remain. I need to use a dynamic specification to allow for the inertia in the dependent variable that is likely to be present in the annual balance-sheet information. In this scenario, lagged levels of the variables are weak instruments for the regressions in differences. Therefore, I proceed with an auto-regressive dynamic model and use a relatively efficient GMM system specification that adds equations in levels to the ones in differences. In this case, the instruments are given by the lagged differences of the corresponding variables.

In addition, the small sample size raises a concern for overfitting bias. For robustness, I alternate the number of lags used as instruments and restrict the quantity of explanatory variables in the estimations. For further reference, a detailed explanation of all the econometric methodology to be used in this study can be found in Arellano and Bond (1991), Blundell and Bond (1997) and Bond (2002).

The dynamic specification consists on an autoregressive-distributed lag model:

$$y_{it} = \alpha y_{i,t-1} + \beta x_{it} + (\eta_i + \varepsilon_{it}) \qquad | \alpha | < 1 \qquad i = 1, 2, ..., N; \qquad t = 2, 3, ..., T. \tag{1}$$

Where y_{it} denotes net interest margins (NIM) for country i in period t. x_{it} is a vector of current and lagged values of additional explanatory variables and is assumed to be endogenous. η_i is an stochastic unobserved country-specific time-invariant effect. ε_{it} is a disturbance term that is assumed to be serially uncorrelated and independent across individuals. For robustness, time dummies are included to account for time-specific effects.

Net interest Margins -Time Series Properties- I first consider a very simple AR(1) specification. The first two columns of Table 2 report OLS levels and Within Group estimates of the auto-regressive parameter α, along with heteroskedasticity-consistent estimates of the asymptotic standard errors. For panels (like this one) in which the number of time periods available is small, standard results for omitted variable bias indicate that the OLS estimator is biased upwards and the within groups one is biased downwards.[8] Therefore, a consistent candidate estimator must lie between the OLS and within groups estimates. However, the differenced GMM estimator is found to be significantly below the lower bound indicated by the within groups estimator. These downward biases in differenced GMM estimates of the AR(1) are consistent with the finite sample biases expected in the case of highly persistent series. The preferred specification is clearly the GMM system estimator. With the introduction of the equations in levels, I obtain a remarkable improvement in the precision of the parameter estimates. The results indicate a large degree of persistence in the net interest margins (.728). The inertia may arise from lagged effects of the explanatory variables, which is to be expected in balance sheets data with annual frequency. We can also observe that the assumption that the disturbances are serially uncorrelated cannot be rejected. As this model is overidentified, I use the Sargan statistic to test the validity of the overidentifying restrictions. In this case, I obtain a chi-square statistic, which gives the reported p-value of 0.123. The null hypothesis that these moment conditions are valid is not rejected at any conventional level, consistent with the first and second order serial correlations tests for first-differenced residuals.

Basic Model A simple approach, and a first step, in studying the pattern of the margins throughout the business cycle is to include GDP growth in the AR(1) model. Prior to presenting the results, I would like to clarify the interpretation. To the extent that the assumptions regarding the instruments employed are correct, the econometric methodology is designed to isolate the effect of the exogenous component of the explanatory variable on the interest margins. Hereafter, when I mention the impact or effect of a given variable on the margins, I am referring to this isolated exogenous component and not

[8] If T is small, the within transformation induces a non-negligible negative correlation between the transformed lagged dependent variable $y_{i,t-1} - \frac{1}{T-1}(y_{i1} + .. + y_{it} + ... + y_{i,T-1})$ and the transformed error term $\varepsilon_{i,t-1} - \frac{1}{T-1}(\varepsilon_{i2} + .. + \varepsilon_{i,t-1} + ... + \varepsilon_{i,T})$.

merely describing the association between them. In Table 3, *GROWTH* has a highly significant and negative effect on the margins despite their persistence. According to the preferred system specification, an increase in income of 10% causes the margins to fall by approximately 1% (100 bp) on impact.

The literature presents extensive evidence of a strong link between long-run economic growth and financial development. It may be that financial development that enhances competition explains the negative relationship mentioned above. In order to assess and control for financial development, I include a three-year-overlapping average of private credit offered by commercial banks (*PRIV.CRED (avg)*) in the conditional set.[9] Notice that by computing averages of this dependent variable we are artificially constructing a persistent series that should not be affected by short-run fluctuations. Finite sample bias is therefore expected in the differenced specification. In contrast, the preferred system specification shows this variable to have a significant and sizeable negative effect on the margins. *GROWTH* remains significantly at a 10% level though with a slightly lower coefficient. This result supports the hypothesis of fluctuating margins at a business cycle frequency.

Sensitivity analysis I use various conditioning information sets to assess the strength of the countercyclical nature of the margins depicted in the basic model. I start by introducing a proxy for concentration as a control variable, the assets of the three largest banks as a share of the assets of all the commercial banks in the system. The variable is significant and again enters with a negative sign. Refer to Table 4. These results support the implications of the E-S hypothesis which predicts operational efficiency gains from banking consolidation. The large number of explanatory variables accompanied by a relative large p-value for the Sargan test estimates raises a concern about overfitting bias. However, no clear pattern in the coefficient estimates is observed when reducing or increasing the number of instruments.

If the negative effect of concentration on margins is explained by efficiency gains, I would expect this impact to vanish when controlling for operating costs. Therefore, I expand the conditioning set and include overhead costs *(OVERCOSTS)*. They are defined as personnel expenses (mostly wages) and other non interest expenses divided by the total bank assets in the local bank

[9] *Three-year-overlapping* averages are calculated as: $x_{it(avg)} = \frac{x_{it}+x_{i,t-1}+x_{i,t-2}}{3}$.

system.

As expected, large operating costs cause margins to increase. The variable enters significatively and with a sizeable positive coefficient. In support to the E-S hypothesis, the inclusion of this variable causes the concentration coefficient to become insignificant and small in sign. Although the size of *GROWTH* is reduced, this cyclical component remains significant.

Saunders and Shumacher (2000) show that interest rate volatility, usually observed in a context of high and variable inflation, is positively related to margins. Thus, I proceed by adding inflation and real interest rates to the conditioning set. Neither any of these variables turns out to be significant, nor explain the cyclicality of the markups. See Table 4 again.[10]

The Role of Foreign Entry Up to this point, I have shown that the exogenous cyclical component of economic growth is negatively associated with net interest margins. Moreover, this link is not due to potential biases induced by omitted variables (including that derived from unobserved country specific effects), simultaneity, or reverse causation. In the next step, I test the main hypothesis proposed in the paper. That is, countercyclical markups are the result of a limit pricing strategy aimed at deterring procyclical entry of competitors in a segmented local financial system. As I explained in the introduction, although the threat of entry is a non-measurable concept, foreign penetration may be considered a good proxy for it. Consequently, I would expect the negative association between margins and economic growth to vanish when controlling for foreign entry.

Thus, I introduce foreign entry in the conditioning set. The covariate $ForeignBanks$ refers to the number of foreign banks divided by the total number of banks in a given country. Foreign bank entry is measured as a *change* in foreign bank presence (i.e. $\Delta ForeignBanks_{it}$). The first experiment, not reported here, consisted in introducing $\Delta ForeignBanks_{it}$ into the extended model presented in the last subsection. Its influence turns out to be negligible and statistically insignificant. The results are different when

[10]It is puzzling to observe that the coefficients for *GROWTH* and *PRIV.CRED (avg)* actually increase when these covariates are included to the conditioning set. It may be the result of money-based disinflation programs being accompanied by short-lived recessions. These events would imply, at the same time, higher margins due to the recession but lower margins and more credit availability resulting from stable and low inflation. Thus, if we do not control for inflation, we would expect margins to be less countercyclical and less sensible to variables linked with growth indicators.

I consider $\Delta ForeignBanks_{i,t-1}$. It may be the case that the beginning of wholesale operations occurs some time after the official entry registration occurs. Thus, if limit pricing exists, it happens one year after the entry decision is effectively taken. These results may also provide support for the supposition that entry occurs at a wholesale level and then spreads to retail niches with a time lag.

In order to understand the notation, notice that in this last case the proposed model is:

$$y_{it} = \alpha y_{i,t-1} + \beta x_{it} + \gamma \Delta ForeignBanks_{i,t-1} + (\eta_i + \varepsilon_{it}) \quad i = 1, 2, ..., N;\ t = 2, 3, ..., T. \tag{2}$$

where y_{it} is the dependent variable, x_{it} any of the controlling sets already introduced, and $\Delta ForeignBanks_{i,t-1} = ForeignBanks_{i,t-1} - ForeignBanks_{i,t-2}$.

Alternatively, (2) can be expressed as:

$$y_{it} = \alpha y_{i,t-1} + \beta x_{it} + \gamma_1 ForeignBanks_{i,t-1} + \gamma_2 ForeignBanks_{i,t-2} + (\eta_i + \varepsilon_{it}). \tag{3}$$

Where $\gamma_1 = -\gamma_2$.

With this procedure I would like to eliminate the possibility that the negative effect on margins is a consequence of pro-efficiency gains from a larger presence of foreign banks in the local financial structure. In other words, if the results are driven by *entry* we expect the coefficients, γ_1 and γ_2, preceding $ForeignBanks_{i,t-1}$ and $ForeignBanks_{i,t-2}$ to be significant and of the same magnitude, but with opposite signs (i.e. the first one negative and the second one positive). In contrast, if the results are driven by the *presence* of foreign banks, γ_2, must either be negative or at least small. That is, the long-run or steady state effect $(\gamma_1 + \gamma_2)/1 - \alpha$ should significantly differ from zero.

Once again, the results do not reject my hypothesis. Refer to Table 5. The two coefficients are opposite in sign and do not significantly differ in absolute value. These lagged variables not only exert a significant negative effect on the margins, but also break down the independent impact of

GROWTH by turning it small and insignificant. The remarkable colinearity of the covariates involved, affects the statistical significance of all of them. To conclude, the pro-competitive effect of entry in the local banking system is short-lived and vanishes after one year.

4 Theoretical model

In this section I present a simple general equilibrium model designed to highlight the role of the proposed bank-supply channel in the economy. I start from a standard DSGE Real Business Cycle model with variable labor supply in the spirit of Hansen (1985). Then, I introduce imperfect competition with limit pricing in the financial system. This modification creates a disintermediation between borrowers and entrepreneurs that amplifies the response of the real variables to technology shocks.

4.1 Households

The household sector is conventional. There is a continuum of households of unit mass. Each household works, consumes, and invests its savings in regular deposits.

The representative household maximizes:

$$E_t \sum_{t=0}^{\infty} \beta^t \left[\frac{1}{1-\gamma} C_t^{1-\gamma} - \frac{a_n}{1+\gamma_n} N_t^{1+\gamma_n} \right]. \tag{4}$$

Subject to the budget constraint:

$$C_t + D_{t+1} = W_t N_t + (1 + r_t) D_t + \Pi_t. \tag{5}$$

Where C_t is consumption; N_t is labor supply; W_t denotes the real wage; D_t are deposits (in real terms) held at commercial banks and $(1 + r_t)$ is the gross real interest rate paid to depositors. Π_t are real dividends payments received from ownership of these financial intermediaries.

4.1.1 Optimality Conditions

Household behavior obeys:

consumption and saving intertemporal allocation:

$$1 = \beta E_t \left\{ \left(\frac{C_t}{C_{t+1}} \right)^{\gamma} (1 + r_{t+1}) \right\}. \tag{6}$$

labor allocation:

$$W_t C_t^{-\gamma} = a_n N_t^{\gamma_n}. \tag{7}$$

4.2 The Entrepreneurial Sector

Entrepreneurs construct capital in each period for use in the subsequent period. Capital is used in combination with labor to produce output. Entrepreneurs are risk neutral. Assuming CRS, Cobb-Douglas technology, the aggregate production function is:

$$Y_t = A_t K_{t-1}^{\alpha} N_t^{1-\alpha}, \tag{8}$$

where Y_t is aggregate output, K_{t-1} is the aggregate amount of capital constructed by entrepreneurs in period $t-1$, N_t is the labor input, and A_t is an exogenous technology shock.

Thus labor demand satisfies:

$$(1 - \alpha) \frac{Y_t}{N_t} = W_t. \tag{9}$$

Demand for New Capital The construction of new capital is determined by the level of investment I_t. Thus, the capital stock obeys:

$$K_t = I_t + (1 - \delta)K_{t-1}, \tag{10}$$

where δ is the depreciation rate.

The gross return to holding one additional unit of capital from t to $t+1$, can be written as:

$$(1 + r^k_{t+1}) = E_t \left\{ \frac{\alpha Y_{t+1}}{K_t} + (1 - \delta) \right\}. \tag{11}$$

Supply for New Capital In equilibrium, the allocation for capital satisfies the following optimality condition:

$$(1 + r^k_{t+1}) = (1 + \Xi_{t+1})\,(1 + r_{t+1}), \tag{12}$$

where the real interest rate, $(1 + r_{t+1})$, is the gross cost of funds absent imperfect competition in the financial system and $(1 + \Xi_{t+1})$ is the gross markup charged by the intermediary bank. I assume that new equity and bond issues are prohibitively expensive, or not available for local firms, so that all investment finance is done with bank credit. I will ignore the presence of the bank multiplier and the existence of reserves. Therefore, the overall amount of credit in the economy must be equal to the overall amount of new household deposits:

$$D_{t+1} = I_t. \tag{13}$$

4.3 The Resource Constraint

The resource constraint for the economy is:

$$Y_t = C_t + I_t. \tag{14}$$

4.4 The Banking System

I assume that the banking system is highly segmented into a large number, n, of sectors or regions (niches).

The size of each niche is the same, and each of them is served by an established bank (incumbent), i, that possesses a local monopoly and therefore finances an equal fraction $\frac{I_t}{n}$ of the total investment. Each incumbent can serve only its own niche because of an implicit collusion agreement that is described later.

This intermediary chooses a net markup for its niche, Ξ_{t+1}, at the beginning of period t. I assume that the cost of serving the niche for each bank i is:

$$\upsilon_i \left(\frac{I_t}{n}\right)^{1-\tau}. \tag{15}$$

The constant υ_i is the cost-efficiency level, and captures any idiosyncratic operational (in)efficiency and information (dis)advantages any bank may have. I assume that υ_i is drawn from a common uniform distribution $U(\upsilon)$ with support on $[0, \lambda]$, at the beginning of the bank operations. υ_i is private information and is unknown to banks outside the niche.

The cost of serving the niche for each bank i depends on the amount of credit financed (the size of the market). In addition, the banking system possesses operational economies of scope and scale over operating costs. Thus, I assume that $0 < \tau < 1$.

In period $t+1$ the bank obtains the following ex-post real profits for carrying the bank contract at period t :

$$\pi_{i,t+1} = (1+\Xi_{t+1})(1+r_{t+1})\left(\frac{I_t}{n}\right) - \left[(1+r_{t+1})\left(\frac{D_{t+1}}{n}\right) + \upsilon_i \left(\frac{I_t}{n}\right)^{1-\tau}\right]. \tag{16}$$

The first term are entrepreneur payments and the term in brackets captures the cost of funds (i.e. payments to depositors) plus operating costs.

Using the fact that $D_{t+1} = I_t$, and that $\Xi_{t+1} r_{t+1} \approx 0$ for the parameter values I consider here, we can express (16) as:

$$\pi_{i,t+1} = \Xi_{t+1} \left(\frac{I_t}{n} \right) - \upsilon_i \left(\frac{I_t}{n} \right)^{1-\tau}. \tag{17}$$

Entry and mergers I assume that entry is possible in this banking system, but that it occurs in successive stages. Entrants in the "banking system" at time t only start competing in the "niche" at time $t+1$, which introduces a one-period time-to-build lag in the model. Right after the entry decision is effectively taken (when the sunk costs are incurred), the entrant is already inside the banking system, but only at the "wholesale level." Hence, during period t, it is able to temporarily serve any of the n niches until it is finally established in one of them in $t+1$. The aim is to capture the idea of entry taking place in the wholesale market first with the ultimate goal of spreading later to the retail segment (niches).[11]

The entry stages are as follows:

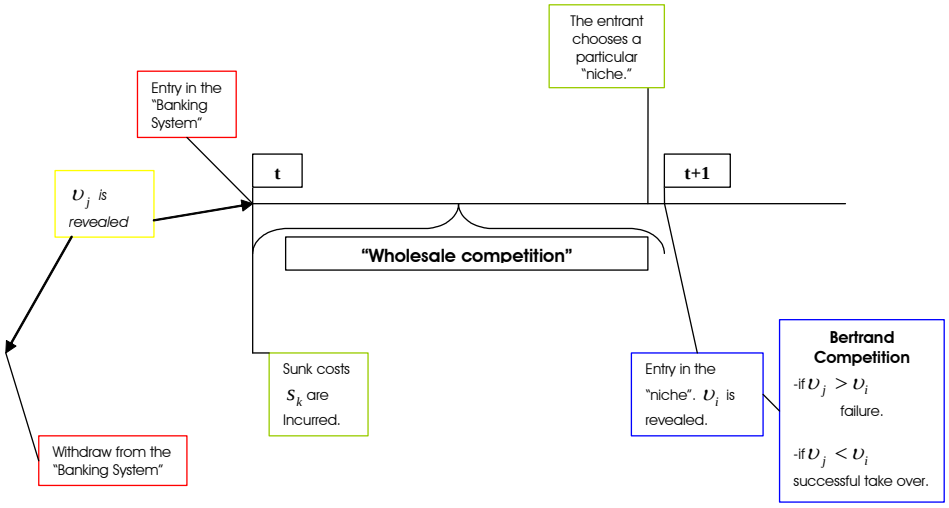
(A) At the beginning of period t, a potential competitor, j, attempts to enter the banking system. At no cost, it draws its cost-efficiency level, υ_j, from the same common uniform distribution $U(\upsilon)$.

(B) After learning its own υ_j, the potential competitor chooses whether to enter the banking system and fight for one of the niches next period or withdraw from the banking system. The closer υ_j is to zero, the more efficient the potential entrant is, and the easier to take over a niche. I will assume that the number of total draws is large enough that at least some potential competitors enter the banking system every period.

(C) To enter the banking system (and eventually fight for one of the niches) an outsider has to incur fixed sunk entry costs, z_t, at the beginning

[11] In addition, we could say that entrants need to incur in a one-period learning process to make their idiosyncratic cost-efficiency level at the regional level effective.

Figure 1: Entry Stages



of period t.[12] z_t is exogenous and measured in real output units. We can also interpret changes in z_t as changes in entry regulations.

(D) In principle, during period t, entrants are able to serve any (or even all) of the n niches at the wholesale level until finally established in one of them. The cost of serving other niches at the wholesale level is:

$$\lambda \left(\frac{I_t}{n} \right)^{1-\tau}, \tag{18}$$

where $\lambda \geq \upsilon_i$ for every i; given the common uniform distribution $U(\upsilon)$ with support on $[0, \lambda]$. As in Petersen and Rajan (1994), I assume that retail banks that are physically closer to their customers have lower costs of transacting with both firms and depositors.

[12] As I said, we can include in them advertisement costs or the construction of a network of branches and ATMs.

(E) For simplicity, I assume that any entrant is able to fight for only one niche (i.e. multi-sectorial entry is not possible). The collusion agreement described later implies that the potential competitor knows the cost-efficiency level distribution of the banking system, $U(\upsilon)$, but cannot infer the particular $\upsilon_i's$ of each incumbent. Hence, entrants are indifferent about the particular niche to fight for. I assume that once inside the banking system they randomly choose which niche to enter at the end of period t.

(F) At the very beginning of period $t+1$, the entrant is inside the niche an is able to learn the incumbent's υ_i. Bertrand competition occurs and the following proposition holds:

PROPOSITION 1 *Under Bertrand competition, only two possible outcomes are possible. If $\upsilon_j > \upsilon_i$, the entrant fails and is forced to merge. If $\upsilon_j < \upsilon_i$ the entrant successfully displaces the incumbent and forces it to merge. The optimal strategy for the loser is to merge immediately and not to compete. The only visible outcome is the possible change of the incumbent at the very beginning of $t+1$.*[13]

Proof. See Appendix B**.**

(G) If successful the new incumbent keeps the niche until it is hit by an exit-inducing shock that occurs with probability $\delta_D \in (0,1)$ in every period. For simplicity, I do not model endogenous exit that is not driven by the afore mentioned Bertrand competition. The "death" shock is independent of the bank's efficiency level. I assume that an entrant immediately fills the empty niche left by every dead bank. Right after drawing an efficiency level, the entrant is able to use the network left by the dead bank (avoiding any sunk costs as well as the time-to-build lag). The number of banks and the frequency of "death" is high enough so that $E(\upsilon_i) = \frac{\lambda}{2}$, and $U(\upsilon)$ nests the cost-efficiency distribution of all incumbents in the financial system.

Implicit Collusion Agreement and Limit Pricing I assume that entrants are liquidity constrained and cannot make losses after incurring sunk costs. In these circumstances, the incumbents' pricing strategy, Ξ_{t+1}, must ensure that none of the new competitors at the wholesale level can obtain

[13] By definition the point likelihood of $\upsilon_j = \upsilon_i$ is null.

any expected positive profits if they decide to offer a net markup below Ξ_{t+1} and serve the niche.[14] That is:

$$\Xi_{t+1}\left(\frac{I_t}{n}\right) \leq \lambda\left(\frac{I_t}{n}\right)^{1-\tau}. \tag{19}$$

Notice, however, that low cost-efficiency incumbents have the incentive to "signal" their idiosyncratic efficiency to new entrants by offering a markup below the level that makes (19) hold as an equality (hereafter, the binding limit). From (17), entrants in the banking system know that only more efficient incumbents can offer a markup, Ξ_{t+1}, well below $\lambda\left(\frac{I_t}{n}\right)^{-\tau}$and still make profits. Therefore, these incumbents have incentives to offer markups levels somewhat below the binding limit in (19) to influence and redirect entrants' decisions toward less-efficient niches. The higher is the amount of entry in the banking system, the higher the incentives to protect the niche by lowering current markups and profits. In this scenario, incumbents would "compete" to deter entry in their own niches. Instead, I assume that there exists an implicit collusion agreement among the incumbents that enforces the secrecy of the idiosyncratic cost-efficiency levels.

I assume that the implicit collusion agreement must necessarily satisfy all the incumbents to be possible.[15] Consequently, a cartel markup below the binding level in (19) does not work. The uniform distribution with support on $[0, \lambda]$, and the assumption that n is very large, implies that such cartel markup level can result in losses for members with cost-efficiency levels in the neighborhood of λ. The negative profits force defections from the agreement; defections that actually reveal the high cost-efficiency level of those defectors. Therefore, the arrangement must consist of markup equal to the binding level in (19):

$$\Xi_{t+1}\left(\frac{I_t}{n}\right) = \lambda\left(\frac{I_t}{n}\right)^{1-\tau} \tag{20}$$

[14] By assumption, the customers remain loyal to the local incumbent bank if the level of the markup offered is the same.

[15] I assume that a single defector can transform the tacit agreement into a an explicit one. As in Rotemberg and Woodford (1992), I assume that such scenario carries incommensurable legal sanctions for the members of the cartel.

If some of the banks attempt to charge a markup below the binding limit, one of the members of the cartel immediately serves such niche at the wholesale level. The punishment consists of establishing a price just below the one chosen by the defector, $\Xi_{t+1}^{def} - \varepsilon$ (such that ε is negligible in size). The resulting negative profits for serving the niche under this condition are equally distributed among the members of the cartel. That is,

$$\frac{\left(\Xi_{t+1}^{def} - \varepsilon\right)(I_t/n) - \lambda(I_t/n)^{1-\tau}}{n-1} < 0. \tag{21}$$

I assume that, in principle, such punishment would take place only if there is a single monopolistic bank serving the niche (so that Proposition 1 holds). In other words, the cartel allows Bertrand competition to occur inside the niche to guarantee a monopolistic structure in which the number of banks in the banking system never exceeds n (one bank per niche). Finally, I assume that the amount of entry and the exogenous exit inducing shock (positively associated with the discount factor) is high enough so that incumbents are better off when committing to the collusive level in (20).

Therefore, the pricing decision is the same in all niches. Since all the niches are of the same size, we can interpret this relationship as the pricing decision taken by the representative bank of this economy. Hence, for every period t, expected profits for each incumbent i are:

$$\pi_{i,t+1} = (\lambda - \upsilon_i)\left(\frac{I_t}{n}\right)^{1-\tau} > 0. \tag{22}$$

Equations (20) and (22) can be interpreted as follows: The greater the aggregate investment, the bigger the size of all niches, and the higher the competitive pressure of the new entrants. In turn, this forces the incumbent to offer lower markups. These countercyclical markups constitute the bank-supply channel that propagates and amplifies shocks to the economy.

Entry decision Banks are forward looking and correctly anticipate their expected stream of profits. After drawing a υ_j, a potential entrant

decides to enter the banking system only if the expected post-entry present discounted net value of the expected stream of profits $\{\pi_{j,t}\}_{t=1}^{\infty}$ is positive:

$$V_{j,t} = \left\{ E_t \sum_{s=t}^{\infty} \left[\beta\left(1-\delta_D\right)\right]^{s-t} \left(\frac{C_{s+1}}{C_s}\right)^{-\gamma} \pi_{j,s+1} \right\} \left(1 - \frac{\upsilon_j}{\lambda}\right) - z_s > 0. \quad (23)$$

Banks discount future profits using the household's stochastic discount factor, adjusted for the probability of survival. The pre-entry probability of "defeating" the incumbent and taking-over the niche is $1 - \frac{\upsilon_j}{\lambda} = \Pr(\upsilon_j < E(\upsilon_i))$. Equations (23) and (22) imply that entry is procyclical (i.e. entry increases when the amount of credit, purchase of new capital and the economic activity are high). The larger the discount factor and the probability of the exit-inducing shock, the stronger the procyclicality.

Entry is affected by market regulation that alters the value of z_t.[16] Equation (23) implies that the higher is z_t, the lower the resulting entry threshold value of υ_j, and thus the lower the amount of entry in the banking system (and vice versa). But then, the higher is z_t, the more likely entries are successful when fighting for the niche. These results are in line with the empirical evidence that entry exerts a sizable impact in small, underdeveloped, and regulated markets.

The government can effectively prohibit entry in the banking system by setting $z_t \to \infty$. In this case, countercyclical limit pricing is not necessary, and incumbents are able to establish a standard collusive agreement.

4.5 Model Parametrization

The only distinctive aspect of the general equilibrium model relative to a benchmark RBC setup is the limit pricing scheme in the financial system, characterized by equations (12) and (20). The former characterizes how imperfect competition in the financial system influences capital demand. The latter describes the limit pricing strategy chosen by the representative incumbent bank. If we restrict the net financial markup Ξ_{t+1} to zero in equation

[16] As in Ghironi and Melitz (2005), changes in sunk entry costs alter the free-entry condition.

(12), we effectively shut off the bank-supply channel and the model reverts to a conventional RBC model.

I set the quarterly discount factor β to 0.99 (which also pins down the steady state quarterly real interest rate depositors receive since $R = \beta^{-1}$). Average hours worked relative to total hours available are set equal to $\frac{1}{3}$. I set the elasticity of intertemporal-substitution, $\frac{1}{\gamma}$, equal to one, and γ_n equal to zero. Following Hansen (1985), I set the standard deviation of the productivity innovations to 0.712. The capital share, α, is 0.36. The quarterly depreciation, δ, is assigned the value of 0.025. From the descriptive statistics for developing countries, I set the quarterly steady-state net financial markup equal to 142 basis points and choose $\tau = 0.70$.

4.6 A Negative Technology Shock

I consider an unanticipated one percent decrease in technology to stress the role of the bank-supply channel deepening a recession. I assume further that the shock obeys a first order auto-correlation process that persists at the rate of 0.95 per quarter. In Figure 2, I plot the response of the eight endogenous variables under both perfect and imperfect competition in the financial system. As I said, the former exactly resembles the basic RBC specification. In this case, there are no financial markups and the natural or wicksellian interest rate depositors and entrepreneurs face are the same.

In the competitive model, a negative technology shock reduces output, factor productivity, and consumption today by more than in future periods. Output and consumption fall today and return later to their original levels. Households want to smooth their consumption and attempt to shift resources away from future periods to the current period. For this reason, we would expect the natural real interest rate to increase.

Investment demand goes down because the technology shock has decreased production. By itself, this pushes down the natural interest rate, offsetting the pressure that comes from households' desire to substitute consumption away from future periods. The net effect of these counteracting pressures is to slightly decrease the natural interest rate by just 7 basis points.

The results change with imperfect competition. The monopolistic intermediary has the possibility of providing credit after charging a markup over

the interest rate paid to depositors. The intermediary banks allow households to substitute consumption away from other periods toward this period by substantially decreasing the interest rate paid on deposits. As a result, consumption does not initially fall as much as in the competitive model. But this relatively higher consumption lowers the marginal utility of income and reduces work effort even more.

A decrease in the labor input negatively affects production and the productivity of capital. This is the cause for an even lower demand for investment relative to the baseline case. Under perfect competition, a resulting lower investment demand and lower interest rates paid to depositors would be reflected in a sharp decrease in the interest rate entrepreneurs face. The fact that investment falls and the financial market shrinks causes the threat of entry to decline, and higher markups are compatible with the limit pricing scheme. The financial markup increases 9.42% (13 bp) on impact. The higher markup does not allow the costs of borrowing for entrepreneurs to fall much, and thus, the optimal capital stock is smaller than in the competitive case and the volatility of all real variables is higher.

4.7 Volatility and Welfare

Macroeconomic Variability and Sensitivity Analysis Quantitative results presented in Tables 6 and 7 confirm that the presence of monopoly power and countercyclical markups in the banking sector ends up increasing the volatility of all real variables relative to the simple RBC model. In the RBC model, the standard deviation of output, consumption and investment is 1.80, 0.52 and 5.74 respectively. With a monopolistic banking system, the corresponding values for the same variables are 2.31, 0.70 and 12.08.

The role of τ is critical for the countercyclical nature of the markups. The larger τ, the larger the banking economies of scale and the higher (lower) the probability of outsiders operating at an efficient scale in a booming (recessionary) economy. In turn, this causes the incumbent to set relatively lower (higher) markups. As expected, in Table 8 we can observe how the volatility of real variables monotically increases when τ increases.

Welfare Results To measure how the welfare of the representative household is affected by the presence of monopolistic power in the banking system, I solve the model using a second-order approximation as in Collard

and Juilliard (2001). Otherwise, conventional linearization can generate approximation errors that may be the cause of possible welfare reversals (see Kim and Kim, 2003 for details). The welfare criterion considered here is based on a second-order Taylor expansion of the representative household's expected utility function (4), around the deterministic steady-state values.

$$\begin{aligned} W_t &= \frac{1}{1-\gamma}\bar{C}^{1-\gamma} - \frac{a_n}{1+\gamma_n}\bar{N}^{1+\gamma_n} + \bar{C}^{1-\gamma}E(\hat{c}_t) \\ &\quad -a_n\bar{N}^{1+\gamma_n}E(\hat{n}_t) - \frac{1}{2}\gamma\bar{C}^{1-\gamma}E(\hat{c}_t^2) - \frac{1}{2}\gamma_n a_n\bar{N}^{1+\gamma_n}E(\hat{n}_t^2). \end{aligned} \tag{24}$$

Where $\bar{C}$ and $\bar{N}$ are the steady-state values of consumption and labor and hats denote percentage deviations from the steady state. In evaluating the welfare criterion, I find that the percent increase in steady-state consumption that would make the household as well off as it would be with perfect competition in the banking system is 10.23%. A monopolistic environment affects welfare of the household through two different channels. Firstly, the financial markup generates a permanent disintermediation between borrowers and entrepreneurs that results in lower steady-state levels of capital accumulation, output, and consumption. Secondly, the countercyclical pattern of such markups increases the volatility of real variables and thus reduces welfare.

5 Conclusions

The contestability of the retail banking sector is restricted by the requirement that entrants must incur large sunk entry costs in highly segmented markets. This implies that the banks must capture a significant fraction of the market right after entering to make the entry profitable. The idea of this paper is that limit pricing strategies aimed at deterring competition in banking retail niches are adopted when incumbents face an entry threat. During recessions the actors in the local banking system are more able to exert their monopolistic power, but boom periods lead to an expansion of the financial system that allows potential entrants to operate at an efficient scale.

Contestable markets force incumbents to lower markups so as to deter entry. In turn, this generates countercyclical financial markups.

Using annual aggregate bank data for a large set of countries for the period 1990-2001, I find that financial markups are strongly countercyclical even after controlling for simultaneity, financial development, banking concentration, operating costs, and inflation. Since the threat of entry is a not a measurable concept, I use foreign penetration as a proxy. I exploit the evidence that foreign bank entry initially takes place in the wholesale market with the intention to expand later to the retail niches. I find that the entry (and not the presence) of foreign banks is the omitted variable that disentangles the cyclicality of the markups in the empirical models .

The modeling of the banking system captures several of the features of the empirical evidence. In the theoretical model, entry occurs at the wholesale level and then spreads to the retail level. The retail market is highly segmented into niches and the more efficient entrants end up taking over current incumbents. Entry is procyclical and more likely to occur in deregulated markets, but is more effective and successful if markets are regulated. Changes in the market structure do not affect the markups. Instead, the markups change because the threat of entry forces incumbents to set rates that deter entry. Finally, economies of scale facilitate entry in boom periods, and vice versa, generating countercyclical markups. At a general equilibrium level, the behavior of this imperfectly competitive financial system generates a bank-supply channel that increases the volatility of real variables, amplifies the business cycle, and reduces welfare. Credit is more expensive during recessions, and firms and households postpone investment and work decisions, thereby deepening the recession.

There are several extensions of the analysis that can be pursued in future work. Not having access to bank-level disaggregated data was a considerable handicap for this study. For instance, it would be interesting to study whether the regional markets that are more concentrated, or have a lower degree of financial development, or are more regulated, have different cyclical patterns. Additionally, the model could be extended to capture the consequences of long-term relationships between banks and customers. Efficiency gains from financial liberalization and market de-segmentation may be offset by some important negative effects not considered in this study. An example is that regional banks are engaged in long-term relationships with small domestic entrepreneurs that otherwise would have no access to the credit market. Entry threats that force low margins can increase the degree of

banking fragility and disrupt these relationships.

Appendix B: Proof of Proposition 1.

Define the break-even level of markups θ_i and θ_j for the incumbent and the entrant. The break-even level is equal to the value of the net margin that provides them zero profits when serving the niche. That is:

$$\theta_i \left(\frac{I_t}{n}\right) - \upsilon_i \left(\frac{I_t}{n}\right)^{1-\tau} = 0, \text{ and} \qquad \theta_j \left(\frac{I_t}{n}\right) - \upsilon_j \left(\frac{I_t}{n}\right)^{1-\tau} = 0. \tag{A.1}$$

Now, let's analyze the case in which $\upsilon_j > \upsilon_i$, and thus $\theta_j > \theta_i$.

Consider for example, $\Xi^i_{t+1} > \Xi^j_{t+1} > \theta_j$. The bank i has no demand and its profits are zero. If bank i charges $\Xi^i_{t+1} = \Xi^j_{t+1} - \varepsilon$ (where ε is positive but nil), it gets the entire niche and has a positive profit $\Xi^j_{t+1} - \varepsilon - \theta_i > 0$.

Therefore bank j cannot be acting in its own interest by charging Ξ^j_{t+1}. Now suppose $\Xi^i_{t+1} = \Xi^j_{t+1} > \theta_j$. In that case they share the niche, and each one serves half of it. But if bank j reduces its price slightly to $\Xi^j_{t+1} - \varepsilon$, it gets all the niche. Nonetheless, bank j will never charge $\Xi^j_{t+1} < \theta_j$, because it would make a negative profit. It follows that bank i can charge $\Xi^i_{t+1} = \theta_j - \varepsilon$ and guarantee for itself all the niche while obtaining a positive profit $\theta_j - \varepsilon - \theta_i > 0$.

Therefore bank j is indifferent between staying or leaving the niche, since will not be able to serve it. If bank i offers bank j a negligible but positive amount of output ε so as to merge, it is in the best interest of bank j to accept it. A symmetric analysis holds when $\upsilon_j < \upsilon_i$.■

References

[1] Angelini, P. and Cetorelli, N. "Bank Competition and Regulatory Reform: The case of the Italian Banking Industry" *Banca d' Italia* WP 99-32. 2000.

[2] Angelini, P. and Cetorelli, N."The Effects of Regulatory Reform on Competition in the Banking Industry." *Journal of Money, Credit and Banking.* 35 (5). 663-684. 2003.

[3] Arellano, M. and Bond, S. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies.* 58. 277-297. 1991.

[4] Bain J. "Barriers to New Competition." Cambridge, Massachussets: Harvard University Press, 1956.

[5] Bank for International Settlements. "The Banking Industry in the Emerging Market Economies: Competition, Consolidation, and Systematic Stability." BIS Papers 4, Basle. 2001.

[6] Baumol, W. "Contestable Markets: An Uprising in the Theory of Industry Structure." *American Economic Review* 72(1). 1-15. 1982.

[7] Berger, A. and Hannan, T. "The Price-Concentration Relationship in Banking." *Review of Economics and Statistics.* 71(2). 291-299. 1989.

[8] Bernanke, B. and Blinder, A. "Credit, Money, and Aggregate Demand." *American Economic Review.* 78. 455-39. 1988.

[9] Blundell, R. and Bond, S. "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models." University College London Discussion Paper. 97-07. 1997.

[10] Bond, S. "Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice." *Portuguese Economic Journal.* 1. 141-162. 2002.

[11] Cerasi, V.; Chizzolini, F. and Ivaldi, J. "Sunk Costs and Competitiveness of European Banks after Deregulation." Toulouse-GREMAQ Series Papers. 97. 473. 1997.

[12] Chevallier, J. and Scharfstein, D. "Capital-Market Imperfections and Countercyclical Markups: Theory and Evidence." *American Economic Review.* 86 (4). 703-725. 1996.

[13] Claessens, S.; Demirguc-Kunt, A. and Huizinga, H. "How does foreign entry affect domestic banking markets?" *Journal of Banking & Finance.* 25. 891-911. 2001.

[14] Collard, F. and Juilliard, M. "Accuracy of stochastic perturbation methods: The Case of Asset Pricing Models." *Journal of Economic Dynamics and Control.* 25. 979-999. 2001.

[15] Demirguc-Kunt, A. and Huzinga, H. "Determinants of Interest Rate Margins and Profitability: Some International Evidence." Working Paper, Development Research Group, World Bank. 1998.

[16] Demirguc-Kunt, A.; Laeven, L.; and Levine R."Regulations, Market Structure, Institutions, and the Cost of Financial Intermediation," *mimeo.* 2003.

[17] Demsetz, H. "Industry Structure, Market Rivalry, and Public Policy." *Journal of Law and Economics.* 16. 1-9. 1973.

[18] Doornik, J.; Arellano, M. and Bond, S. "Panel Data estimation using DPD for Ox," *mimeo.* Nuffield College. 2002.

[19] Edwards, S. and Vegh, C. "Banks and Macroeconomic Disturbances Under Predetermined Exchange Rates." WP 5977. NBER. 1997.

[20] Focarelli, D.; Panetta, F. and Salleo, C. "Why do Banks Merge?" *mimeo.* Federal Reserve Bank of Chicago. May 1999.

[21] Galeotti, M and Schiantarelli, F. "The Cyclicality of Markups in a Model with Adjustment Costs: Econometric Evidence for U.S. industry," *Oxford Bulletin of Economics and Statistics.* 60 (2), 121-142. 1998.

[22] Gertler, M. and Hubbard, R. "Financial Factors in Business Fluctuations," *mimeo.* Columbia. Graduate School of Business. 1988.

[23] Ghironi, F. and Melitz, M. "International Trade and Macroeconomic Dynamics with Heterogeneous firms." *Quarterly Journal of Economics.* CXX. 865-915. 2005.

[24] Gilligan, T.; Smirlock, M. and Marshall, W. "Scale and Scope Economies in the Multiproduct Banking Firm." *Journal of Monetary Economics.* 13. 393-405. 1984.

[25] Grady, D. and Kyle, R. "The Simultaneity of Bank Decision-Making, Market Structure and Bank Performance." *Journal of Finance.* 34. 1-18. 1979.

[26] Hannan, T. and Berger, A."The Rigidity of Prices: Evidence from the Banking Industry." *American Economic Review.* 81. 938-945. 1991.

[27] Hansen, G. "Indivisible Labor and the Business Cycle." *Journal of Monetary Economics.* 16. 309-327. 1985.

[28] Jackson, W. "The Price-Concentration Relationship in Banking: A Comment." *Review of Economics and Statistics.* 74. 371-376. 1992.

[29] Keynes, J. M. "Relative Movements of Real Wages and Output." *Economic Journal.* 24. 34-51. 1939.

[30] Kim, J. and Kim, S. H. "Spurious welfare reversals in international business cycle models." *Journal of International Economics.* 60(2). 471-500. 2003.

[31] Levine, R.; Loayza, N. and Beck, T. "Financial Intermediation and Growth: Causality and Causes." *Journal of Monetary Economics.* 46. 31-77. 2000.

[32] Neumark, D. and Sharpe, S. "Market Structure and the Nature of Price Rigidity: Evidence from the Market for Consumer Deposit." *Quarterly Journal of Economics.* 107(2). 656-680. 1992.

[33] Olivero, M. "Trade, Non-Competitive Banking and the International Transmission of Business Cycle," *mimeo,* Duke University. 2004.

[34] Petersen, M. and Rajan, R. "The Benefits of Lending Relationships: Evidence from Small Business Data." *Journal of Finance.* 49. 3-37. 1994.

[35] Pigou, A. C. *Industrial Fluctuations.* London: Macmillan and Co. 1927.

[36] Rhoades, S. "Structure-Performance Studies in Banking: A Summary and Evaluation." Staff Economic Studies No 92, Federal Reserve Board. 1977.

[37] Ribon, S. and Yosha, O. "Financial Liberalization and Competition in Banking: An Empirical Investigation." Tel Aviv University, Working Paper. 23-99. 1999.

[38] Rotemberg, J. and Saloner, G. "A Supergame Theoretic Model of Business Cycles and Price Wars During Booms." *American Economic Review.* 76 (3). 390-407. 1986.

[39] Rotemberg, J. and Woodford, M. "Oligopolistic Pricing and the Effects of Aggregate Demand on Economic Activity." *Journal of Political Economy.* 100. 1153-1207. 1992.

[40] Saunders, A. and Shumacher, L. "The Determinants of Bank Interest Rate Margins: an International Study." *Journal of International Money and Finance.* 19. 813-832. 2000.

[41] Shaffer, S. "A Test of Competition in Canadian Banking." *Journal of Money, Credit and Banking.* 25. 49-61. 1993.

[42] Smirlock, M. "Evidence on the (Non) Relationship between Concentration and Profitability in Banking." *Journal of Money Credit and Banking.* 17. 69-83. 1985.

[43] Spiller, P. and Favaro, E. "The Effects of Entry Regulation on Oligopolistic Interactions: The Uruguayan Banking Sector." *RAND Journal of Economics.* 9. 305-327. 1984.

[44] Vives, X. "Regulatory Reform in European Banking". *European Economic Review.* 35. 505-515. 1991.

STATISTICAL APPENDIX

Table 1: Descriptive Statistics –Mean (M) and Standard Deviation (SD)-

	NET INTEREST MARGINS	PRIVATE CREDIT	CONCENTRATION	No. of Foreign Banks	Obs.
Developing countries	M 0.0571 SD 0.0351	M 0.2532 SD 0.2125	M 0.6645 SD 0.2280	M 0.3226 SD 0.1839	91
Developed countries	M 0.0268 SD 0.0119	M 0.7653 SD 0.3331	M 0.5751 SD 0.2206	M 0.2769 SD 0.2144	33

Table 2: Time Series Properties.
Dependent Variable: Net Interest Margins (NIM_t)

	OLS LEVELS	WITHIN GROUPS	GMM DIF	GMM SYS
NIM_{t-1}	0.835 (0.000)	0.423 (0.000)	0.296 (0.021)	0.728 (0.000)
m1			-2.240 (0.025)	-3.341 (0.001)
m2			-0.675 (0.500)	0.4228 (0.672)
Sargan			0.178	0.123

Sample: 124 Countries (1990-2000).
-Year Dummies included in all models.
-m1 and m2 are test for first and second order serial correlation for first-differenced residuals, asymptotically N (0, 1). They are reported from the first-step estimations.
-The Sargan test for over-identifying restrictions for the GMM estimators is asymptotically χ^2. It is reported from the two-step estimations.
-P- values are reported in parentheses.

Table 3: Basic Model

Dependent Variable: Net Interest Margins (NIM_t)

	GMM DIF	GMM SYS	GMM DIF	GMM SYS
NIM_{t-1}	-0.068 (0.680)	0.678 (0.000)	0.017 (0.903)	0.576 (0.000)
$GROWTH_t$	-0.212 (0.004)	-0.108 (0.009)	-0.145 (0.005)	-0.066 (0.088)
$PRIV.CRED.(avg)_t$	-----	-----	-0.040 (0.218)	-0.027 (0.000)
.m1	-2.795 (0.005)	-3.189 (0.001)	-2.633 (0.008)	-3.013 (0.003)
.m2	-1.591 (0.112)	0.219 (0.826)	-1.522 (0.128)	0.5301 (0.596)
Sargan	0.453	0.071	0.077	0.194

Sample: 109 Countries (1990-2000). For further information, see notes to Table 2.

Table 4: Sensitivity Analysis -GMM System Estimator-.

Dependent Variable: Net Interest Margins (NIM_t)

NIM_{t-1}	0.548 (0.000)	0.430 (0.000)	0.512 (0.000)
$GROWTH_t$	-0.101 (0.011)	-0.053 (0.069)	-0.064 (0.068)
$PRIV.CRED.(avg)_t$	-0.032 (0.000)	-0.017 (0.001)	-0.035 (0.000)
$CONCENTR._t$	-0.018 (0.085)	-0.012 (0.139)	-0.008 (0.437)
$OVERCOSTS_t$	-----	0.455 (0.002)	-----
$INFLATION_t$	-----	-----	-0.008 (0.486)
$REALRATE_t$	-----	-----	0.0002 (0.345)
.m1 .m2 Sargan	-3.167 (0.002) 0.522 (0.602) 0.805	-3.009 (0.003) 0.8251 (0.409) 0.949	-2.703 (0.007) 1.213 (0.225) 1.000

Table 5: Entry and the counter-cyclicality of the net interest margins -GMM System Estimator-.

Dependent Variable: Net Interest Margins (NIM_t)

NIM_{t-1}	0.678 (0.000)	0.548 (0.000)	0.518 (0.000)
$GROWTH_t$	-0.108 (0.009)	-0.101 (0.011)	-0.069 (0.120)
$PRIV.CRED.(avg)_t$	-----	-0.032 (0.000)	-0.037 (0.001)
$CONCENTR._t$	-----	-0.018 (0.085)	-----
$Foreign\ Banks_{t-1}$	-----	-----	-0.014 (0.412)
$Foreign\ Banks_{t-2}$	-----	-----	0.015 (0.263)
.m1 .m2 Sargan	-3.189 (0.001) 0.219 (0.826) 0.071	-3.167 (0.002) 0.522 (0.602) 0.805	-2.509 (0.012) 1.374 (0.170) 0.832

Sample: 109 Countries (1990-2000). For further information, see notes to Table 2.

Table 6: Standard Deviation

VARIABLE	RBC	ICFM
Output	1.80	2.31
Consumption	0.52	0.70
Investment	5.74	12.08
Capital	0.49	0.94
Employment	1.37	2.19

Notes: Theoretical second moments (as percentage deviations from steady-state values) are reported. RBC refers to the standard RBC model and ICFM to the monopolistic financial market setup. The method used was the frequency domain technique described in Uhlig (1999). The series are H-P filtered with a smoothness parameter of 1600 so that only the cyclical components remain.

Table 7: Relative Standard Deviation

VARIABLE	RBC	ICFM
Output	1.00	1.00
Consumption	0.29	0.30
Investment	3.19	5.23
Capital	0.28	0.41
Employment	0.76	0.95

Notes: Standard deviations relative to output. For further information, see notes to Table 6.

Table 8: Sensitivity Analysis

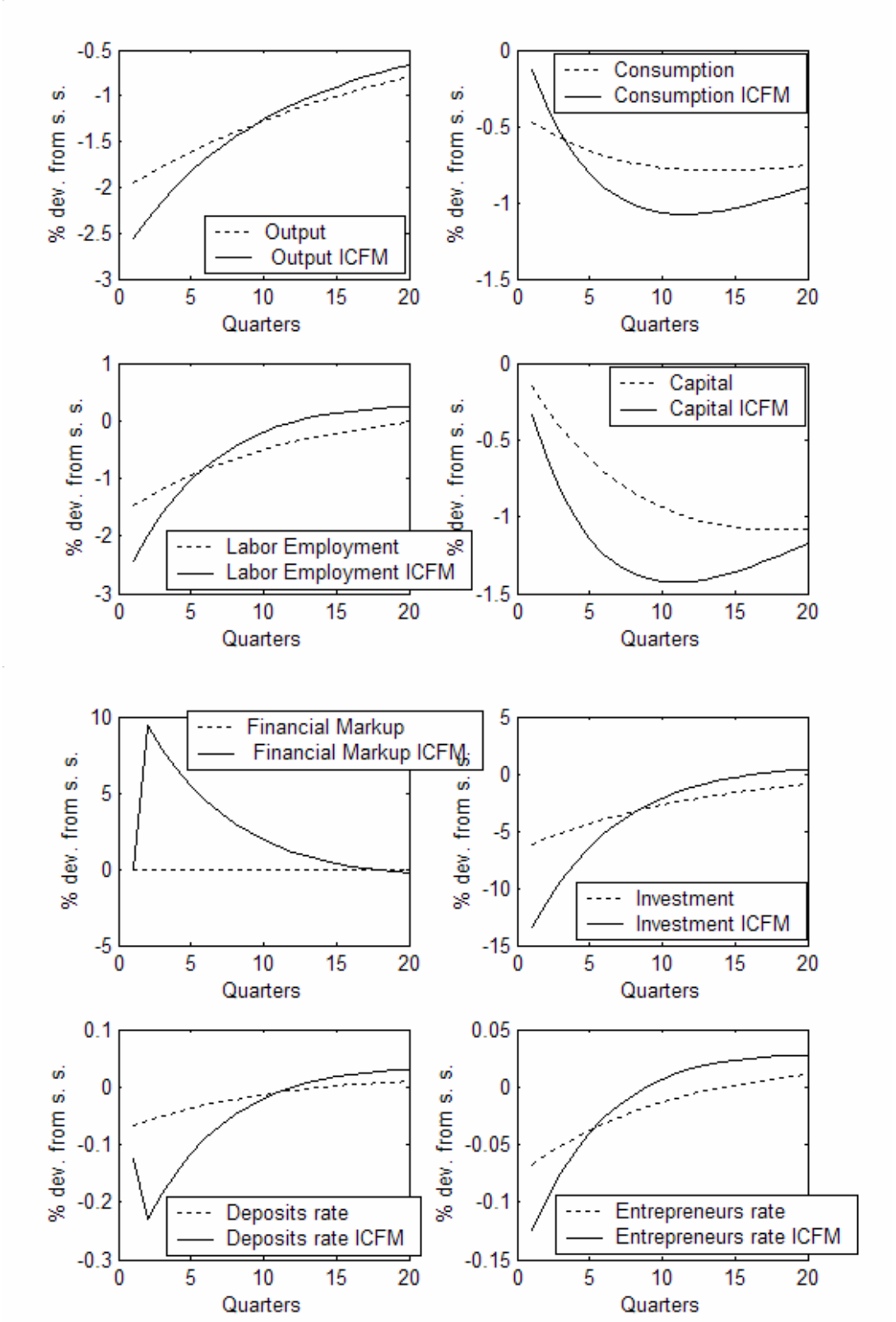
VARIABLE	$\tau = 0.20$	$\tau = 0.40$	$\tau = 0.60$	$\tau = 0.80$
Output	1.84	2.02	2.21	2.41
Consumption	0.60	0.62	0.66	0.74
Investment	8.24	9.67	11.26	12.89
Capital	0.68	0.78	0.89	0.98
Employment	1.43	1.71	2.03	2.35

Notes: Volatilities for different parameter values of τ . For further information, see notes to Table 6.

Variable Definitions

NIM: Net interest income minus interest over total assets (after subtracting defaulted loans).
GROWTH: Annual growth rate of real GDP.
PRIV. CREDIT.: Private Credit by deposit money banks to GDP calculated using the following deflation method: $\{(0.5)*[F_t / P_{et} + F_{t-1} / P_{et-1}]\}/[GDP_t / a_t]$
CONCENTRATION: Assets of the three largest banks as a share of the assets of all the commercial banks in the system.
OVERCOSTS: Accounting value of a bank's overhead costs as a share of its total assets.
INFLATION: Annual inflation from the GDP deflator.
REALRATE: Real interest rate.
Foreign Banks: Number of foreign banks to total number of banks. A bank is defined to be a foreign bank if it has at least fifty percent foreign ownership.
GDP: Real GDP per capita.

Figure 2



Percentage point response of the Monopolistic Financial Market (solid line), and RBC (dashed line) models' to an unanticipated one percent decrease in technology.