



INTERNATIONAL JOURNAL OF CENTRAL BANKING

An Empirical Evaluation of Structural
Credit-Risk Models

Nikola A. Tarashev

Inflation Targets as Focal Points

Maria Demertzis and Nicola Viegi

Transaction Pricing and the Adoption of Electronic Payments:
A Cross-Country Comparison

Wilko Bolt, David Humphrey, and Roland Uittenbogaard

Using Securities Market Information for Bank
Supervisory Monitoring

John Krainer and Jose A. Lopez

Capital Regulation and Banks' Financial Decisions

Haibin Zhu

Preference Heterogeneity in Monetary Policy Committees

Alessandro Riboni and Francisco J. Ruge-Murcia



An Empirical Evaluation of Structural Credit-Risk Models <i>Nikola A. Tarashev</i>	1
Inflation Targets as Focal Points <i>Maria Demertzis and Nicola Viegi</i>	55
Transaction Pricing and the Adoption of Electronic Payments: A Cross-Country Comparison <i>Wilko Bolt, David Humphrey, and Roland Uittenbogaard</i>	89
Using Securities Market Information for Bank Supervisory Monitoring <i>John Krainer and Jose A. Lopez</i>	125
Capital Regulation and Banks' Financial Decisions <i>Haibin Zhu</i>	165
Preference Heterogeneity in Monetary Policy Committees <i>Alessandro Riboni and Francisco J. Ruge-Murcia</i>	213

The contents of this journal, together with additional materials provided by article authors, are available without charge at www.ijcb.org.

Copyright © 2008 by the Association of the International Journal of Central Banking.
All rights reserved. Brief excerpts may be reproduced or translated provided the source
is cited. Consult www.ijcb.org for further information.

The *International Journal of Central Banking* is published quarterly
(ISSN: 1815-4654). Online access to the publication is available free of charge
at **www.ijcb.org**. Individual print subscriptions are available at an annual rate of
\$100 (USD).

Print subscription orders may be placed online at www.ijcb.org, by phone
(+49 69 1344 7623), via fax (+49 69 1344 8553), or by e-mail (editor@ijcb.org).

Renewals, claims, address changes, and requests for permission to reprint material
from this journal should be addressed to:

International Journal of Central Banking
DG Research
European Central Bank
Postfach 16 03 19
D-60066 Frankfurt
Germany
Phone: +49 69 1344 7623
Fax: +49 69 1344 8553
E-mail: editor@ijcb.org

The views expressed in this journal do not necessarily represent the views of the
Association of the International Journal of Central Banking or any of its members.

ISSN: 1815-4654

International Journal of Central Banking

Board of Directors

Chairman

Charles Bean, *Bank of England*

Board Members

Abdulrahman Al-Hamidy, *Saudi Arabian Monetary Agency*

Jan Marc Berk, *The Nederlandsche Bank*

Afonso Bevilaqua, *Central Bank of Brazil*

Claudio Borio, *Bank for International Settlements*

Mohammed Tahar Bouhouche, *Bank of Algeria*

Malcolm Edey, *Reserve Bank of Australia*

Khor Hoe Ee, *Monetary Authority of Singapore*

Manuel Ramos Francia, *Bank of Mexico*

Hans Genberg, *Hong Kong Monetary Authority*

Heinz Herrmann, *Deutsche Bundesbank*

Nobuo Inaba, *Bank of Japan*

Tor Jacobson, *Sveriges Riksbank*

He Jianxiong, *People's Bank of China*

Juan F. Jimeno, *Bank of Spain*

Ali Hakan Kara, *Central Bank of Turkey*

Ulrich Kohli, *Swiss National Bank*

Donald Kohn, *Federal Reserve Board*

Ana Leal, *Bank of Portugal*

Vyacheslav Morgunov, *Central Bank of Russian Federation*

John Murray, *Bank of Canada*

Tom O'Connell, *Central Bank of Ireland*

Fabio Panetta, *Bank of Italy*

Jan Qvigstad, *Norges Bank*

Lucrezia Reichlin, *European Central Bank*

Ivan Ribnikar, *Bank of Slovenia*

Arnor Sighvatsson, *Central Bank of Iceland*

Jan Smets, *National Bank of Belgium*

Marc-Olivier Strauss-Kahn, *Bank of France*

Juha Tarkka, *Bank of Finland*

George Tavlas, *Bank of Greece*

Dobieslaw Tymoczko, *National Bank of Poland*

Editorial Board

Managing Editor

Frank Smets

European Central Bank

Co-editors

Hyun Shin
Princeton University

Andrew Levin
Federal Reserve Board

Kazuo Ueda
The University of Tokyo

Carl E. Walsh
University of California

Associate Editors

Franklin Allen
The Wharton School of the
University of Pennsylvania

Michael D. Bordo
Rutgers University

Guy Debelle
Reserve Bank of Australia

Michael B. Devereux
University of British Columbia

Douglas W. Diamond
University of Chicago Graduate
School of Business

Michael Dotsey
Federal Reserve Bank of
Philadelphia

Douglas Gale
New York University

Jordi Galí
Centre de Recerca en Economia
Internacional (CREI)

Marvin Goodfriend
Carnegie Mellon University

Charles A.E. Goodhart
London School of Economics

Michael B. Gordy
Federal Reserve Board

Luigi Guiso
European University Institute

Andrew G. Haldane
Bank of England

Takatoshi Ito
University of Tokyo

Philip Lane
Trinity College Dublin

Francesco Lippi
University of Sassari

Carmen M. Reinhart
University of Maryland

Rafael Repullo
Centro de Estudios Monetarios
y Financieros (CEMFI)

Eli M. Remolona
Bank for International
Settlements

Hélène Rey
London Business School

Andrew K. Rose
University of California,
Berkeley

Klaus Schmidt-Hebbel
Central Bank of Chile

Lars E.O. Svensson
Sveriges Riksbank

Jürgen von Hagen
University of Bonn

Tsutomu Watanabe
Hitotsubashi University

An Empirical Evaluation of Structural Credit-Risk Models*

Nikola A. Tarashev
Bank for International Settlements

This paper evaluates the capacity of five structural credit-risk models to forecast default rates. In contrast to previous studies with similar objectives, the paper employs firm-level data and finds that model-based forecasts of default rates tend to be unbiased and to deliver point-in-time errors that are small in both statistical and economic terms. In addition, in- and out-of-sample regression analysis reveals that the models account for a significant portion of the variability of credit risk over time but fail to fully reflect its dependence on macroeconomic cycles.

JEL Codes: G33, E44, G28, C13.

1. Introduction

Accurate measures of credit risk are a key prerequisite for sound risk management and for meaningful supervisory evaluation of the vulnerability of lender institutions. In an appreciation of this, the new capital adequacy framework (known as Basel II) will require banks to validate the appropriateness of their credit-risk models, some of which are likely to incorporate—explicitly or implicitly—constructs developed in the academic literature.

*I am indebted to Claudio Borio and Kostas Tsatsaronis for their encouragement and for several insightful discussions on the topic. An anonymous referee, Jeff Amato, Ingo Fender, Michael Gordy, Hirotaka Hideshima, Ricardo Schechtman, Richard Vasicek, Haibin Zhu, and seminar participants at the Bank for International Settlements (BIS), Banque de France, the Vienna Workshop of the Basel Committee's Research Task Force, and Karlsruhe University provided useful comments on earlier drafts of this paper. I also thank Dimitrios Karampatos for valuable help with the data and Melanie Sykes for putting the Word document together. The views expressed in this paper are my own and do not necessarily reflect those of the BIS. Author contact: Bank for International Settlements, Centralbahnplatz 2, CH-4052 Basel, Switzerland; e-mail: nikola.tarashev@bis.org.

The academic literature has produced a variety of theoretical models of credit risk. This paper examines the family of *structural* models, which consider the stochastic process of a corporate borrower's assets and postulate that a default occurs when these assets fall below a threshold value. The structural models can be divided into an "endogenous-default" group and an "exogenous-default" group. Models in the former group let borrowers choose strategically the timing of default. In contrast, models in the latter group impose an ad hoc default threshold but develop richer stochastic environments in order to capture empirical regularities of credit markets that are unaccounted for by endogenous-default models.

The paper analyzes the extent to which model-based default-rate expectations—defined as the cross-sectional averages of model-based probabilities of default (PDs)¹—account for the level and intertemporal evolution of actual default rates in three (BBB, BB, and B) rating classes. The analysis considers five structural models, which are calibrated to firm-specific borrower characteristics that are allowed to change over time. Two of the models incorporate endogenous defaults and are developed in Leland and Toft (1996) (henceforth, LT) and Anderson, Sundaresan, and Tychon (1996) (AST). The other three are exogenous-default models, developed in Longstaff and Schwartz (1995) (LS), Collin-Dufresne and Goldstein (2001) (CDG), and Huang and Huang (2003) (HH).²

One of the main conclusions of this paper is that the structural credit-risk models tend to account for the level of default rates. First, model-based expectations of default rates are generally unbiased at both the one- and five-year horizons. Second, the point-in-time errors in these expectations tend to be small in statistical terms. Third, these errors also appear to be of little economic importance, as they tend to translate into small discrepancies between (i) capital buffers derived under the internal-ratings-based (IRB) approach of Basel II on the basis of model-implied expected default rates and (ii) capital

¹The paper abstracts from other components of credit risk, such as loss given default and exposure at default.

²It is of natural interest to compare the performance of these academic models with the performance of commercial models, such as the one underpinning the popular expected default frequencies of Moody's KMV. Data limitations, however, do not allow for including such comparisons in this paper.

buffers derived under the same approach but incorporating actual default rates as (ex post) measures of credit risk.³

These general findings mask important differences across models, which are due mainly to differences in the way that the models incorporate two borrower characteristics: the default threshold and the dead-weight cost at default. The endogenous-default LT and AST models allow the two characteristics to be extracted simultaneously from the data. Under the LT model, such an extraction leads to expected default rates that match the level of actual default rates closely and consistently across horizons and rating classes. Under the AST model, however, the default thresholds appear to be overly sensitive to changes in asset volatility, which leads to underpredictions of one-year default rates.

In turn, a calibration scheme that is applicable to the exogenous-default LS, HH, and CDG models needs to incorporate an ad hoc parameter value. In the specific application, such a value is assigned to the dead-weight cost, which is set equal to the average of its counterpart in the LT model and then is used for calculating exogenous-default thresholds. Under the LS and HH models, this leads to generally accurate default-rate expectations. By contrast, the CDG model, which forces the default threshold to gravitate too strongly toward a level that guarantees solvency, consistently underestimates default rates.

Owing to their calibration to borrower-specific parameters, all five models deliver expectation errors that are orders of magnitude smaller than those reported by Leland (2004). That paper calibrates the LT and LS models to the *representative* borrower and concludes that they substantially underestimate default rates at short—e.g., one-year—horizons. The difference between the conclusion of Leland (2004) and the finding of this paper is due to the strongly nonlinear relationship between inputs to the models and model-implied PDs. The nonlinearities give rise to the so-called Jensen inequality effect, whereby the average PD across borrowers—i.e., the model-based expectation of the default rate—is larger than the PD of the average—i.e., representative—borrower.

³See Basel Committee on Banking Supervision (2004).

The Jensen inequality effect is also the reason why the models underpredict credit risk when their calibration differentiates borrowers only on the basis of time averages of borrower-specific characteristics. Such calibration decreases the effective cross-sectional heterogeneity of the borrowers, which dampens the Jensen inequality effect and, thus, excessively depresses model-based expectations of default rates.

In addition, this paper conducts regression analysis in order to study the extent to which structural credit-risk models account for the evolution of default rates over time. All regressions include model-based expected default rates as explanatory variables. Furthermore, the analysis incorporates macroeconomic control variables pointed out by Estrella and Hardouvelis (1991) and Smets and Tsatsaronis (1997), who identify predictors of economic activity, and by Borio and Lowe (2002), who identify predictors of banking-system distress.

The results of the regression analysis, which are confirmed in out-of-sample exercises, indicate that structural credit-risk models account to an extent for the intertemporal evolution of credit risk but fail to reflect fully its dependence on the business and credit cycles. The models perform best when calibrated to time-varying, as opposed to average, borrower-specific parameters. In addition, simultaneously using the expected default rates implied by several models improves on the fit of single-model forecasts. Exploiting these regularities reveals “models-only” regressions that attain adjusted R^2 of 22 percent, 66 percent, and 63 percent in the BBB, BB, and B rating classes, respectively, which is in line with a finding of van Deventer, Li, and Wang (2006). At the same time, PDs implied by the structural models miss important information regarding the impact of real and financial activity on credit risk. This is seen in the finding that the macroeconomic controls add to and, for the BBB rating class, even substitute for the information content of the models.

The rest of the paper is organized as follows. After a review of the related literature in section 2, section 3 presents the five models and highlights their key assumptions. Then, section 4 describes the data, and section 5 outlines and discusses how the data are employed for the calibration of the models. Section 6 presents and discusses model-based predictions of one- and five-year default rates.

Considering mainly the one-year horizon, sections 7 and 8 evaluate, respectively, the statistical and economic significance of errors in the models' default-rate forecasts. Finally, section 9 analyzes the significance of one-year model-based PDs as in-sample explanatory variables and out-of-sample forecasters of default rates.

2. Related Literature

The analysis in this paper—which is in the spirit of the aforementioned Leland (2004) and van Deventer, Li, and Wang (2006)—contributes to a large and growing literature on the performance of structural credit-risk models.

A sizable portion of this literature consists of studies conducted by Moody's KMV staff. These studies perform two general exercises that analyze the capacity of the structural model underpinning Moody's KMV PD estimates to (i) account for the level of observed default rates and (ii) accurately differentiate credit risk across firms.⁴ The first exercise is paralleled in this paper in the context of academic models. The second exercise rests on a large proprietary data set of default events and is thus beyond the scope of this paper. Importantly, these exercises reveal that the performance of the Moody's KMV model depends crucially on its flexible empirical implementation. Consequently, it is useful to keep in mind that the default-rate forecasts examined in this paper are based on a *direct* implementation of academic models.

In addition, a number of articles have examined the degree to which (variants of) the Black-Scholes-Merton (BSM) model can help predict credit risk. Hillegeist et al. (2004), e.g., incorporate BSM PDs in a hazard model and find that they contain significantly more information on credit risk than popular accounting-based measures. In related exercises, Bharath and Shumway (2004) and Duffie, Saita, and Wang (2007) reach the conclusion that a volatility-adjusted leverage—which is a key parameter in the BSM model and is also known as distance to default (DD)—plays an important role in, respectively, a sufficient statistic for the PD and forecasts of the PD term structure. Duffie, Saita, and Wang (2007) also find that the

⁴See, e.g., Stein (2000), Kealhofer and Kurbat (2001), and Bohn, Arora, and Korablev (2005).

information content of the DD variable is usefully complemented by macroeconomic covariates, a result echoed by the regression analysis below.

Furthermore, several articles have examined the models considered in this paper in order to evaluate their capacity to accurately price credit risk. Huang and Huang (2003), e.g., calibrate the models to observed default rates and then analyze the implied credit-risk premiums. For their part, Eom, Helwege, and Huang (2004) study individual bonds in order to compare their credit spreads to the predictions of structural credit-risk models. While these articles examine theoretical pricing implications, which are based on preference-weighted probabilities that should reflect bond holders' risk appetite, the analysis herein focuses on actual (or physical) PDs.

Finally, given that the considered models pertain to single-name credit risk, the analysis in this paper only touches on issues related to the interdependence of default events across borrowers. For recent advances in the treatment of such issues, which play a key role in estimates of portfolio credit risk, see, e.g., Basel Committee on Banking Supervision (2006) and the references therein, as well as Das et al. (2007).

3. The Models

Constructed in the spirit of the Merton (1974) framework, the five structural credit-risk models examined in this paper focus on an individual borrower that defaults when its assets, V , fall below a particular threshold, V^* . All models assume that the asset value evolves as follows:

$$dV_t/V_t = (r + \lambda - \delta)dt + \sigma dW, \quad (1)$$

where r denotes the risk-free rate, λ is the asset risk premium, δ is the asset payout ratio (reflecting, e.g., dividend and coupon payments), W is a standard Wiener process, and σ is the instantaneous asset volatility.

The PD over a particular horizon equals the probability that the first passage of V below V^* occurs over this horizon. The nature of the default threshold differs across models and makes it natural to divide them into two categories. Those in the first category—i.e., the

HH, LS, and CDG models—adopt an exogenous threshold, whose value has to be determined outside of the model. In contrast, the models in the second category—i.e., the LT and the AST models—derive the decision to default endogenously, as part of the borrower's optimization problem. Thus, the default threshold in these models is a function of borrower characteristics.

Given an initial value of assets, V_0 , all five models imply the following:

$$\frac{\partial PD}{\partial V^*} > 0, \quad \frac{\partial PD}{\partial \sigma} > 0, \quad \frac{\partial PD}{\partial r} < 0, \quad \frac{\partial PD}{\partial \lambda} < 0, \quad \frac{\partial PD}{\partial \delta} > 0. \quad (2)$$

Intuitively, the PD increases in the default threshold, V^* , and in the level of risk, as captured by asset volatility σ . The implications of the remaining three parameters, which determine the drift in the value of assets, are best considered together. Tight credit conditions—caused by contractionary monetary policy that raises the risk-free rate, r , and/or by stronger aversion to risk that raises the risk premium, λ —seem to counterintuitively depress PDs. At the same time, however, tight credit conditions would tend to raise the payout ratio, δ , which increases PDs.

3.1 *Exogenous-Default Models*

Besides leaving the value of the default threshold to the discretion of the empirical researcher, the three exogenous-default models are similar in that they assume that all debt is of infinite maturity. This assumption delivers analytic tractability but makes it impossible to capture the empirical regularity that borrowers are less likely to default over a given horizon if they are to repay the debt principal further in the future.

What differs across exogenous-default models are their assumptions regarding the overall stochastic environment. The remainder of this subsection discusses these assumptions, paying particular attention to their implications for theoretical PDs.

3.1.1 *The Model of Longstaff and Schwartz (1995)*

In order to capture the cooling effect of rises in the interest rate on the macroeconomy, the LS model allows for a stochastic risk-free

rate of return that correlates negatively with shocks to asset returns:

$$dr_t = k_r(\bar{r} - r_t)dt + \sigma_r dW_t^r, \quad (3)$$

where $\bar{r}$ is the long-run risk-free rate of return, k_r reflects the speed of mean reversion, σ_r is the instantaneous volatility of the risk-free rate of return, and $\text{corr}(dW_t^r, dW_t) \equiv \sigma_{rV} < 0$.

In such a setting, a change of r has an ambiguous impact on the PD. By expression (1), a higher interest rate increases the deterministic drift in the value of assets and, ceteris paribus, lowers the PD. At the same time, the assumption $\sigma_{rV} < 0$ implies that a higher r tends to be associated with a negative shock to asset returns, which raises the PD. The relative importance of the latter impact increases when asset volatility, σ , is higher and/or the correlation $\frac{\sigma_{rV}}{\sigma_r\sigma}$ is closer to -1 .

3.1.2 The Model of Collin-Dufresne and Goldstein (2001)

Collin-Dufresne and Goldstein observe that debt issuance increases when the value of assets increases, as a result of which the leverage ratio (i.e., the ratio of debt to assets) exhibits mean reversion. The CDG model accommodates this empirical regularity, which implies that the default threshold, V_t^* , moves in step with the value of assets, V_t . Under the maintained assumption that V_t^* is a constant fraction of debt,

$$d \ln V_t^* = k_l (\ln V_t - \ln V_t^* - \nu) dt, \text{ where } k_l > 0, \nu > 0. \quad (4)$$

The parameters ν and k_l affect theoretical PDs as follows. The restriction $\nu > 0$ implies that the borrower is inherently solvent, as its assets would stay above the default threshold in the absence of shocks. That said, the closer ν is to zero, the closer V_t^* and V_t tend to stay to each other and, thus, the higher the risk of V_t falling below V_t^* , which causes a default. In turn, a higher k_l implies that the ratio $\ln(V_t/V_t^*)$ is more likely to stay close to its long-run value ν . Since $\nu > 0$, an increase in k_l lowers the PD.

3.1.3 The Model of Huang and Huang (2003)

There is empirical evidence that equity risk premiums are negatively correlated with returns on broad equity indices. On the basis

of such evidence, the HH model incorporates a stochastic risk premium that correlates negatively with shocks to the assets of the typical borrower:

$$d\lambda_t = k_\lambda(\bar{\lambda} - \lambda_t)dt + \sigma_\lambda dW_t^\lambda, \quad \text{corr}(dW_t^\lambda, dW_t) \equiv \sigma_{\lambda V} < 0. \quad (5)$$

A higher $\bar{\lambda}$ implies a higher long-run drift in the value of assets, which, ceteris paribus, lowers the PD. The impact of $\bar{\lambda}$ is stronger the larger is the mean-reversion parameter k_λ . In addition, since $\sigma_{\lambda V} < 0$, a negative value of dW_t (which puts upward pressure on the PD) tends to be counteracted by an increase of the drift in the value of assets.

3.2 *Endogenous-Default Models*

The two endogenous-default models, which let the borrower decide when to default, differ in their assumptions regarding the debt contract. While the AST model allows the borrower to renege on and then alter the terms of its contract, no renegotiation is possible in the LT model, in which a defaulting firm is surrendered to creditors. In addition, while the AST framework focuses on perpetual bonds, the LT model assumes continuous issuance of debt that is of constant but finite time to maturity.

3.2.1 *The Default Threshold in the Model of Anderson, Sundaresan, and Tychon (1996)*

At the time of default, creditors in the AST model choose between two options. They can either liquidate the borrowing firm and seize its assets (net of a bankruptcy cost) or accept the terms of a new debt contract. Since a liquidation of the borrowing firm is the worst possible outcome for its equity holders, they propose a postdefault contract that is acceptable to creditors.

In this setup, the default threshold is determined by the borrower's capacity and willingness to renegotiate its debt contract without being liquidated. Thus, a higher (fixed) dead-weight cost, K —incurred by creditors only if they liquidate the borrower—boosts

the borrower's bargaining power and allows it to set a higher default threshold V_{AST}^* . In addition, V_{AST}^* is also raised by an upward shift in the predefault value of debt, which—being equivalent to an upward shift in the debt-service burden—prompts borrowers to negotiate a more advantageous contract earlier. Consequently, excluding K , the impact of the remaining model parameters on V_{AST}^* is a result of their impact on the predefault value of debt. In concrete terms, this value is raised by a higher risk-neutral drift in the asset process (i.e., a higher r^5 or a lower δ); a higher debt principal, P ; a higher coupon rate, c ; a lower asset volatility, σ ; and a lower monitoring cost, m .⁶ In sum,

$$\begin{aligned} \frac{dV_{AST}^*}{dK} &> 0, \frac{dV_{AST}^*}{dr} > 0, \frac{dV_{AST}^*}{d\delta} < 0, \frac{dV_{AST}^*}{dP} > 0, \\ \frac{dV_{AST}^*}{dc} &> 0, \frac{dV_{AST}^*}{d\sigma} < 0, \frac{dV_{AST}^*}{dm} < 0. \end{aligned} \quad (6)$$

3.2.2 *The Default Threshold in the Model of Leland and Toft (1996)*

In the LT model, a borrower forfeits its equity value as soon as it does not fulfill a contracted obligation. Thus, the willingness to service debt increases—i.e., the default threshold V_{LT}^* decreases—with increasing value of equity, which equals the value of the firm net of the value of its debt.⁷ On the one hand, the value of the firm decreases in the default cost (which is a fraction α of assets) but, being assessed over the infinite horizon, is insensitive to the time to maturity, T , of the continuously issued debt contracts. On the other hand, the value of debt contracts decreases in α (but by less than the value of the firm) and in T , a rise in which heightens the risk that the borrower defaults before the contract matures. The upshot

⁵In the empirical exercises reported below, this positive effect of a rise in r on the predefault value of debt is weakened but never reversed by the fact that a rise in r also lowers the present value of coupon payments.

⁶The original AST model does not incorporate monitoring costs. The introduction of such costs generalizes the original formula for V_{AST}^* by replacing the coupon payments cP with $(1 - m)cP$.

⁷In the LT model, the (market) value of the firm equals the asset value plus the value of tax benefits, less the value of bankruptcy costs, over the infinite horizon.

is that the default threshold V_{LT}^* increases in the default cost but decreases in the time to debt maturity.

The implications of the other model parameters are rationalized similarly. A higher coupon rate, c ; principal, P ; or asset payout rate, δ ; decreases the value of equity, while a higher risk-free rate, r ; asset volatility, σ ; or tax benefits, τ ; increases the value of equity. Thus,

$$\begin{aligned} \frac{dV_{LT}^*}{d\alpha} &> 0, \frac{dV_{LT}^*}{dT} < 0, \frac{dV_{LT}^*}{dc} > 0, \frac{dV_{LT}^*}{dP} > 0, \frac{dV_{LT}^*}{d\delta} > 0, \\ \frac{dV_{LT}^*}{dr} &< 0, \frac{dV_{LT}^*}{d\sigma} < 0, \frac{dV_{LT}^*}{d\tau} < 0. \end{aligned} \quad (7)$$

4. Data: Sources and Descriptive Statistics

Data availability limits the analysis to corporate borrowers domiciled in the United States. Firm-specific borrower and debt characteristics are provided by Moody's (rating of senior unsecured debt,⁸ coupon rate, and time to maturity of outstanding bond issues), Bloomberg (book value of total debt⁹ and market capitalization), and DataStream (price of equity and dividend rate). In addition, Moody's provides data on the face value of defaulted debt and its price thirty days after default, which help estimate default recovery rates.

The paper also uses Moody's data on default rates over the one- and five-year horizons.^{10,11} The cohorts of firms tracked by Moody's for the calculation of a particular default rate are chosen according to the rating of senior unsecured debt. Not surprisingly, only a few AAA, AA, and A firms fail on debt obligations, implying that the default history of these firms carries insufficient information with

⁸When a firm does not have senior unsecured debt, Moody's estimates the rating of such debt via interpolation.

⁹Total debt includes all interest-bearing obligations.

¹⁰Specifically, the paper uses the Default Risk Service and Credit Risk Calculator databases of Moody's Investors Service. These databases provide information about all bond issues in Moody's rating universe as well as ratings and default data.

¹¹Moody's definition of default—i.e., the event in which a creditor incurs an economic loss—is in accordance with all the models considered in this paper. See Moody's Investors Service (1998) for further detail.

respect to changes in their creditworthiness. In addition, Moody's coverage of firms rated C or below is limited and prevents meaningful analysis. In light of this, the paper focuses on firms with a BBB, BB, or B rating.¹² Owing to rating migrations, the size of cohorts changes over time, averaging 517 (BBB-rated firms), 389 (BB-rated firms), and 482 (B-rated firms).

The macroeconomic variables used in this paper are provided by the International Monetary Fund, the Congressional Budget Office, and the Bank for International Settlements. These variables comprise an index of U.S. asset prices,¹³ the U.S. GDP gap, the U.S. credit-to-GDP ratio, the term spread in the Treasury rate, and the one-year Treasury rate. The first two variables are deflated by the U.S. Consumer Price Index. The credit-to-GDP ratio and the asset-price index are used as gaps from their respective stochastic trends, which, following Borio and Lowe (2002), are calculated on the basis of data available in real time.¹⁴ The term spread is set equal to the difference between the ten-year and three-month Treasury rates. In turn, the GDP gap is calculated as the difference between the log of real GDP and the log of potential real GDP. Finally, the one-year Treasury rate is used as a proxy for the risk-free rate of return.

The calculation of model-based PDs requires data from different sources, and it is the intersection of the Moody's and Bloomberg data sets that restricts the sample size. The upshot is that the available data allow for obtaining firm-specific PDs at a quarterly frequency, from 1990:Q1 to 2003:Q2. The smallest cross-sections of PDs are at the beginning of the sample: seventeen BBB-, fifteen BB-, and seven B-rated firms. The cross-sections expand over time and attain an average (maximum) size of 78 (140) for BBB-, 80 (127) for BB-, and 67 (166) for B-rated firms. The sample is dominated by nonfinancial

¹²In this paper, BBB refers to a Moody's rating between Baa1 and Baa3, BB to a rating between Ba1 and Ba3, and B to a rating between B1 and B3.

¹³The index is a weighted geometric mean of equity prices, and residential and nonresidential (commercial) property prices. The weights change through time and are based on households' annual net wealth. The data sources are the S&P Corporate 500 (equity prices), the National Council of Real Estate Investment Fiduciaries (commercial property prices), and the U.S. Office of Federal Housing Enterprise Oversight (residential property prices).

¹⁴Specifically, the date- t value of the trend is calculated via a Hodrick-Prescott (HP) filter, which uses data only up to time t . The parameter of the HP filter is set to 1600.

firms, which constitute 86 percent, 91 percent, and 92 percent of, respectively, the BBB-, BB-, and B-rated firms.^{15,16}

5. Calibration Methodology

This section consists of five subsections. The first three subsections divide model parameters into three groups and describe their calibration in detail. The fourth subsection summarizes the calibration results. Finally, the fifth subsection provides justification of the adopted calibration scheme and, in the process, reports and discusses the implications of alternative calibrations.

5.1 *Time-Invariant Parameters Common to All Firms*

Two parameters of the asset process are fixed both across firms and over time. One of them is the initial value of assets, V_0 , which is set to 100 without loss of generality because assets follow a geometric Brownian motion. The second parameter is the risk-free rate, which is set equal to the average one-year Treasury rate over the sample period: $r = 4.7$ percent.

In calibrating the LT and AST models, the applicable tax rate, τ , and the monitoring cost, m , are also held fixed. Following Leland (2004), $\tau = 0.15$. Since m and τ have similar implications for AST and LT PDs, respectively, the maintained assumption is that $m = 0.15$ as well.¹⁷

The final set of constant parameters are those specific to the exogenous-default models. Following Longstaff and Schwartz (1995),

¹⁵Financial firms enter Moody's calculation of default rates, and for consistency, such firms are considered for the derivation of theoretical PDs as well. That said, excluding financial firms from the sample leaves the results virtually unchanged.

¹⁶The reported sample sizes are obtained after cleaning the data in order to exclude leverage ratios, dividend rates, and asset volatilities that do not belong to the interval (0,1). (The calibration of leverage and asset volatility is described in section 5.2.) Such a filter removes a relatively small number of observations and is unlikely to influence the analysis. In addition, firm-quarter observations, which imply a default threshold larger than 90 percent of the assets' initial value, are also filtered out.

¹⁷When m is lowered to zero, the one-year PDs implied by the AST model increase on average by 0.3, 2, and 5 percentage points for BBB-, BB-, and B-rated firms, respectively.

Huang and Huang (2003), and Collin-Dufresne and Goldstein (2001), the parameters of the second stochastic process in the LS, HH, and CDG models are calibrated as follows: $k_r = 0.226$, $\bar{r} = 4.7$ percent, $\sigma_r = 0.0468$, and $\sigma_{rV} = -0.25$; $k_\lambda = 0.202$, $\sigma_\lambda = 0.031$, $\bar{\lambda} = 0.04165$, and $\sigma_{\lambda V} = -0.35$; $k_l = 0.2$ and $\nu = 0.6$.¹⁸

5.2 Time-Varying Firm-Specific Parameters

Most of the model parameters are calibrated at the firm level and change over time. This is the case for the coupon rate, c , and the time to maturity, T , which are obtained at the yearly frequency directly from data on bond issues.¹⁹ The maintained assumption is that c and T are representative of all (bank and nonbank) debt. In addition, leverage, l , is calibrated quarterly as the ratio of the book value of total debt to the sum of total debt and market capitalization. Then, following Huang and Huang (2003), the payout ratio is determined as $\delta = l * c + (1 - l) * d$, where d is the dividend rate paid out to the firm's equity holders and is obtained at the quarterly frequency.²⁰ For its part, the debt principal P is just rescaled leverage: $P = l * V_0$.

Firm-specific values of the asset risk premium and volatility are calculated at the quarterly frequency on the basis of stock market data and the model of Leland (1994). This model, which is a special case of the LT model and incorporates only perpetuities, implies that equity value, E , equals

$$E_t = V_t - \frac{(1 - \tau)cP}{r} + \left[\frac{(1 - \tau)cP}{r} - V^* \right] \left(\frac{V_t}{V^*} \right)^{-x},$$

¹⁸Alternatively, $\bar{\lambda}$ and ν could be calibrated to firm-specific data; i.e., to the time average of a firm's asset risk premium and leverage ratio, respectively. Switching to these alternative parameterizations has virtually no effect on the implied PDs.

¹⁹For each firm-year pair in the sample, c and T equal, respectively, the average coupon rate and time to maturity of the firm's outstanding bond issues. The averages use weights proportional to the face values of the corresponding bonds. The average time to maturity typically declines over time on a firm-by-firm basis and, thus, a calibrated value of T is roughly twice as large as the average time to maturity over the remaining life of the firm's outstanding debt. This is consistent with the maturity structure assumed in the LT model.

²⁰Ideally, d would also incorporate sales and repurchases of equity shares. The employed data sources do not, however, provide comprehensive coverage of these variables.

where
$$x = \frac{r - \delta}{\sigma^2} - \frac{1}{2} + \left[\left(\frac{r - \delta}{\sigma^2} - \frac{1}{2} \right)^2 + \frac{2r}{\sigma^2} \right]^{1/2} \text{ and}$$

$$V^* = \frac{(1 - \tau)cP}{r} \frac{x}{1 + x}.$$

The calibration of asset volatility, σ , relies on this equation and a numerical procedure described in Bharath and Shumway (2004). For each firm-quarter pair, this procedure requires an initial guess for σ , which is set equal to a leverage-adjusted volatility of equity returns over the previous year. Given a value for σ and a time series of E over the previous year, the above equation delivers a time series of asset returns. Then, this time series leads to an update of the value of σ and, thus, to an update of the time series of asset returns. The process continues until the absolute difference between two consecutive values of σ is less than 0.001.

The calibration of the asset risk premium λ incorporates a quarterly estimate of the corresponding equity risk premium, λ^E . This estimate is obtained in two steps. The first step derives a time series of marketwide equity premiums that (i) averages 8 percent over the sample period and (ii) exhibits the quarter-on-quarter growth rates estimated by Tarashev and Tsatsaronis (2006) for the S&P 500 stock market index. For each quarter, the second step employs results of Bhandari (1988) in order to estimate firm-specific equity premiums on the basis of the associated leverage ratios. Each cross-section of equity premiums obtained at the second step is rescaled so that its average matches the corresponding marketwide risk premium obtained at the first step.²¹

Given estimates of λ^E and σ , the estimate of λ relies on the Leland (1994) equation for the firm's equity value. Applying Ito's lemma to that equation delivers²²

$$\lambda = \delta - r + \left[(\lambda^E + r - d)(1 - l) - \frac{\sigma^2}{2} x(1 + x)\Omega \right] \frac{1}{1 - x\Omega},$$

²¹The transformation is necessitated by the fact that the results of Bhandari (1988) abstract from time variation in the equity market risk premium.

²²Note that the variability of the equity drift over time is assumed to be absorbed by λ^E .

where $(\lambda^E + r - d)$ is the equity drift, $\Omega = ((1 - \tau)\frac{cP}{rV} - \frac{V^*}{V})(\frac{V^*}{V})^x$, and $(1 - l)$ proxies for $\frac{E}{V}$.

5.3 Default Threshold and Dead-Weight Cost of Default

The last two parameters that remain to be set are the default threshold, V^* , and the fraction of assets lost in default, α (or the fixed bankruptcy cost, K , in the AST model). The calibration of these parameters incorporates an estimate of the default recovery rate, ρ , which is defined as the price of debt thirty days after default divided by the associated face value. The estimate of ρ is set to 47 percent,²³ which is the average of the ex post recovery rates in Moody's default database.²⁴

Under an endogenous-default model, the values of V_{LT}^* and α (LT model) or V_{AST}^* and K (AST model) are determined *simultaneously*. Specifically, each of these parameter pairs comprises the two unknowns in two equations:

$$\begin{aligned} V_{LT,it}^* &= V_{LT}^*(\alpha_t; T_{it}, c_{it}, P_{it}, \delta_{it}, r, \sigma_{it}, \tau) \text{ and} \\ \rho &= \frac{1}{N_t} \sum_{i=1}^{N_t} \frac{(1 - \alpha_t) V_{LT,it}^*}{P_{it}} \text{ for the LT model and} \\ V_{AST,it}^* &= V_{AST}^*(K_t; c_{it}, P_{it}, \delta_{it}, r, \sigma_{it}, m) \text{ and} \\ \rho &= \frac{1}{N_t} \sum_{i=1}^{N_t} \frac{V_{AST,it}^* - K_t}{P_{it}} \text{ for the AST model,} \end{aligned} \quad (8)$$

where t indexes the quarter and i the firm, N_t is the total number of firms in quarter t , and $V_{LT}^*(\bullet; \dots)$ and $V_{AST}^*(\bullet; \dots)$ summarize the endogenous nature of the two default thresholds.

²³Alternatively, it is possible to estimate a time series of ρ so that each value in this series is based on default data from the previous several years. Unfortunately, the scarcity of the data that can be used to measure recovery rates leads to noisy time-varying estimates of ρ . Specifically, there is a total of 561 default events that can be used to measure recovery rates—or, on average, 40 defaults per year from 1990 to 2003.

²⁴For consistency with the way Moody's calculates default rates, as well as with technical assumptions of the theoretical models under study, the calibrated recovery rate is based only on defaults on senior unsecured debt.

By contrast, the exogenous-default models require that the default cost, α , or the default threshold, V^* , be determined in an ad hoc fashion. In this paper, α is set to 45 percent, which is the average value of this parameter under the LT model. Then, the exogenous-default threshold is allowed to vary quarterly and across firms according to a simplified version of equations (8):

$$\rho = \frac{(1 - \alpha)V_{M,it}^*}{P_{it}} = \frac{0.55V_{M,it}^*}{P_{it}}, \quad (9)$$

where M stands for the CDG, the HH, or the LS model. In the CDG model, where the default threshold is time varying, the calibrated value of $V_{CDG,it}^*$ is the *initial* value of the parameter.

5.4 Calibration Summary

Tables 1 and 2 provide alternative perspectives on the calibration of the models. Table 1 lists all the parameters and summarizes features of their calibration. For its part, table 2 reports parameter averages alongside their counterparts calibrated by Huang and Huang (2003) and/or Leland (2004).

The differences between the calibration results of this paper and those in the previous literature are generally small and are due primarily to three factors. First, this paper calibrates separately three rating classes, whereas the earlier papers extract certain parameters from data on the entire spectrum of ratings (table 2, fourth column). Second, previously used parameters are based on data that start in 1970, whereas the data used herein start in 1990. Third, the calibration in this paper incorporates total debt, whereas the literature has relied on total *liabilities*.²⁵ The effect of this surfaces most clearly in the values of leverage and asset volatility, which are, respectively, lower and higher than their analogues in earlier papers.

²⁵In addition to total debt, total liabilities include obligations that do not involve interest payments (e.g., promises for physical deliveries). There are two reasons for choosing to work with total debt as opposed to total liabilities. First, a leverage ratio that is based on total liabilities would lead to an over-estimation of coupon payments and payout rates. Second, total debt tracks more closely the bond instruments underlying the calculation of default recovery rates.

Table 1. Model Parameters and Their Calibration^a

Parameter	Description	Firm Specific Y or N	Time Varying Y (y or q) or N
c	Coupon Rate	Y	Y(y)
T	Time to Maturity	Y	Y(y)
r	Risk-Free Rate of Return	N	N
l	Leverage Ratio	Y	Y(q)
δ	Asset Payout Rate	Y	Y(q)
λ	Asset Risk Premium	Y	Y(q)
σ	Asset Volatility	Y	Y(q)
ρ	Default Recovery Rate	N	N
α or K	Default Cost ^b	N	Y(q) in Endogenous- Default Models N in Exogenous- Default Models
V^*	Default Boundary ^b	Y	Y(q)
τ	Tax Rate	N	N
m	Monitoring Cost	N	N
$k_r, \sigma_r, \bar{r}, \sigma_r V,$	Parameters of the Second Stochastic		
$k_\lambda, \sigma_\lambda, \bar{\lambda}, \sigma_\lambda V,$	Process (Exogenous- Default Models)	N	N
k_l, ν			

Note: Y = yes, N = no, y = yearly, q = quarterly.
^aUnless stated otherwise, the parameter values are the same across models and are based on assumptions adopted by Huang and Huang (2003) and/or Leland (2004).
^bThe calibrated values of the default boundary and the (dead-weight) default cost depend on the type of the underlying model and on the values of other model parameters. In the exogenous-default models, α is fixed at 45 percent and, given an estimate of leverage, V^* is set to be consistent with an estimate of the default recovery rate. In the endogenous-default models, the values of α (or K) and V^* are determined simultaneously by estimates of the following debt characteristics: coupon rate, time to maturity (LT model only), risk-free rate of return, leverage, asset payout rate, asset volatility, default recovery rate, relevant tax rate (LT model only), and monitoring cost (AST model only).

5.5 Discussion of the Calibration Methodology

The calibration scheme described above incorporates several important choices that either need to be justified or compared with alternatives. This is done in the present subsection.

Table 2. Key Parameters of the Structural Credit-Risk Models

	BBB-Rated Firms ^a (Averages)	BB-Rated Firms ^a (Averages)	B-Rated Firms ^a (Averages)	Rep. Firm in the Literature ^b
Drift ($\lambda + r - \delta$)	7.6%	7.4%	5.6%	6%
Payout Rate (δ)	3.6%	3.8%	5.1%	6%
Volatility (σ)	26%	31%	34%	23%
Leverage (l)	30%	37%	50%	43.3% (BBB) 53.5% (BB) 66% (B)
Coupon Rate (c)	7%	8%	9%	8%
Time to Maturity (T)	10 Years	8 Years	7 Years	10 Years

^aBased on a calibration to firm-specific time-varying parameters (see section 5).

^bParameter values adopted by Huang and Huang (2003) and/or Leland (2004).

One of the important choices lies in the decision to calibrate time-varying parameters even though the models assume that (virtually) all of their parameters stay constant over time. Since key parameters reflect risk premiums, debt-service payments, equity volatility, etc., this assumption of the models should be interpreted as referring to steady-state borrower characteristics that convey a long-term level of risk. Thus, the assumption should not be taken literally when the objective is to evaluate the time path of model forecasts with a short—e.g., one-year—horizon, over which credit risk depends significantly on transitory shocks. This motivates calibrating model parameters to their time-varying short-term estimates. Provided that debt characteristics can be expected to change little over the PD horizon, the adopted calibration procedure will be roughly consistent with the models’ “constant parameter” assumption and will allow for intertemporal variability in each firm’s credit risk.²⁶

Nevertheless, it is natural to ask if the incorporation of time-varying parameters affects adversely the performance of the models.

²⁶For the endogenous-default models, it is also necessary to assume that parameter variability is small enough to have a negligible impact on the default threshold. In general, the latter value would depend on both short- and long-term borrower characteristics.

In light of this, the paper also analyzes the implications of an alternative calibration, which sets parameter values to firm-specific time averages.

Another important choice is the calibration of a constant risk-free rate of return, r . The main reason for this choice is that it is inherently difficult to pin down the value of r used by the market at any specific point in time (see Feldhütter and Lando 2004). In addition, a particular calibration of a time-varying risk-free rate—which sets r to the average Treasury rate over the corresponding quarter—does not alter the overall level of model-based PDs but impairs, albeit slightly, the explanatory power of these PDs for intertemporal changes in default rates.²⁷

The calibration of the asset risk premium, λ , and volatility, σ , also warrants an explanation. There are two main reasons for basing this calibration on the model of Leland (1994), as opposed to on models whose default-rate forecasts are analyzed in this paper. First, the Leland (1994) model delivers endogenous debt and equity valuation, which is key for relating λ and σ to observable stock market data but is beyond the scope of the exogenous-default HH, LS, and CDG models. Second, the default threshold in the Leland (1994) model is independent of default costs, which is key for operationalizing the entire calibration scheme but is not true in the AST and LT models. In addition to these reasons, keeping λ and σ constant across models allows for an evaluation of how the models process the same information.

As far as the asset risk premium λ is concerned, it may seem reasonable to estimate it—just like asset volatility—on the basis of model-implied asset returns over the previous year. Unfortunately, this alternative approach delivers extremely noisy estimates. To see this, one can start by calculating the mean and the standard deviation in the time series of asset risk premiums estimated for each firm under the alternative approach. The resulting cross-section of means averages 0.05, while the cross-section of standard deviations exhibits a significantly higher average of 0.3. For comparison, the corresponding statistics are 0.07 and 0.014 for the estimates of λ used in this paper and 0.3 and 0.065 for the estimates of σ .

²⁷The results of this background exercise are available from the author upon request.

Finally, the calibration of the default threshold and the default dead-weight cost follows Leland (2004). In the context of the AST and LT models, this calibration incorporates the fact that (i) there is no consensus in the literature regarding the value of the default-cost parameter,²⁸ and (ii) this value can be derived simultaneously with the associated default threshold on the basis of data on recovery rates. An alternative scheme, which ignores information on recovery rates and incorporates an ad hoc value of default costs, is bound to introduce much arbitrariness in endogenous-default thresholds.

In contrast to the LT and AST models, the default dead-weight cost α —or, by equation (9), the default threshold—*must* be set in an ad hoc fashion under the CDG, HH, and LS models. The adopted calibration approach, explained in section 5.3, attains two objectives. First, by endowing these models with the average α obtained under the LT model, the approach avoids a predetermined wedge between the PD predictions of the exogenous-default and endogenous-default models. Second, by keeping the exogenous-default α constant over time, in contrast to its counterpart in the LT model, the calibration reflects the fact that the latter model provides a researcher with more information about the value of α . Importantly, it turns out that an alternative approach that endows the LT, CDG, HH, and LS models with the same time-varying value of α blurs in effect any distinction among the first three models.²⁹

6. Model-Implied Expectations of Default Rates

The analysis in this paper draws parallels between model-based PDs and actual default rates. Implicitly, this analysis assumes that the parameter values used here for the calculation of model-based PDs accurately characterize the firms considered by Moody's for the calculation of default rates. Under such an assumption, a model's failure to account for default rates is attributed to the specification of the model.

²⁸At one extreme, Andrade and Kaplan (1998) argue that α should not exceed 20 percent; at the other extreme, Leland and Toft (1996) set this parameter to 50 percent.

²⁹The same point is made in a slightly different context by Leland (2004).

With this in mind, it is reassuring that model-based expectations of one-year default rates—i.e., cross-sectional averages of firm-specific PDs—are, in general, quite in line with actual one-year default rates. This is seen in figure 1, which reveals that the time series of such expectations tends to match the time series of default rates realized over the following year. In addition, table 3 (panel 1) reveals that—with the exception of the CDG and, to a lesser extent, the AST model—the structural credit-risk models exhibit little bias

Figure 1. Expected One-Year Default Rates (by Model) versus Actual Default Rates

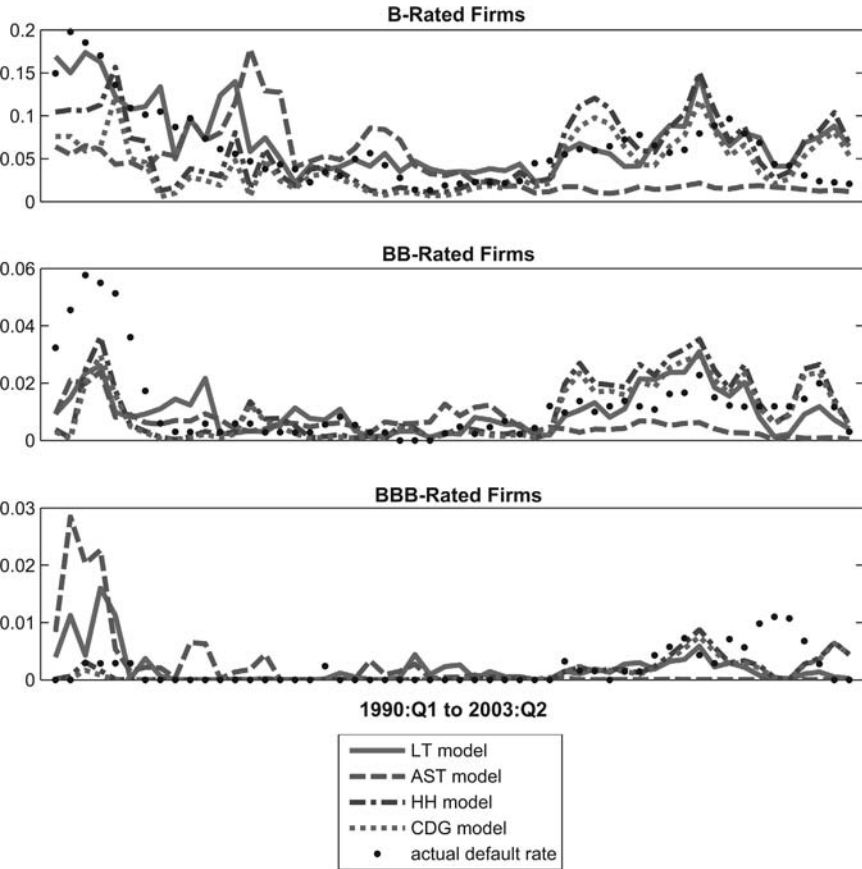


Table 3. Expected and Actual Default Rates (in percentage points)

One-Year Horizon														
B-Rated Firms					BB-Rated Firms					BBB-Rated Firms				
	Mean	Median	Min.	Max.		Mean	Median	Min.	Max.		Mean	Median	Min.	Max.
Actual Default Rate	6.3	5.5	1.3	19.8		1.20	0.90	0.00	5.80		0.19	0.00	0.00	1.10
<i>Panel 1:</i> Expected Default Rate ^a	LT	5.9	2.2	17.4	LT	0.99	0.81	0.11	3.10	LT	0.19	0.09	0.00	1.61
	AST	4.3	3.1	1.0	AST	0.62	0.57	0.02	2.40	AST	0.23	0.02	0.00	2.85
	HH	5.8	4.4	1.1	HH	1.10	0.52	0.03	3.62	HH	0.13	0.01	0.00	0.88
	LS	5.8	4.4	1.0	LS	1.10	0.52	0.03	3.69	LS	0.13	0.01	0.00	0.88
	CDG	4.3	3.3	0.6	CDG	0.94	0.45	0.02	2.97	CDG	0.12	0.02	0.00	0.76
<i>Panel 2:</i> Expected Default Rate, Typical Firm Parameters ^b	LT	6.7	5.8	3.8	LT	0.51	0.48	0.20	1.12	LT	0.03	0.01	0.00	0.12
	AST	3.9	3.7	1.0	AST	0.42	0.42	0.07	0.86	AST	0.09	0.03	0.00	0.56
	HH	3.4	4.0	0.5	HH	0.35	0.15	0.01	1.05	HH	0.03	0.01	0.00	0.31
	LS	3.4	4.0	0.5	LS	0.36	0.16	0.01	1.06	LS	0.03	0.01	0.00	0.32
	CDG	2.5	2.8	0.4	CDG	0.34	0.12	0.01	0.97	CDG	0.03	0.01	0.00	0.32
<i>Panel 3:</i> Expected Default Rate, Representative Firm ^c	LT	1.1	0.9	0.1	LT	0.03	0.00	0.00	0.31	LT	0.0001	0.0000	0.0000	0.0017
	AST	0.8	0.2	0.0	AST	0.01	0.00	0.00	0.20	AST	0.0006	0.0000	0.0000	0.0097
	HH	2.0	1.0	0.0	HH	0.11	0.00	0.00	0.98	HH	0.0009	0.0000	0.0000	0.0141
	LS	2.0	1.0	0.0	LS	0.11	0.00	0.00	0.97	LS	0.0009	0.0000	0.0000	0.0137
	CDG	1.5	0.7	0.0	CDG	0.09	0.01	0.00	0.65	CDG	0.0012	0.0000	0.0000	0.0177
<i>Panel 4:</i> Expected Default Rate, Rep. Firm in the Literature ^d	AST	0.250			AST	0.0289				AST	0.002910			
	HH	0.055			HH	0.0012				HH	0.000010			
	LS	0.052			LS	0.0011				LS	0.000010			
	CDG	0.032			CDG	0.0010				CDG	0.000015			

^aBased on a calibration to firm-specific time-varying parameters (see section 5).
^bBased on average parameters (over time) for each firm.
^cBased on average parameters (across firms) in each quarter.
^dBased on parameter values adopted by Huang and Huang (2003) and/or Leland (2004).

^aBased on a calibration to firm-specific time-varying parameters (see section 5).^bBased on average parameters (over time) for each firm.^cBased on average parameters (across firms) in each quarter.^dBased on parameter values adopted by Huang and Huang (2003) and/or Leland (2004).

vis-à-vis actual default rates.³⁰ This is evidence in support of the models because the intertemporal mean of the default-rate expectations implied by an accurate model converges to the corresponding mean of the actual default rates as the length of the sample time series expands.

Next, it is important to examine whether the performance of the models is robust to changes in the forecast horizon. In light of this, table 4 (panel 1) reports summary statistics for model-based expectations of five-year default rates alongside the corresponding statistics for actual default rates. The expectations are based on typical firm-specific characteristics, captured by time averages of model parameters. Typical firm characteristics relate to the “steady-state” credit outlook, which is likely to dominate in the determination of long-horizon PDs.

Five-year expected default rates, which would be a key input into risk-management schemes with a long-term focus, tend to match closely actual default rates. The endogenous-default LT and AST models attain a close match across all rating classes considered. In comparison, the performance of the exogenous-default models is poorer: the CDG model consistently underpredicts default rates in the B rating class, while the HH and LS models consistently underpredict in the BBB rating class.

6.1 *The Importance of Firm-Specific Calibration*

In a recent paper, Leland (2004) finds that—when calibrated to the representative firm—the LT model and (a simplified version of) the LS model consistently and substantially underpredict actual one-year default rates. The data used in the present paper confirm this finding. Employing these data to calibrate the five structural credit-risk models to the representative firm—which is endowed with average borrower characteristics—delivers PDs that are *orders of magnitude* smaller than actual default rates. As revealed by table 3, this result, which is robust across models and rating classes, obtains

³⁰Figure 1 and most of the subsequent analysis abstracts from the output of the LS model because, as testified by the results in table 1, this output is virtually indistinguishable from the output of the HH model (see also section 6.3 below).

Table 4. Expected and Actual Default Rates (in percentage points)

Five-Year Horizon														
B-Rated Firms					BB-Rated Firms					BBB-Rated Firms				
	Mean	Median	Min.	Max.		Mean	Median	Min.	Max.		Mean	Median	Min.	Max.
Actual Default Rate	22.2	21.7	12.5	34.8		6.9	7.0	1.5	14.7		1.22	0.72	0.00	3.70
<i>Panel 1:</i> Expected Default Rate ^a	LT	24.3	20.2	34.2	LT	6.4	5.3	3.9	11.2	LT	0.93	0.61	0.29	1.63
	AST	22.6	23.3	17.4	AST	6.3	5.7	4.5	10.1	AST	1.29	0.81	0.50	4.05
	HH	18.6	16.5	11.5	HH	5.3	4.0	2.2	11.0	HH	0.65	0.47	0.07	1.82
	LS	17.6	14.8	10.7	LS	5.3	3.9	2.1	10.8	LS	0.69	0.52	0.09	1.83
	CDG	13.1	11.5	6.3	CDG	5.3	4.4	1.8	10.2	CDG	0.94	0.74	0.09	2.80
<i>Panel 2:</i> Expected Default Rate, Representative Firm ^b	LT	21.4	20.4	14.1	LT	3.7	2.6	1.8	8.0	LT	0.26	0.24	0.07	0.67
	AST	19.0	19.0	14.4	AST	4.3	3.8	2.4	7.7	AST	0.57	0.39	0.24	1.53
	HH	18.3	16.4	8.4	HH	3.3	1.9	0.9	8.6	HH	0.16	0.08	0.02	0.69
	LS	17.1	14.6	7.4	LS	3.2	1.9	0.9	8.1	LS	0.18	0.10	0.03	0.73
	CDG	9.1	6.0	2.5	CDG	2.7	1.8	0.7	6.5	CDG	0.23	0.10	0.04	1.01
<i>Panel 3:</i> Expected Default Rate, Rep. Firm in the Literature ^c	LT	8.7			LT	3.2				LT	1.00			
	AST	12.4			AST	6.9				AST	3.77			
	HH	9.7			HH	3.9				HH	1.30			
	LS	7.1			LS	2.7				LS	0.85			
	CDG	2.6			CDG	1.4				CDG	0.73			

^aBased on average parameters (over time) for each firm.

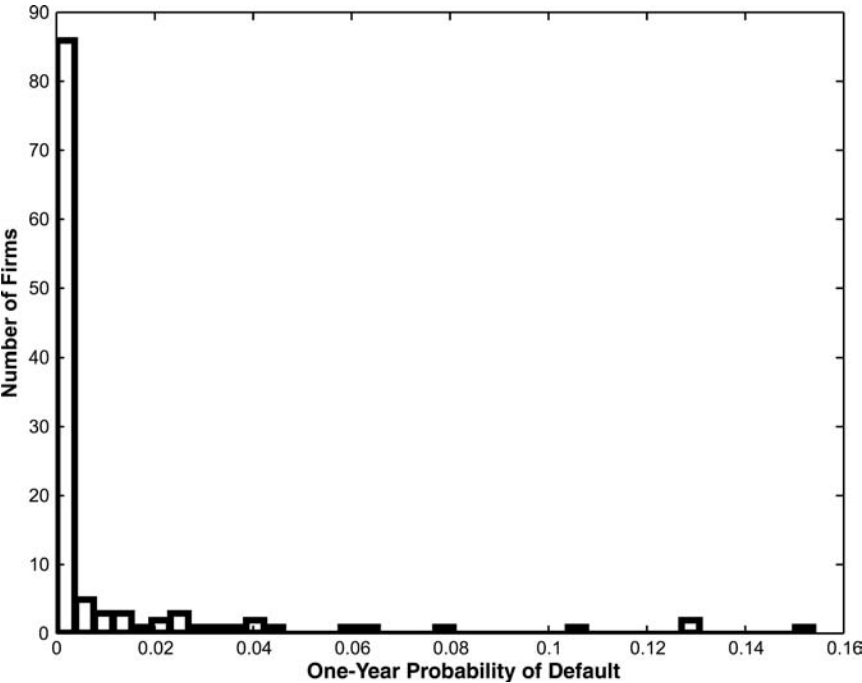
^bBased on average parameters (across firms) in each quarter.

^cBased on parameter values adopted by Huang and Huang (2003) and/or Leland (2004).

irrespective of whether one allows the characteristics of the representative firm to change from quarter to quarter (panel 3) or considers a time-invariant representative firm, as done in the extant literature (panel 4).

As conjectured by Leland (2004), a calibration to the representative firm biases substantially downward model-based forecasts of one-year default rates. Figure 2 illustrates this point by focusing on a particular quarter—1998:Q4—and portraying the distribution of PDs implied by the LT model under the firm-specific calibration. This distribution has a pronounced right tail, implying that the mean (1.1 percent) is substantially larger than the median (0.0393 percent). The former statistic equals the expected default rate, implied by the LT model on the basis of the firm-specific

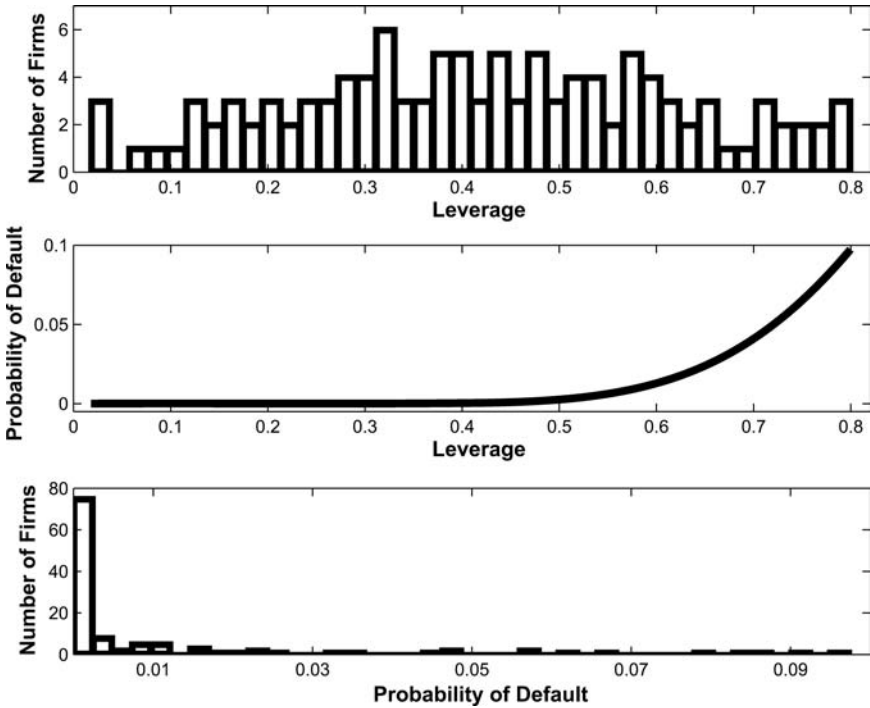
Figure 2. PDs Implied by the LT Model for Firms Rated BB in 1998:Q4



calibration, while the latter statistic is extremely close to the PD of the average (i.e., representative) firm (0.0395 percent).³¹

The strong dependence of expected default rates on whether the underlying model is calibrated to firm-specific parameters or to the representative firm is a particular manifestation of the so-called Jensen inequality effect. The Jensen inequality effect, which arises because of the convex mapping from firm characteristics into model-based PDs, can be best appreciated if such a PD is considered as a function of a single characteristic. Figure 3 illustrates such

Figure 3. The Jensen Inequality Effect (LT Model, BB-Rated Firms, 1998:Q4)



³¹The close match between the median of firm-specific PDs and the PD of the representative firm reveals also that firm-specific parameters (i) have an approximately symmetric distribution in the cross-section and (ii) exhibit a generally monotonic impact on model-based PDs.

a hypothetical scenario by focusing on the leverage ratios of the firms behind the plot in figure 2. In particular, figure 3 reveals that, *ceteris paribus*, the LT model translates a virtually symmetric distribution of leverage ratios (with a skewness of 0.015) into a PD distribution that has a long right tail (i.e., a skewness of 2.6). It is this tail that drives the large positive difference between the average PD and the PD of the average (representative) firm.

Likewise, the five-year PDs of the representative borrower are invariably lower than the expected default rates based on firm-specific PDs (see table 4, panels 2 and 3). However, again in line with the findings of Leland (2004), the resulting negative bias in representative-borrower PDs is smaller than at the one-year horizon.

6.2 *The Impact of Time-Varying Parameters*

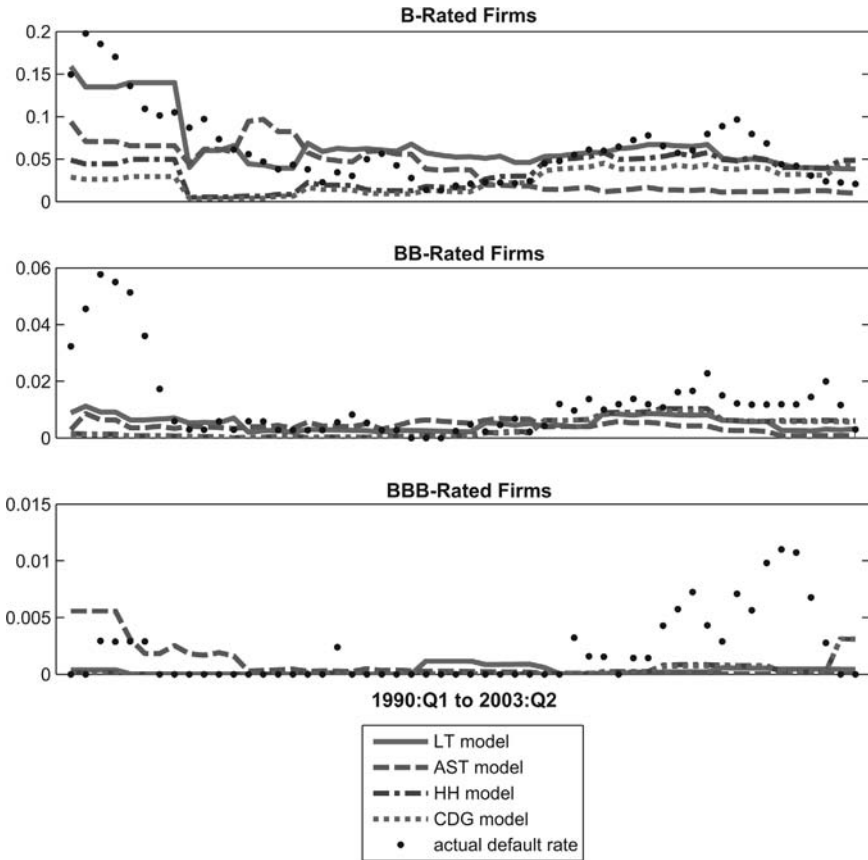
In order to see how the calibration of time-varying parameters affects the performance of the models, an alternative set of one-year expected default rates are calculated on the basis of typical (or average) firm-specific characteristics.³² As revealed by table 3 (panel 2) and figure 4, these expected default rates are (i) generally lower and (ii) exhibit less variation over time than the expected default rates obtained under time-varying parameters. The first feature reveals that the switch from time-varying to typical parameters reduces the dispersion of borrower characteristics in the cross-section, which dampens the Jensen inequality effect discussed above and, as a result, depresses model-based PDs. The second feature is due to the fact that changes in the alternative expected default rates abstract from changes in credit risk at the firm level and are driven solely by exits from and entries into rating classes.

6.3 *Comparison across Models*

As illustrated by figure 1, default-rate expectations differ considerably across models. This section relates some of these differences to

³²Recall that all model-based five-year PDs incorporate time-invariant parameters.

Figure 4. Expected One-Year Default Rates (by Model, Constant Parameters) versus Actual Default Rates



the specification of the models and the way in which specification differences affect calibration.

When considering the endogenous-default LT and AST models, it is necessary to note that differences in their default forecasts could stem solely from differences in the implied default thresholds, V_{LT}^* and V_{AST}^* . As outlined in section 5.3, V_{LT}^* and V_{AST}^* are calibrated together with the associated default-cost parameters, α and K , to match a time-invariant recovery rate. Background analysis (available upon request) unveils two stylized facts about this calibration. First,

model-implied recovery rates decrease in α (or K).³³ Second, from the remaining parameters, only changes in asset volatility σ have a first-order impact on model-implied recovery rates. These two stylized facts imply that, given a target recovery rate, a rise in σ —which puts downward pressure on model-implied recovery rates by reducing V_{LT}^* (V_{AST}^*)—necessitates a downward revision in α (K), leading to a further reduction of V_{LT}^* (V_{AST}^*).³⁴ Formally, denoting an endogenous-default model by M , $\frac{\partial V_M^*}{\partial \sigma} < 0$. In conjunction with (2), this implies that the impact of asset volatility on model-based PDs, $\frac{dPD_M}{d\sigma} = \frac{\partial PD_M}{\partial V_M^*} \frac{\partial V_M^*}{\partial \sigma} + \frac{\partial PD_M}{\partial \sigma}$, could be of either sign.

The impact of asset volatility on endogenous-default thresholds drives the wedge between PDs implied, respectively, by the LT and AST models. This impact is much stronger in the context of the latter model and leads to AST PDs that are *negatively* correlated with asset volatility. This negative correlation underpins, e.g., the AST model's underprediction of default rates in the second half of the sample (when σ is on the rise) and the overprediction of default rates in the BBB rating class at the beginning of the sample (when σ is relatively lower). By contrast, changes in σ have a sufficiently small (negative) impact on V_{LT}^* so that LT PDs correlate positively with asset volatility.

At the same time, by equation (9), the calibrated values of exogenous-default thresholds are insensitive to changes in asset volatility. This drives a wedge between the predictions of the endogenous-default and the exogenous-default (i.e., LS, HH, and CDG) models. In addition, in light of the discussion in the previous two paragraphs, it is not surprising to find that the differences between PDs based on an endogenous-default model and PDs based on an exogenous-default model are correlated negatively with asset volatility.

Finally, a comparison across the three exogenous-default models reveals remarkable similarities. As reported in table 3, the expected default rates delivered by the HH and LS models are virtually indistinguishable. By extension, this implies that the two models'

³³Note that this stylized fact does not have to hold a priori because the endogenous V_{LT}^* and V_{AST}^* increase, respectively, in α and K , which puts upward pressure on model-implied recovery rates.

³⁴To see this, recall expressions (6)–(8).

second stochastic processes—which are the only sources of potential differences—have a negligible impact on implied PDs. Thus, in what is to follow, analysis of the HH model’s predictions applies equally to the LS model. Turning to the CDG model, the expected default rates it implies correlate strongly (with a correlation coefficient of 0.98 or higher) with their counterparts obtained under the HH or LS models. That said, the CDG model specifies strong mean reversion of the ratio V_{CDG}^*/V to a level above unity, which depresses CDG PDs relative to HH and LS PDs.³⁵

7. Statistical Significance of Model-Based Forecast Errors³⁶

Focusing on the one-year horizon, this section examines whether actual default rates support the probability distribution of default rates implied by model-based firm-specific PDs. Unfortunately, a key determinant of such a distribution, i.e., the interdependence of defaults across firms, cannot be estimated on the basis of the available data and has to be calibrated to some rule-of-thumb value.

Concretely, the probability distribution of default rates is constructed as follows. First, for a particular model and quarter in the sample, it is assumed that there are firm-specific standard normal variables that generate defaults—by falling below firm-specific thresholds—with *marginal* probabilities equal to the corresponding model-based PDs. Second, in line with Bohn, Arora, and Korablev (2005), the correlation among these random variables is assumed to be 20 percent.³⁷ This allows for generating

³⁵ As the mean-reversion parameter, k_t , decreases toward zero, CDG PDs converge to HH and LS PDs.

³⁶ This subsection builds on the analysis in Kurbat and Korablev (2002) and the related numerical procedure outlined in Bohn, Arora, and Korablev (2005).

³⁷ There are two reasons for implementing this correlation coefficient. First, Kurbat and Korablev (2002) report that a homogeneous median correlation works well as a proxy for a heterogeneous correlation structure. Motivated by this finding, Bohn, Arora, and Korablev (2005) examine proprietary data and settle on correlations of 18 percent and 19 percent in deriving the probability distribution of default rates. Second, robustness checks reveal that changing the homogeneous correlations of the default generators between 5 percent and 25 percent, which is the range invoked in the literature on portfolio credit risk, has little material impact on the main results.

interdependence across default events while circumventing the use of asset return correlations, which could be incorporated into the models but cannot be estimated from the available data. On the basis of this structure, a large number of draws (concretely, 100,000) from the joint distribution of the default generators delivers the probability distribution of default rates implied by the model in focus.

This exercise reveals that the model-implied probability distributions of default rates are generally consistent with actual default rates, even though there are noteworthy differences across models. Figures 5–8 display quarterly time series of the fifteenth and ninety-fifth percentiles (with the associated 80 percent confidence interval) and the medians of the model-implied probability

Figure 5. Probability Distribution of Default Rates (Based on LT Model) versus Actual Default Rates

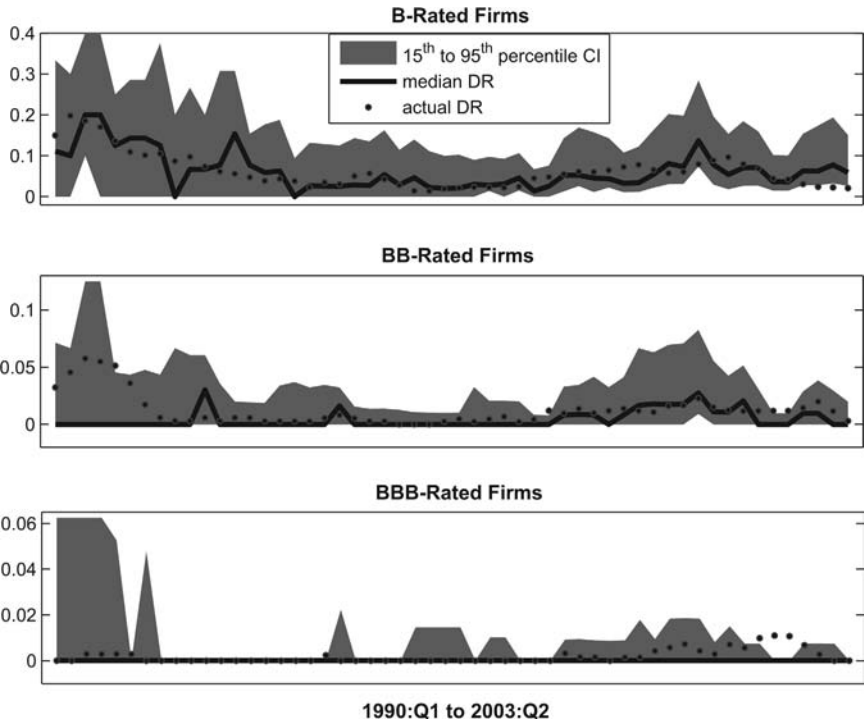
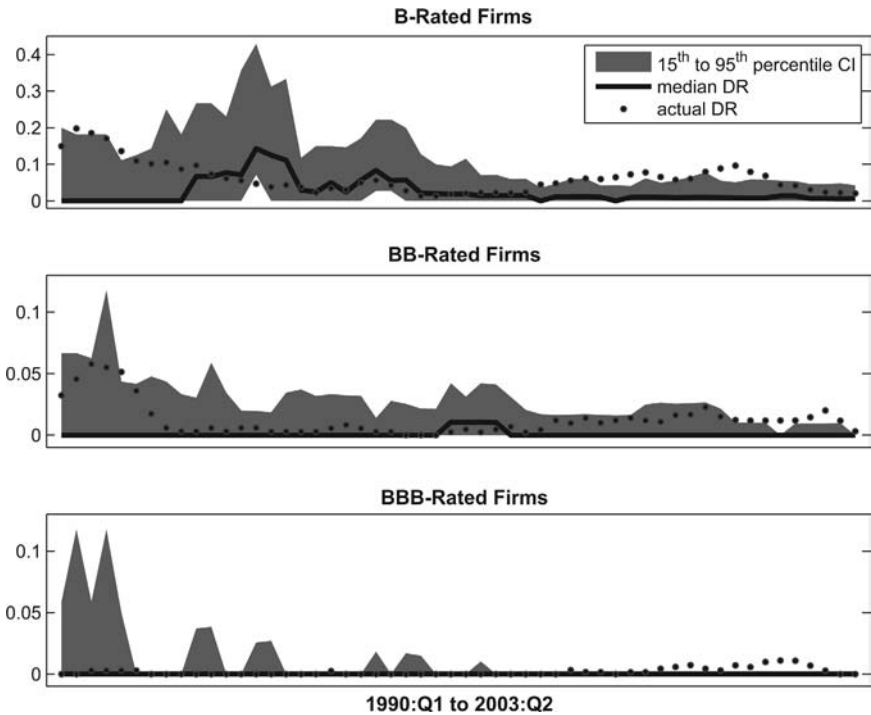


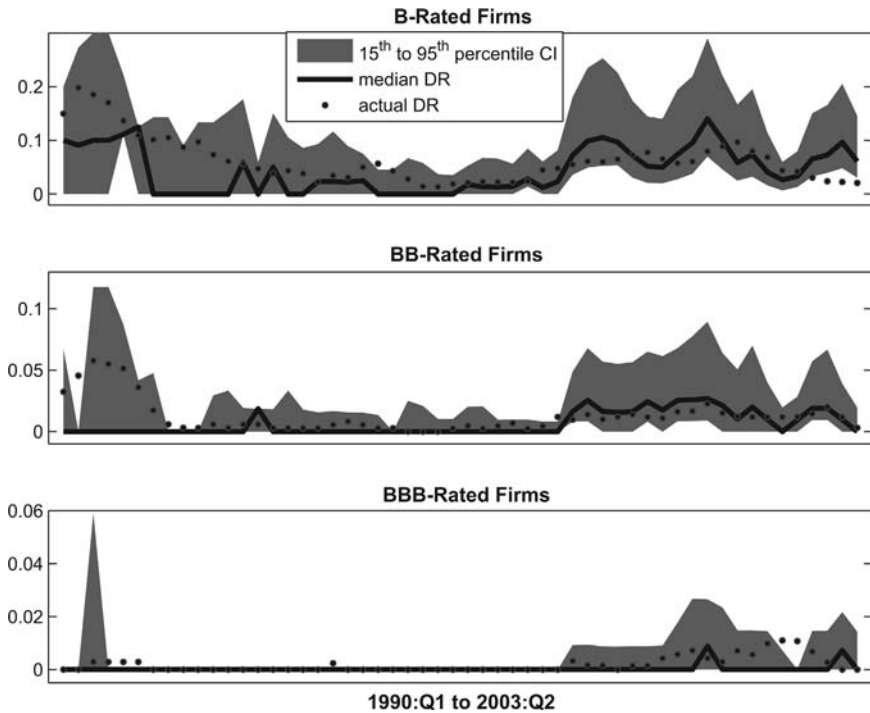
Figure 6. Probability Distribution of Default Rates (Based on AST Model) versus Actual Default Rates



distributions,³⁸ alongside the corresponding actual default rates. The vast majority of these default rates fall within the 80 percent confidence intervals implied by the LT and HH models. The former model misses, at most, five (out of fifty-four) default rates in a rating class, and the misses of the latter model do not exceed six. By contrast, the performance of the AST and CDG models is much poorer, with the former missing as many as nineteen default rates in the BBB rating class and the latter following behind with ten misses in the BB rating class. One would note that most of the misses are violations of the upper bound of the confidence

³⁸Note that the means of these distributions equal the cross-sectional averages of PDs, which are plotted in figure 1.

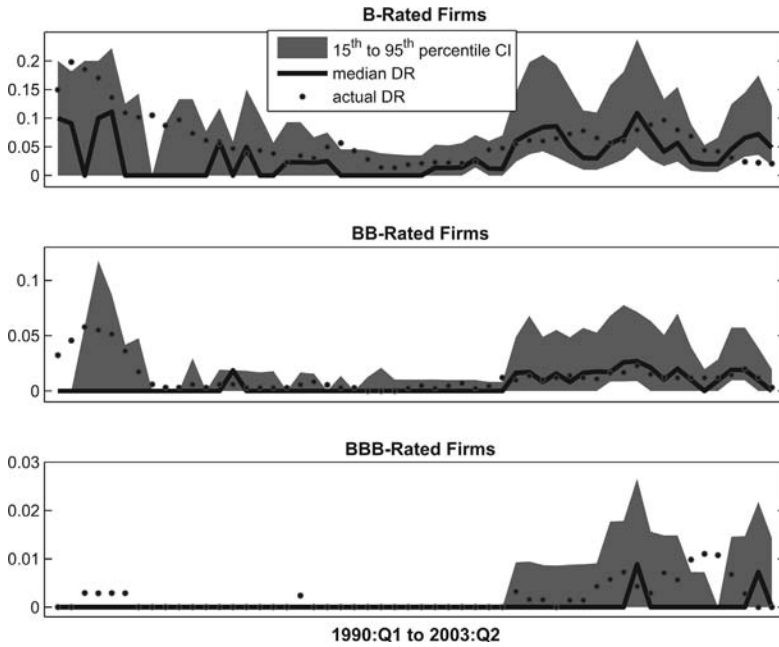
Figure 7. Probability Distribution of Default Rates (Based on HH Model) versus Actual Default Rates



intervals—driven by too-low model-based PDs—whereas even the unconventionally high lower bound (set to the fifteenth percentile) is violated seldom and only in the context of the B rating class. Even in this class, however, the more conventional fifth percentile of the distributions of default rates is invariably zero and is, thus, never violated.³⁹

³⁹At the five-year horizon, actual default rates almost always belong to the 80 percent confidence intervals implied by model-based PDs. The exceptions, which emerge only in the BBB rating class, occur in five, eight, and four quarters under the AST, HH, and CDG models, respectively, and are always associated with default rates that are higher than the corresponding model-based ninety-fifth percentile.

Figure 8. Probability Distribution of Default Rates (Based on CDG Model) versus Actual Default Rates



8. Economic Significance of Model-Based Forecast Errors

Having discussed the statistical significance of forecast errors, it is also of interest to evaluate their economic significance. This section conducts such an evaluation in a stylized setting in which lenders determine the amount of capital to set aside on the basis of the foundation IRB approach of Basel II and default-rate expectations implied by the firm-specific calibration of a particular credit-risk model. These lenders' capital is compared to two benchmarks: (i) capital based on ex post information about credit risk and (ii) capital based on the credit rating of the exposures.

Consider six hypothetical banks that invest in a bond at time t . The true one-year PD of that bond is assumed to equal the default rate in a particular rating class (BBB, BB, or B), which is realized over the year starting at t . Two of the banks are used as benchmarks. One of them, the “perfect-foresight” bank, calculates

its regulatory capital in a particular quarter by setting the PD equal to the one-year default rate realized over the following year and using the foundation IRB approach.⁴⁰ The other benchmark bank adopts the standardized approach (SA) of Basel II. The capital requirements of the SA bank are thus based on the rating of the bond issuer and do not change over time. Each one of the remaining four banks relies respectively on the LT, AST, HH, or CDG model. In each quarter, such a bank determines its regulatory capital on the basis of the expected one-year default rate (i.e., the average PD) implied by the adopted model and the foundation IRB approach.

In order to be able to draw sharp conclusions, it is assumed that the optimal capital levels are those calculated by the perfect-foresight bank.⁴¹ At the other extreme is the SA bank. This bank uses an approach that relies on publicly available credit ratings that, in principle, reflect borrowers' *relative* creditworthiness as opposed to probabilities of default.⁴² In this context, a credit-risk model provides value added only if the bank relying on it matches the optimal capital level better than the SA bank.⁴³

Table 5 reports results of this exercise. The first row of the table contains descriptive statistics of the optimal capital levels, set by the perfect-foresight bank. The remaining rows contain the means of the capital levels calculated by the other five banks, as well as the average and maximum absolute discrepancies between these capital levels and the optimal capital.

⁴⁰Under the foundation IRB approach of Basel II, the capital charge for a particular exposure is determined exclusively by an estimate of the associated one-year PD.

⁴¹This stylized exercise makes no distinction between regulatory (or required) capital and economic (or optimal) capital.

⁴²The officially announced objective of credit-rating agencies is to rate firms according to their long-term financial characteristics. This means that, conditional on the latter characteristics, credit ratings should not vary across the business cycle, even though PDs are largely countercyclical. In addition, credit ratings are supposed to distinguish riskier firms from safer firms and, thus, need to provide an ordinal ranking of firms but need not reflect an absolute measure of credit risk. For further discussion on this topic, refer to Amato and Furfine (2004).

⁴³Note that the optimal capital is based on ex post default rates and thus may incorporate more information about PDs than what could possibly be known ex ante. In this sense, model-implied capital cannot be expected to match exactly the optimal one.

Table 5. Capital Measures (averages over time, in percentage points)

[illegible]

Since the models were seen above to differ in their bias vis-à-vis actual default rates, it is not surprising to find that the models also differ in the success with which they match optimal capital. On average over time, the LT model is quite successful across all of the three rating classes considered. In the BB rating class, e.g., the LT bank experiences a mean discrepancy vis-à-vis the optimal capital of 0.1 percentage points and a mean absolute discrepancy of 1.9 percentage points. These discrepancies amount to 0.9 percent and 17 percent, respectively, of the optimal capital's range, which equals 11.4 percentage points. In comparison, the AST and CDG models perform more poorly, with the banks that adopt these models experiencing mean absolute discrepancies equal to 25 percent and 19 percent, respectively, of the range of optimal capital in the BB rating class. Finally, the HH bank exhibits intermediate performance.

At the same time, again in line with differences between model-based expectations and actual default rates (recall figure 1), all models lead to substantial discrepancies vis-à-vis the level of optimal capital at specific points in time. The LT model, e.g., generates maximum absolute errors that equal 45 percent to 102 percent of the optimal capital ranges, depending on the rating class. Under the other models, these errors tend to be even larger.

Finally, the performance of the SA bank is mixed. The SA bank matches quite closely the optimal capital in the B rating class, where it performs as well as or better than the banks that adopted credit-risk models. However, the performance of the SA bank deteriorates significantly as the credit rating improves to BB and then to BBB. This result reflects the fact that the standardized approach assigns the same regulatory capital (8 percent per unit of exposure) to both BBB- and BB-rated exposures. This results in conservative requirements that overshoot somewhat the optimal regulatory capital even for the riskier category. In the context of a BBB rating, however, the average overshooting is substantial: at 5.3 percentage points, it amounts to 83 percent of the optimal capital's range.

9. Model-Based PDs and Turning Points in the Outlook of Credit Risk

Since the analysis in sections 6–8 examined forecast errors through the prism of intertemporal aggregates, it shed little light on the

models' explanatory power with respect to turning points in the outlook of credit risk. Such explanatory power is of natural interest to both market participants and supervisors, not least because, in the context of structural credit-risk models, it can be easily traced to borrower characteristics that drive the evolution of PD forecasts over time.⁴⁴

The evaluation of the explanatory power of credit-risk models relies on time-series regressions (see tables 6–8). Unless stated otherwise, one-year actual default rates are regressed on corresponding model-implied expected default rates that incorporate firm-specific time-varying parameters whose calibration is described in sections 5.1–5.3.⁴⁵ In all regressions, period- t expectations forecast the default rate realized over the year *starting* in t . In addition, several control variables allow for dissecting the models' explanatory power: the lagged default rate (realized over the year ending in t) and macroeconomic indicators (realized *prior* to the year ending in t).

The regressions focus on one rating class at a time and include fifty-four quarterly observations (from 1990:Q1 to 2003:Q2). For the B and BB rating classes, the analysis relies on OLS regression results. In the BBB rating class, a significant number of the observed default rates equal 0, which prompts the use of a censored-regression specification based on the Tobit model.⁴⁶ In each case, the estimated standard errors account for serial correlation in the residual.⁴⁷ Finally,

⁴⁴Background analysis (the results of which are available from the author upon request) reveals that intertemporal movements of model-based expectations of default rates are driven primarily by changes in leverage, asset volatility, and coupon rates and to a much smaller extent by the drift in the asset process or the time to maturity (in the LT model only).

⁴⁵Given that the available data spans less than thirteen years, the regression analysis carried out in this section abstracts from PDs and observed default rates at the five-year horizon.

⁴⁶Twenty-two of the fifty-four default rates in the BBB rating class equal 0. To account for this, observed default rates are assumed to be “censored” at a low positive value (specifically, 0.03 percent). Under the censored-regression model, the reported adjusted R^2 reflects the goodness of fit vis-à-vis the latent dependent variable, which is a linear function of the regressors.

⁴⁷Since the time series are at the quarterly frequency and the credit-risk horizon is one year, the regression errors should be expected to be serially correlated. To account for this, the p -values are based on Newey-West robust covariance matrices (for the regressions pertaining to BB- and B-rated firms) or on Huber-White robust covariance matrices (for the regressions pertaining to BBB-rated firms).

Table 8. B-Rated Firms

	Dependent Variable: Default Rate							
	Models Only			Adding Macro Explanatory Variables				
	1	2	3	4	5	6	7	8
Constant	-0.00 (0.87)	0.02 (0.03)	0.00 (0.86)	-0.00 (0.76)	0.00 (0.59)	-0.01 (0.39)	0.00 (0.96)	0.02 (0.15)
Default Rate (4-Quarter Lag)	0.18 (0.06)	0.61 (0.00)	0.50 (0.00)	0.21 (0.07)	0.14 (0.30)	0.30 (0.01)	0.27 (0.02)	0.33 (0.00)
LT PDs	0.72 (0.00)			0.62 (0.03)	0.61 (0.01)	0.61 (0.03)	0.60 (0.02)	0.51 (0.01)
AST PDs		-0.03 (0.83)		-0.03 (0.81)	0.07 (0.54)	0.09 (0.58)	0.17 (0.21)	0.21 (0.03)
HH PDs			0.47 (0.01)	0.13 (0.24)	-0.03 (0.84)	0.03 (0.84)	-0.06 (0.64)	0.03 (0.85)
Credit/GDP Gap (16-Quarter Lag)					0.45 (0.16)			
Asset-Price Gap (12-Quarter Lag)						0.27 (0.02)		
GDP Gap (4-Quarter Lag)							1.04 (0.00)	
Term Spread (2-Quarter Lag)								-1.69 (0.00)
Adjusted R^2	63%	34%	52%	63%	65%	66%	71%	75%

Note: Based on one-year actual default rates and cross-sectional averages of one-year theoretical PDs. The sample period is from 1990:Q1 to 2003:Q2 (fifty-four quarterly observations). The p -values (in parentheses) are based on Newey-West robust covariance matrices (three lags). Entries in bold indicate coefficients that are statistically significant at the 10 percent level.

in light of section 6.3, strong multicollinearity across regressors is avoided by focusing on the default-rate expectations implied by only one of the exogenous-default models: the HH model.

9.1 “Models-Only” Regressions

The first four columns in tables 6–8 report the estimated coefficients from regressions that incorporate default-rate expectations implied by the LT, AST, and/or HH models but abstract from macroeconomic indicators. The first three of these regressions focus on one of the models at a time, whereas the fourth one represents a “horse race” among the three models.

The results of the single-model regressions reveal that the structural credit-risk models do exhibit explanatory power vis-à-vis default rates. That said, the size of this explanatory power changes substantially both across models and rating classes. The HH model is the only one that attains statistically significant positive coefficients in all three rating classes. This model accounts for roughly half of the intertemporal variability of default rates in the B and BB rating classes but for only one-fifth of this variability in the BBB rating class. In comparison, the LT model performs somewhat better for B-rated firms, slightly worse for BB-rated firms, and has no explanatory power for BBB-rated firms. The worst performer in the single-model regressions is the AST model, which exhibits significant explanatory power only in the BB-rating class.

Importantly, the different models appear to complement each other in explaining the evolution of default rates over time. This is seen in the context of BBB- and BB-rated firms, where using the expected default rates of several models improves the fit of the regressions.⁴⁸ The improvement is greater in the context of the BB rating class, where the AST and HH PDs complement each other in capturing the three phases of credit risk in the sample: (i) the spike in default rates in 1990–91, (ii) their subsequent drop until 1998,

⁴⁸In the BB rating class, the predictions of the LT model are not combined with those of the other two models in order to avoid multicollinearity issues. In particular, the expected default rates delivered by the LT and HH models have a correlation coefficient of 77 percent, which renders parameter estimates quite noisy. Including the LT instead of the HH model in column 4 of table 7 reduces the goodness-of-fit measure to 51 percent.

and (iii) their moderate pickup thereafter. On the one hand, the HH PDs are consistently close to observed default rates but miss their relative levels in phases (i) and (iii); on the other hand, the AST PDs underpredict over most of the sample but exhibit a global peak in the early 1990s, just like observed default rates (recall figure 1).

By contrast, using expected default rates from several models does not improve the fit of the regression in the B rating class. In this rating class, the LT model exhibits strong explanatory power on its own—leading to an adjusted R^2 of 63 percent—and dominates the explanatory power of the other models.

Finally, the above findings are put into perspective by examining how the regression results change when the models are calibrated to the typical (i.e., average) parameters of each firm. As revealed by figure 4, such a calibration dampens the time variability of default-rate expectations, which should be expected to impair the capacity of the models to account for turning points in credit risk. This is confirmed by table 9, which reports goodness-of-fit measures that are lower than their counterparts in tables 6–8 (columns 1–4). The only exception is a slight improvement of the already good fit delivered by the LT model in the B rating class.

9.2 *Macroeconomic Control Variables*

As seen in figure 1, the model-based expected default rates often miss or erroneously predict important changes in actual default rates. In light of the firm-level focus of the models, this finding suggests that they might be ignoring important marketwide determinants of credit risk. In order to investigate the validity of this hypothesis, this subsection incorporates macroeconomic variables in the regression analysis.

The choice of macro variables is motivated by the extant literature that has identified predictors of turning points in the credit and business cycles. In particular, Borio and Lowe (2002) discover that positive deviations of the credit-to-GDP ratio and real asset prices from their respective trends reflect the buildup of financial market imbalances and forecast well banking-system distress years down the road. To the extent that such distress is associated with a deterioration of banks' lending portfolios, the two financial-side variables would be useful in the context of the present paper, as they would

help predict changes in observed default rates. In addition, in order to account for real-side imbalances that might impact lending criteria, the analysis herein considers the deviation of real GDP from its potential level as a harbinger of future changes in credit risk. Finally, the Treasury term spread, which helps predict changes in real economic activity according to Estrella and Hardouvelis (1991) and Smets and Tsatsaronis (1997), is also incorporated in the regressions. The term spread should help predict default rates insofar as the business cycle affects firms' willingness or capacity to service their debt.

With the above motivation, the asset-price, credit-to-GDP, and GDP gaps and the Treasury term spread are used together with model-based PDs as explanatory variables of default rates.⁴⁹ The results, reported in columns 5–8 in tables 6–8, shed light on two related questions: (i) Could the forecast errors of theoretical PDs be attributed to factors related to the business or credit cycles? (ii) Is the explanatory power of theoretical PDs robust to controlling for variables that capture marketwide phenomena?

Focusing on the two financial-side variables, the results suggest an affirmative answer to the first question. The credit-to-GDP gap and/or the real asset-price gap enter the regressions with statistically significant coefficients and improve the fit, especially for the BBB and BB rating classes.⁵⁰ In the context of the BBB rating class, e.g., the adjusted R^2 reaches 50 percent, which is 28 percentage points higher than the highest R^2 in the models-only regressions. Moreover, in line with the conclusions of Borio and Lowe (2002), positive shocks to the financial-side variables indicate a buildup of financial vulnerabilities and tend to increase default rates three to five years in the future.⁵¹

⁴⁹In order to avoid overfitting, each macro explanatory variable is used separately.

⁵⁰The lag of a macro explanatory variable is chosen to minimize the Akaike information criterion. The results are robust to changing the adopted lags by one or two quarters in either direction.

⁵¹The similarities between the results of the present exercise and the findings of Borio and Lowe (2002) notwithstanding, one should keep in mind that the latter paper builds a nonlinear indicator of banking-system distress. The length of the sample size employed herein limits the analysis to the linear case.

Examination of the real-side macro variables also suggests that the forecast errors of the structural credit-risk models can be attributed, at least partially, to factors related to the business cycle. The GDP gap and the term spread improve substantially the goodness of fit of the regressions in the BBB and B rating classes. For example, the introduction of these variables in the context of BBB-rated firms raises the adjusted R^2 by 52 and 44 percentage points, respectively. Positive shocks to the GDP gap indicate an overheating of the economy, which tends to be associated with an increase of default rates one to two years into the future. In turn, negative shocks to the term spread are associated with higher default rates of BBB- and B-rated firms within the following two years. This is consistent with findings of Estrella and Hardouvelis (1991) and Smets and Tsatsaronis (1997) that a tightening of the term spread heralds recessions up to two years in the future.

As regards the second question posed above, the explanatory power of the structural credit-risk models is largely robust to controlling for macro variables. This is seen both in the B rating class, where the magnitude and significance of the LT coefficient is maintained across all regressions, and in the BB rating class, where the same is true for the AST and, to a lesser extent, HH coefficients. In comparison, the findings are quite different in the BBB rating class. Since the goodness-of-fit measures of the models-only regressions are quite low in this rating class, it is not surprising that the explanatory power of macro variables clearly dominates that of the models.

9.3 Out-of-Sample Analysis

The above regression results convey two messages regarding the in-sample performance of structural credit-risk models. First, the models help account for the intertemporal evolution of default rates. Second, models-only accounts of default rates are substantially improved upon on the basis of macroeconomic variables. In order to investigate the robustness of these two messages, this section considers out-of-sample forecasts that rely on model-based

expected default rates and macroeconomic variables available in real time.⁵²

Specifically, there are four exercises that (i) perform out-of-sample forecasts of the default rates realized over the last eight quarters of the time period under study (i.e., from 2001:Q3 to 2003:Q2) and (ii) differ in the amount of information incorporated in these forecasts. For the first two exercises, models-only specifications include expected default rates from several models (see tables 6–8, regression 4) and forecast the default rate in quarter t on the basis of regression coefficients estimated from data covering 1990:Q1 to 2001:Q2 (i.e., without updates) or 1990:Q1 to $t - 1$ (i.e., with updates). The regression coefficients underpinning the other two exercises rely on data spanning the same time periods and incorporating, in addition to model-based expected default rates, the macro explanatory variable that has led to the best in-sample fit of actual default rates for the rating class in focus.⁵³ Table 10 reports results of the four out-of-sample exercises alongside corresponding results obtained in sample.

The out-of-sample performance of the models-only regressions is consistent with the degree to which the structural credit-risk models account for default rates in sample. In the context of BB- and B-rated classes, the *with*-updates exercise reveals out-of-sample mean absolute errors (MAEs) that are, respectively, 8 and 25 basis points higher than their in-sample counterparts. This worsening of the fit is small relative to the mean of the actual default rates between 2001:Q3 and 2003:Q2 and, thus, translates into only slight increases of mean average percentage errors (MAPEs). In comparison, the worsening of the fit is larger in the BBB rating class, where the in-sample performance of the models is poorest.⁵⁴

⁵²The interpretation of this exercise abstracts from the fact that some of the parameters underpinning model-based expected default rates (i.e., the risk-free rate of return and the recovery rate) are averaged over the entire sample period.

⁵³In light of tables 6–8, this macroeconomic variable is the GDP gap (for BBB-rated firms), the gap in the credit-to-GDP ratio (BB-rated firms), and the term spread (B-rated firms).

⁵⁴The MAEs and MAPEs for the BBB rating class incorporate the censoring specification outlined in footnote 46.

Table 10. Out-of-Sample Performance

BBB-Rated Firms						
Models Only			Adding a Macro Variable ^d			
	Without Updates ^a	With Updates ^b	<i>In-Sample</i> ^c	Without Updates ^a	With Updates ^b	<i>In-Sample</i> ^c
MAE MAPE	0.0062 105%	0.0074 125%	0.0055 92%	0.0046 78%	0.0035 59%	0.0016 26%
BB-Rated Firms						
Models Only			Adding a Macro Variable ^d			
	Without Updates ^a	With Updates ^b	<i>In-Sample</i> ^c	Without Updates ^a	With Updates ^b	<i>In-Sample</i> ^c
MAE MAPE	0.0065 54%	0.0057 47%	0.0049 41%	0.0032 26%	0.0033 27%	0.003 25%
B-Rated Firms						
Models Only			Adding a Macro Variable ^d			
	Without Updates ^a	With Updates ^b	<i>In-Sample</i> ^c	Without Updates ^a	With Updates ^b	<i>In-Sample</i> ^c
MAE MAPE	0.027 65%	0.0255 61%	0.023 56%	0.0108 26%	0.0096 23%	0.0072 17%

Note: Summary statistics of the difference between the actual default rates and the forecast or fitted default rates from 2001:Q3 to 2003:Q2. MAE = mean absolute error; MAPE = mean absolute percentage error. The percentage errors are calculated with respect to the mean default rate realized from 2001:Q3 to 2003:Q2.

^aBased on regression coefficients (regression 4, tables 6-8) that incorporate data from 1990:Q1 to 2001:Q2.

^bBased on regression coefficients (regression 4, tables 6-8) that incorporate data from 1990:Q1 to t , where t is a quarter from 2001:Q2 to 2003:Q3 and occurs one quarter before the forecast default rate.

^cBased on regression coefficients (regression 4, tables 6-8) that incorporate the full sample period (from 1990:Q1 to 2003:Q2).

^dThe added macro variable is the best performer according to regressions 5 to 8 in tables 6-8: GDP gap (BBB-rated firms), credit-GDP gap (BB-rated firms), and term spread (B-rated firms).

Again in line with the in-sample results reported in section 9.2, adding a macroeconomic explanatory variable improves substantially the out-of-sample forecasts of default rates. This is seen most clearly in the fact that, irrespective of the rating class, the MAEs and MAPEs of the *without*-updates forecasts incorporating a macro variable are lower than those of the best (i.e., in-sample) forecasts based on models-only regressions. Of course, out-of-sample forecasts that incorporate a macro variable and quarterly updates of the regression parameters tend to perform even better. These with-updates forecasts fit actual default rates of B- and, especially, BB-rated firms virtually as well as the corresponding in-sample forecasts. In comparison, out-of-sample forecasts incorporating a macro variable perform more poorly in the BBB rating class.

10. Conclusion

This paper has used firm-level data in order to evaluate the degree to which five structural credit-risk models account for the level and intertemporal evolution of actual default rates. In contrast to previous analyses, which have addressed similar objectives on the basis of aggregate data, the paper finds that PDs implied by the models tend to match the level of actual default rates. In addition, the models explain a substantial portion of the variability of default rates over time. That said, the model-based forecasts of default rates are substantially improved upon by the introduction of macroeconomic variables, which are not modeled explicitly but reflect the business and credit cycles.

Data limitations have imposed several restrictions on the analysis, which are to be kept in mind when interpreting its conclusions. The mitigation of such limitations would, by itself, lead to useful future contributions to the subject matter of this paper. On the one hand, longer data series, incorporating several credit cycles, would shed further light on the robustness of the models' success in accounting for upturns and downturns in credit risk. On the other hand, larger cross-sections would increase significantly one's confidence in the theoretical forecasts of default rates at specific points in time. For example, larger cross-sections in the historical default

data would allow for a reliable estimate of a time-varying recovery rate, which would improve considerably the calibration of the models.

References

- Amato, J. D., and C. H. Furfine. 2004. "Are Credit Ratings Procyclical?" *Journal of Banking and Finance* 28 (11): 2641–77.
- Anderson, R. W., S. Sundaresan, and P. Tychon. 1996. "Strategic Analysis of Contingent Claims." *European Economic Review* 40 (3–5): 871–81.
- Andrade, G., and S. Kaplan. 1998. "How Costly Is Financial (Not Economic) Distress? Evidence from Highly Leveraged Transactions that Become Distressed." *Journal of Finance* 53 (5): 1443–93.
- Basel Committee on Banking Supervision. 2004. "International Convergence of Capital Measurement and Capital Standards: A Revised Framework." Bank for International Settlements.
- . 2006. "Studies on Credit Risk Concentration: An Overview of the Issues and a Synopsis of the Results from the Research Task Force Project." BCBS Working Paper No. 15.
- Bhandari, L. C. 1988. "Debt/Equity Ratio and Expected Common Stock Returns: Empirical Evidence." *Journal of Finance* 43 (2): 507–28.
- Bharath, S. T., and T. Shumway. 2004. "Forecasting Default with the KMV-Merton Model." Working Paper.
- Bohn, J., N. Arora, and I. Korablev. 2005. "Power and Level Validation of the EDFTM Credit Measure in the U.S. Market." Moody's KMV White Paper.
- Borio, C., and P. Lowe. 2002. "Assessing the Risk of Banking Crises." *Quarterly Review* (Bank for International Settlements) (December): 43–54.
- Collin-Dufresne, P., and R. Goldstein. 2001. "Do Credit Spreads Reflect Stationary Leverage Ratios?" *Journal of Finance* 56 (5): 1929–57.
- Das, S. R., D. Duffie, N. Kapadia, and L. Saita. 2007. "Common Failings: How Corporate Defaults Are Correlated." *Journal of Finance* 62 (1): 93–117.

- Duffie, D., L. Saita, and K. Wang. 2007. "Multi-Period Corporate Default Prediction with Stochastic Covariates." *Journal of Financial Economics* 83 (3): 635–65.
- Eom, Y., J. Helwege, and J.-Z. Huang. 2004. "Structural Models of Corporate Bond Pricing: An Empirical Analysis." *The Review of Financial Studies* 17 (2): 499–544.
- Estrella, A., and G. A. Hardouvelis. 1991. "The Term Structure as a Predictor of Real Economic Activity." *Journal of Finance* 46 (2): 555–76.
- Feldhütter, P., and D. Lando. 2004. "The Credit Risk Component in Corporate Spreads and Swap Spreads." Working Paper.
- Hillegeist, S. A., E. K. Keating, D. P. Cram, and K. G. Lundstedt. 2004. "Assessing the Probability of Bankruptcy." *Review of Accounting Studies* 9 (1): 5–34.
- Huang, J.-Z., and M. Huang. 2003. "How Much of the Corporate-Treasury Yield Spread Is Due to Credit Risk?" Working Paper.
- Kealhofer, S., and M. Kurbat. 2001. "The Default Prediction Power of the Merton Approach, Relative to Debt Ratings and Accounting Variables." Moody's KMV White Paper.
- Kurbat, M., and I. Korablev. 2002. "Methodology for Testing the Level of the EDFTM Credit Measure." Moody's KMV Technical Report No. 020729.
- Leland, H. E. 1994. "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure." *Journal of Finance* 49 (4): 1213–52.
- . 2004. "Predictions of Default Probabilities in Structural Models of Debt." *Journal of Investment Management* 2 (2): 1–16.
- Leland, H. E., and K. B. Toft. 1996. "Optimal Capital Structure, Endogenous Bankruptcy, and the Term Structure of Credit Spreads." *Journal of Finance* 51 (3): 987–1019.
- Longstaff, F. A., and E. S. Schwartz. 1995. "A Simple Approach to Valuing Risky Fixed and Floating Rate Debt." *Journal of Finance* 50 (3): 789–820.
- Merton, R. C. 1974. "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates." *Journal of Finance* 29 (2): 449–70.
- Moody's Investors Service. 1998. "Historical Default Rates of Corporate Bond Issuers, 1920–1998." Special Comment.

- Smets, F., and K. Tsatsaronis. 1997. "Why Does the Yield Curve Predict Economic Activity? Dissecting the Evidence for Germany and the United States." BIS Working Paper No. 49.
- Stein, R. M. 2000. "Evidence on the Incompleteness of Merton-Type Structural Models for Default Prediction." Moody's KMV White Paper.
- Tarashev, N., and K. Tsatsaronis. 2006. "Risk Premia across Asset Markets: Information from Option Prices." *Quarterly Review* (Bank for International Settlements) (March): 93–103.
- van Deventer, D. R., L. Li, and X. Wang. 2006. "Advanced Credit Model Performance Testing to Meet Basel Requirements: How Times Have Changed." In *The Basel Handbook: A Guide for Financial Practitioners*, ed. M. Ong. Risk Books.

Inflation Targets as Focal Points*

Maria Demertzis^a and Nicola Viegi^b

^aDe Nederlandsche Bank and University of Amsterdam

^bUniversity of Cape Town and ERSA

In a world characterized by noisy information and conflicting signals, no central bank is able to affect private-sector expectations at all times. In order to evaluate the effectiveness of any central bank communication strategy, it is important to know what private agents rely on when they form expectations. We model monetary policy as an information game in which individuals form their expectations based on all the information that is available to them (public and private) and are, therefore, subject to the noise that characterizes that information. Individual agents also know that inflation is ultimately affected both by central bank policies and by the average expectation formed by all agents. The way individuals interpret these two components to form their expectations is explained in the context of a higher-order expectations setup and is central to our argument. We then apply Bacharach's (1993) variable-universe methodology to provide a framework for assessing everyone's interpretations. Therefore, our contribution is, first, to describe monetary policy as an information game in which interpretations matter and, second, to provide a way of solving for these interpretations. We show that a monetary policy regime that has explicit quantitative objectives may provide individuals with better anchors for coordinating their expectations. However, that is only true either if no great shocks are anticipated or if all other public information is very unclear, leaving the inflation target as the only clear piece of information available. We derive the conditions under which this is true.

JEL Codes: C71, C78, E52.

*Views expressed are our own and do not necessarily reflect those of De Nederlandsche Bank (DNB). Without wishing to implicate, we would like to

1. Introduction

Modern monetary policy theory emphasizes the central role of private-sector expectations in determining policy outcomes. Recent empirical evidence by Paloviita and Virén (2005) demonstrates this for inflation in the euro area. It is thus widely acknowledged that the success of maintaining a stable monetary environment depends crucially on the ability of the policy regime to control inflation expectations (Blinder et al. 2001). Evidence of that is shown by Orphanides and Williams (2005) in their analysis of U.S. monetary policy history, where they argue that monetary policy failures are connected with changes in public sentiment about the future state of the economy. In other words, policy mistakes alone are not enough to produce long-term negative effects on monetary stability.

The practice of monetary policy in the past ten to fifteen years has thus concentrated on providing institutional setups that provide an explicit information platform for expectations to be formed. The main objectives of such institutional setups are to attain

- credible institutions, mainly through independence and the pursuit of price stability as the principal objective;
- clear policy frameworks captured by well-defined intermediate policy objectives and procedures; and, finally,
- transparent policymaking implemented through publication and distribution of the information set used in the decision-making process (inflation forecasts, modeling strategies, and well-defined assumptions) and a clear demonstration of accountability (publication of minutes, regular appearance

thank Wilko Bolt, Peter van Els, Charles Goodhart, Marco Hoeberichts, Lex Hoogduin, Andrew Hughes Hallett, Kai Leitemo, Eric Schaling, Hyun Shin, an anonymous referee, and seminar participants at DNB, the Riksbank, the University of Amsterdam, the 2005 “Money, Macro and Finance” Conference, and the Economic Society of South Africa for comments and suggestions. The usual disclaimer applies. Corresponding author: Demertzis: Economics and Research, De Nederlandsche Bank, P.O. Box 98, 1000 AB, Amsterdam, the Netherlands. E-mail: m.demertzis@dnb.nl. Tel: +31 (20) 524 2016. Fax: +31 (20) 524 2506. Web site: www1.fee.uva.nl/toe/content/people/demertzis.shtm. Viegi: School of Economics, University of Cape Town, Private Bag, Rondebosch, 7700, South Africa. E-mail: nicola.viegi@uct.ac.za. Tel: +27 21 650 2763. Fax: +27 21 650 2854.

in front of parliamentary committees, and regular press conferences).

Practically every monetary policy authority nowadays defines its policies according to these criteria, emphasizing one or another aspect depending on preferences. The euro-area setup, e.g., has emphasized the importance of building and sustaining credibility and independence from governments as an instrument toward low expected inflation. In the U.S. experience, credibility and independence but also flexibility in following multiple objectives have helped achieve a stable monetary environment. Alternatively, inflation targeting—as implemented first by the Reserve Bank of New Zealand, then the Bank of England, and increasingly by more banks around the world—is understood to provide clear and immediate objectives for monetary policy. Inflation-targeting practitioners argue that the main advantage of an explicit numerical inflation target is its ability to provide a focal point for private-sector expectations. As Mervyn King (2002, 4) has claimed for the UK case, inflation expectations have indeed been anchored to the preannounced target. The ability of explicit quantitative targets to tie down expectations is also confirmed by the empirical analysis of Mishkin and Schmidt-Hebbel (2001, 2007), Johnson (2002), and more recently by Fatás, Mihov, and Rose (2007) and Levin, Natalucci, and Piger (2004).¹ However, conventional monetary policy models (Svensson 1999, 2003; Woodford 2003) allow for no difference in the way inflation targeting is modeled by comparison with other regimes.² There is thus no explicit analysis of the way the provision of a specific numerical target may constitute a better anchor for private-sector expectations. What is the mechanism that makes the inflation target a better anchor, and which conditions are required for this to be true? In our view, to be able to answer that, we need a mechanism of expectations formation more

¹See also Leiderman and Svensson (1995) and Bernanke et al. (1999) for earlier accounts of experiences with inflation targeting.

²Kuttner (2004) also alludes to this fact. The benefits of inflation targeting as a coordination device have been discussed by Hughes Hallett and Viegi (2002) but in the context of two authorities, the policies of which might have strong “spillovers.”

complex than the standard full-information rational-expectations paradigm.

The recent model put forward by Morris and Shin (2002a, 2002b) (and used in Amato, Morris, and Shin 2002 and Amato and Shin 2003) renders itself, first, to identifying how private agents form expectations based on private and public information available to them and, second, to showing how policymakers affect these expectations by providing greater or lesser information. This setup shows that in forming these expectations, private agents care not only about their own views but also about other people's expectations as a means to confirming their own beliefs. In fact, Phelps (1983, 35) noted the importance of general perceptions when he said that "in order to reduce the price level (in relation to the accustomed trend), it is not sufficient that the central bank (CB) persuade each agent to reduce his private expectation of the money supply (in relation to the past trend) by the warranted amount. The prevalence of this expectation must be public knowledge—an accepted fact." And as the "beauty contest" element³ (based on Keynes 1936) is understood to play a greater role in expectations forming, signals provided by public institutions can conceivably become tantamount to coordination devices. This therefore implies that monetary policy can be viewed as a coordination game between the central bank and the private sector but also as a matching game between members of the private sector themselves. Due to the latter, public information then acquires a dual role—"of conveying fundamentals information, as well as serving as a focal point for beliefs" (Morris and Shin 2002b; henceforth, MS 2002). The question that arises following this argument is, which monetary policy regimes provide better signals, and in what way do these signals constitute focal points?⁴ The aim of this paper is to formalize the widely believed

³The British comedy team Monty Python epitomizes the essence of higher-order expectations in the following: "I think that all good, right thinking people in this country are sick and tired of being told that all good, right thinking people in this country are fed up with being told that all good, right thinking people in this country are fed up with being sick and tired. I'm certainly not, and I'm sick and tired of being told that I am."

⁴See Sugden (1995) for a theory on focal points. Bryan and Palmqvist (2006) also apply the term "focal points" to monetary policy, in a similar fashion to ours.

but little-analyzed benefits of inflation targeting in coordinating private individuals' expectations and the conditions necessary for this to be achieved.

The theory on coordination games provides valuable insight into the way that such games are resolved. For example, it is often observed that in matching games, players coordinate much more frequently than they would if they were just randomizing (Casajus 2000). Indeed, according to Wilson and Rhodes (1997), it is to the benefit of all actors to avoid the conflict that escalates as solutions are delayed. To achieve that, players rely heavily on salient features when deciding on their actions. And salience, in this context, can be a "social custom or convention, namely a mode of behavior that finds automatic acceptance" (Dixit and Skeath 1999). A salient item then is "one that stands out from the rest by its uniqueness in some conspicuous respect" (Bacharach 1993). Furthermore, Wilson and Rhodes (1997) argue that all that is required for such salience to be achieved is a signal from somebody who can be recognized as the "leader" in the game. A commonly accepted leader, in a clearly defined leader-follower(s) game, can thus provide such a focal point. In our setup, the central bank arguably acquires such a leadership role by the sheer extent of its contribution to the final inflation outcome. Compared with other potential providers of "focal points" (say, newspapers, the government, or even big financial players in the system), the central bank is better positioned to fulfill such a role. Providing a numerical target, we argue, grants private agents the choice between two actions, in contrast to a regime absent of a numerical target. These actions are either to internalize the announced target and treat it as an additional piece of public information or, driven by their incentive to coordinate, to ignore all other information and fix their expectations at that level. Given the intentions of the central bank, the latter is naturally the preferred option, but only if everybody else pursues this option as well. In deciding between these two options, therefore, the agent remains uncertain as to how others view the target. Trusting the target for oneself is neither necessary nor sufficient for pursuing it; what is actually pivotal to one's choice is how every individual understands others' interpretation of the target and the central bank's ability to achieve it. What the individual needs, therefore, is a framework that will help, first, identify what the options are and, second,

evaluate how these options are understood by all others. To this end, we employ the *variable-universe games* approach put forward by Bacharach (1993), which describes how players evaluate their options in the context of what everyone else might believe about those options.

Our contribution is to merge the MS (2002) framework with that of Bacharach (1993) in order to provide the individual with a framework to identify the best action. We find that numerical targets are effective coordinators of expectations when no great shocks are anticipated, or when all other public information available to individuals is relatively imprecise.

The paper is organized as follows. With the aid of a standard monetary policy model, section 2 describes how monetary policy can be seen as an information game based on the work by MS (2002). We then explain how the provision of numerical targets increases the options available to individuals in section 3 and describe Bacharach's (1993) approach to interpret these options. Section 4 then merges the two models and discusses the conditions for which the individual forms expectations according to the target provided by the central bank. Section 5 concludes.

2. Monetary Policy as an Information Game

The central bank has a standard loss function in which it chooses the rate of inflation x to minimize the distance from the inflation objective set x^T and close the output gap y ,

$$L_{CB}|\xi = \frac{1}{2}\mathbb{E}[(x - x^T)^2 + y^2], \quad (1)$$

subject to a standard Lucas supply function, $y = x - x^e + \xi$, where ξ is a supply shock with zero mean and constant variance, σ_ξ^2 . Note that any central bank will have an objective x^T irrespective of whether it communicates it to the public clearly, or even at all. We assume for simplification that the central bank's instrument is x . Optimization of (1) implies that

$$x|\xi = \frac{x^T}{2} + \frac{x^e}{2} - \frac{\xi}{2}, \quad (2)$$

where x is now the ex post inflation outcome conditional on the shock ξ and x^e is private-sector expectations about the relevant rate of inflation. Representation (2) is of a structural form in the sense that expectations are not replaced (Leitemo 2005).⁵ In a typical commitment game, where the central bank communicates its target x^T and commits to it, expectations formed by all individuals collectively are equal to the CB's objectives, $x^e = x^T$, and the ex post outcome is

$$x|\xi = x^T - \frac{\xi}{2} \quad (3)$$

$$\mathbb{E}(x) = x^T. \quad (4)$$

The objective of this paper, however, is to depart from the assumption that expectations are always equal to the objectives of the central bank and analyze how individuals go about interpreting the information that is available to them when forming expectations. Every individual i will be forming an expectation of inflation x_i such that the collective outcome, which is the one that is relevant to ex post inflation, is $x^e = \int_0^1 x_j dj$ (for a continuum of agents). The timing of the game assumed has the central bank deciding what its objectives are first; shocks occur next, then private agents form expectations based on information available about these shocks and policy objectives, and, finally, the central bank reacts to them.

⁵Note that (2) is specific to the underlying Lucas supply function assumed but demonstrates that the outcome will be a function of both the policy the central bank pursues and what the private sector anticipates. Analogously, had the model been of the standard neo-Keynesian type based on Clarida, Galí, and Gertler (1999),

$$\begin{aligned} x_t &= \beta E_t x_{t+1} + k y_t + \varepsilon_t \\ y_t &= E_t y_{t+1} - \gamma(i_t - E_t x_{t+1}) + \eta_t, \end{aligned}$$

then the structural representation of the ex post inflation outcome would be

$$x_t = \frac{k^2}{1+k^2} x^T + \frac{1}{1+k^2} E_t x_{t+1} + \frac{\varepsilon_t}{1+k^2}.$$

Our point is to show that the ex post outcome is a function of both the central bank's objective and the private sector's expectations.

2.1 *The Formation of Expectations*

We start by arguing that while the central bank itself may be clear about what its objectives are, it is not always possible to assume that private individuals form expectations that are consistent with these objectives. It becomes important then to examine the information that is available to the private sector and how the private sector uses that information to form expectations. Typically, every individual forms expectations based on two information sets—namely, what is publicly available and therefore common to everyone, and what is available to them privately. Furthermore, every individual is aware of the fact that the ex post outcome of inflation x will be determined by (2); in other words, it will be affected equally (given the model we assume) by the policy the central bank pursues to attain its objectives and by the average of expectations formed by the public.

However, as the individual is interested in predicting the ex post level of inflation correctly (in order, e.g., to base her wage negotiations on⁶), she needs to interpret both components of (2) based on the information she has. Her objectives are captured by a standard expected disutility,⁷

$$u_i(\mathbf{x}^e, x^T) \equiv \frac{1}{2} \mathbb{E}_i(x_i - x)^2. \quad (5)$$

x_i is individual i 's expectation of what inflation will be, and x is again the ex post inflation outcome. Note that subscript i in the expectations operator indicates that the individual will be seeking to minimize her expected disutility, given her own perceptions. We use $\mathbf{x}^e$ to refer to the expectations profile over all agents. The

⁶We assume that the individual consumer sets a price variable (individual wage) and supply elastically to the amount of labor demanded. This is just a narrative trick: the argument would work equally well in a setup as in Lucas's island model, in which individuals set the price of a good in an imperfect knowledge setup.

⁷See Canzoneri (1985). Unlike MS (2002), this objective function does not have the "beauty contest" term in it, but as we will show in (6), it delivers the same type of "action," in the sense that the individual's expectation relies on both her evaluation of the "state" and the average "action." The "beauty contest" term is therefore not strictly necessary for higher-order expectations to become relevant (Angeletos and Pavan 2007 make a similar point).

individual decides her action x_i based on the first-order condition of (5). This is

$$\arg \min_{x_i} u_i(\mathbf{x}^e, x^T) = \mathbb{E}_i(x)$$

and, from (2),

$$\begin{aligned} x_i &= \mathbb{E}_i(x) \\ x_i &= \mathbb{E}_i\left(\frac{x^T}{2} + \frac{x^e}{2} - \frac{\xi}{2}\right) \\ x_i &= \frac{1}{2}\mathbb{E}_i(x^T - \xi) + \frac{1}{2}\mathbb{E}_i(x^e). \end{aligned} \quad (6)$$

Based on (6), the optimal action for individual i is thus a function of three things: (i) the objectives of the central bank and hence the policy it will pursue, (ii) the shock that has occurred, and, finally, (iii) the average expectation formed by all individuals. Moreover, in forming expectations x_i , individual i needs to evaluate these three things, captured here by the expectations operator, subscript i . It follows that if $x_i = x_j \ \forall j$, then $x_i = x^e$ and individuals' expectations are matched. However, although desirable, coordination between agents at any level of inflation is not sufficient; the optimal outcome occurs when agents coordinate at the objective pursued by the central bank, x^T . Coordination at any other expectation rate still leaves agents away from the level of inflation that the central bank aims to achieve. We will argue later that knowledge of the central bank's objective is necessary but not sufficient for coordination at it. Following MS (2002), we argue that information used by the agents is available in the form of a public signal that is common to all and a private signal that is specific to each agent in the economy. Individuals therefore observe p and z_i , where

$$\text{Public signal: } p = (x^T - \xi) + \eta \quad (7)$$

$$\text{Private signal: } z_i = (x^T - \xi) + \varepsilon_i. \quad (8)$$

Both η and ε_i have a zero mean and constant variance, σ_η^2 and σ_ε^2 , respectively. Furthermore, the two error terms are independent of x and of each other, such that $\mathbb{E}(\varepsilon_i \varepsilon_j) = 0$ for $i \neq j$. The clarity of

public information is not under the full control of the central bank but is affected by a combination of the central bank's information strategy, the general market information available, and noise. Based on these two types of signals, Morris and Shin show that agent i 's action then is

$$\begin{aligned} x_i &= \frac{2\alpha p + \beta z_i}{2\alpha + \beta} \\ &= x^T - \xi + \frac{2\alpha\eta + \beta\varepsilon_i}{2\alpha + \beta}, \end{aligned} \quad (9)$$

where $\alpha = \frac{1}{\sigma_\eta^2}$ and $\beta = \frac{1}{\sigma_\varepsilon^2}$, the level of precision for the two information sets, respectively.

DEFINITION 1. *We call (9) the “MS action.”*

We assume homogenous agents and calculate expectations across all agents as follows:

$$x^e = \int_0^1 x_j dj = x^T - \xi + \frac{2\alpha\eta}{2\alpha + \beta}. \quad (10)$$

Equation (10) shows that the average expectation across all agents will be distorted by the (lack of) precision of the two signals as well as the preference for the “beauty term,” here equal to $\frac{1}{2}$ (from (6)).

3. The Role of Inflation Targets

As argued earlier, the central bank's objectives are captured by (1). In reflecting central bank practices, “price stability” is at the heart of its communication. We differentiate, however, between banks in terms of the way they communicate this objective. We identify a central bank that communicates a quantitative target x^T as an “inflation targeter.”⁸ On the other hand, a central bank that does not

⁸The benefits or risks of communication have been discussed extensively in the literature. Rudebusch and Williams (2007) carry out an exercise that is similar in spirit to ours by showing that when private agents have imperfect information, central bank communication (in their case, in the form of publishing interest rate projections) can help manage financial-market expectations.

make an attempt to explicitly quantify the term “price stability” and communicate it to the public is, in turn, a “non-inflation targeter.” We investigate next what this implies for the choices available to the individual.⁹

Our analysis so far shows that if the central bank does not explicitly quantify its objective of price stability, then the best individual i can do is (9), which relies on what public and private information she has about the central bank’s inflation objective and the supply shock. Based on (5), we can then calculate individual i ’s disutility:

$$u_i(x_i, x^e) = \frac{\alpha + \beta}{(2\alpha + \beta)^2}. \quad (11)$$

We contrast this now with the case where the central bank *does* announce what it means by the term “price stability.” The individual now receives an extra signal in addition to (7) and (8):

$$\text{Central bank signal: } h = x^T. \quad (12)$$

As a result, every individual is faced with the choice between two alternatives for her action a_i : either weigh all information available to her and thus follow the strategy suggested by Morris and Shin (x_i) *or*, driven by her desire to coordinate, fix her expectations at the inflation target announced by the central bank (x^T). In modeling terms, these alternatives are summarized in (13) and (14), respectively:

$$a_i : x_i = x^T - \xi + \frac{2\alpha\eta + \beta\varepsilon_i}{2\alpha + \beta}, \quad (13)$$

$$a_i : x^T. \quad (14)$$

By analogy, the same applies for the collective action $\bar{a}$ ¹⁰ such that, respectively for the two alternatives, the MS action leads to

⁹We appreciate that this is a crude distinction and not one that necessarily captures all the differences between inflation and non-inflation targeters. On the other hand, inflation-targeting countries have also experienced the regime in very different ways. Our objective here is simply to capture that characteristic that applies to all inflation targeters—namely, of providing a quantitative target—and that differentiates them from non-inflation targeters.

¹⁰In our terminology, individual i is thus confronted with the “average” or “collective” action. We will use these terms interchangeably.

an average inflation expectation of x^e , whereas following the central bank's target leads to x^T . In modeling terms, these amount to (15) and (16), respectively:

$$\bar{a} : x^e = x^T - \xi + \frac{2\alpha\eta}{2\alpha + \beta}, \quad (15)$$

$$\bar{a} : x^T. \quad (16)$$

In the absence of an announced inflation objective, average inflation expectations will be formed based on (15). However, when the central bank announces what its objective is, private agents then have the option of forming expectations according to either (15) or (16). We can see, therefore, that the communication strategy of the central bank with respect to the inflation objective affects the formation of expectations by increasing the number of options available to the agents. The rationale behind such a central bank announcement is the knowledge that agents need to form a view about the average expectation of inflation (as shown in (6)) and are therefore looking to identify salient points for inflation. Inflation targeting, in the way we define it here, aspires to appeal to the agents' coordination motive. It is interesting to note that while an agent that forms expectations according to the target resolves the coordination motive, by bypassing any information noise (α and β), it does so at the cost of throwing away information about the supply shock ξ . The optimal situation for the individual would be $\eta = 0$ —i.e., no information shocks—as the individual's action would then account for both the target and the shock. But in the presence of information imprecision, there are instances where the benefits of coordination more than outweigh the loss of information on shocks. The decisive factor in the individual's decision is the credibility of the announced target, as we will show later in the paper.

For a continuum of agents, the inflation outcome is affected by the collective action $\bar{a}$ only. Ex post inflation (from (2)) is thus

$$\begin{aligned} x|_{\xi, x^e} &= \frac{x^T}{2} + \frac{x^e}{2} - \frac{\xi}{2} \\ &= x^T - \xi + \frac{\alpha\eta}{2\alpha + \beta} \end{aligned} \quad (17)$$

when $\bar{a} : x^e$ and

$$\begin{aligned} x|_{\xi, x^T} &= \frac{x^T}{2} + \frac{x^T}{2} - \frac{\xi}{2} \\ &= x^T - \frac{\xi}{2} \end{aligned} \quad (18)$$

when $\bar{a} : x^T$. Based on (5), we turn next to the general form of individual i 's disutility, affected by both her own action (a_i) and the average action ($\bar{a}$), i.e.,

$$u_i(a_i, \bar{a}) \equiv \mathbb{E}_i(a_i - x)^2. \quad (19)$$

Four different disutilities are then possible outcomes for individual i :

- Choice 1: $a_i = x_i, \bar{a} = x^e$

$$u_i(x_i, x^e) = \frac{\alpha + \beta}{(2\alpha + \beta)^2} \quad (20)$$

- Choice 2: $a_i = x^T, \bar{a} = x^e$

$$u_i(x^T, x^e) = \sigma_\xi^2 + \frac{\alpha}{(2\alpha + \beta)^2} \quad (21)$$

- Choice 3: $a_i = x_i, \bar{a} = x^T$

$$u_i(x_i, x^T) = \frac{1}{4}\sigma_\xi^2 + \frac{4\alpha + \beta}{(2\alpha + \beta)^2} \quad (22)$$

- Choice 4: $a_i = x^T, \bar{a} = x^T$

$$u_i(x^T, x^T) = \frac{1}{4}\sigma_\xi^2 \quad (23)$$

The very provision of an inflation target now increases the number of options available to the individual and thus the number of potential outcomes. Table 1 summarizes the pure-form strategies available to individual i (and the disutility outcomes associated

Table 1. Individual i 's Disutility in Normal Form

NIT		IT		
$a_i \backslash \bar{a}$	x^e	$a_i \backslash \bar{a}$	x^e	x^T
x_i	$\frac{\alpha+\beta}{(2\alpha+\beta)^2}$	x_i	$\frac{\alpha+\beta}{(2\alpha+\beta)^2}$	$\frac{1}{4}\sigma_\xi^2 + \frac{4\alpha+\beta}{(2\alpha+\beta)^2}$
		x^T	$\sigma_\xi^2 + \frac{\alpha}{(2\alpha+\beta)^2}$	$\frac{1}{4}\sigma_\xi^2$

with them) for the two alternative monetary regimes of non-inflation targeting (NIT) and inflation targeting (IT).

Thus, if the central bank does not announce a value for its inflation objective, the best individual i can do is apply the MS action. However, if the central bank does announce a quantitative objective for inflation, then which response is optimal for individual i depends crucially on the relative size of the variance of the supply shock, σ_ξ^2 . In particular, we argue the following.

PROPOSITION 1. *Applying the inflation target becomes individual i 's “dominant” strategy for a relatively low variance of the supply shock; otherwise, expectations forming becomes a “matching” game.*

Proof. Table 1 shows that for any given level of precision for public and private information, adopting the inflation target x^T becomes a dominant strategy for individual i if the variance of the supply shock is below a given threshold, i.e.,

$$\sigma_\xi^2 < \frac{\beta}{(2\alpha + \beta)^2}. \quad (24)$$

As we assume homogenous agents (and, by implication, that everyone adopts their dominant strategy), the game has *one* Nash equilibrium—namely, $E_1 \equiv (x^T, x^T)$. The inflation outcome is equal to (18), and individual i ends up facing disutility equal to $\frac{1}{4}\sigma_\xi^2$. However, if condition (24) is not satisfied, i.e., $\sigma_\xi^2 \geq \frac{\beta}{(2\alpha+\beta)^2}$, then individual i 's optimal response in pure-form strategies requires “matching” the average action. In other words, $a_i = x_i$ is the best response to $\bar{a} = x^e$, and $a_i = x^T$ is the best response to $\bar{a} = x^T$. The game now

has *two* Nash equilibria in pure form—namely, $E_2 \equiv (x^T, x^T)$ and $E_3 \equiv (x_i, x^e)$.¹¹

In the latter case, individual i is unable to differentiate between the two Nash equilibria in pure form. To decide on an action, she needs to compare the *expected* (rather than *actual*) disutility that each of them delivers; in other words, $\mathbb{E}[u_i(a_i, \bar{a})]$, where $a_i \in \{x_i, x^T\}$ and $\bar{a} \in \{x^e, x^T\}$. This in turn depends on the probability with which the average action takes one of its own two respective values.

DEFINITION 2. Define $q \equiv \Pr(\bar{a} = x^T)$ and $(1 - q) \equiv \Pr(\bar{a} = x^e)$, $q \in [0, 1]$.

Expected disutility for individual i for each of her two actions is then

$$\mathbb{E}\{u_i[x_i, (x^e \text{ or } x^T)]\} = (1 - q)u_i(x_i, x^e) + qu_i(x_i, x^T) \quad (25)$$

$$\mathbb{E}\{u_i[x^T, (x^e \text{ or } x^T)]\} = (1 - q)u_i(x^T, x^e) + qu_i(x^T, x^T). \quad (26)$$

The difficulty with this, however, is that probability q is unknown to individual i and she is thus unable to evaluate which action to take based on (25) and (26). She needs to identify an alternative framework that provides her with sufficient information on which she can then base her choice. We will argue that rather than evaluate q , it will suffice for individual i simply to know how the target is perceived (and not necessarily applied) on average. Otherwise put, it will suffice for the individual to know the proportion of people that consider the target credible. We believe that in the context of monetary policy, this information is publicly available and easier for the individual to deduce. Based on that, we will show that she can then identify under which conditions following the target provided by the central bank is to her advantage. We describe the framework for doing that next, based on Bacharach's (1993) variable-universe games.

¹¹The condition is such that if equality holds, then the individual is indifferent between the two.

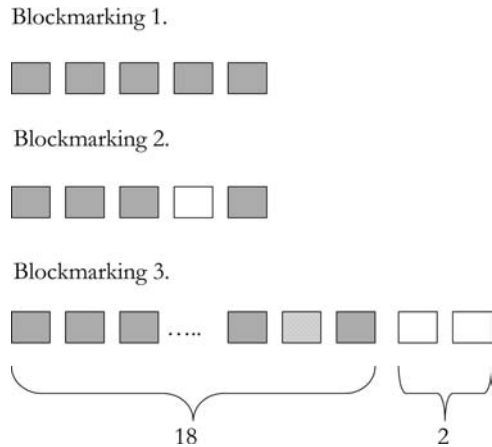
3.1 *A Framework for Interpreting Expectations*

Bacharach’s (1993) variable-universe games framework helps describe how players evaluate their strategies to identify salient points when forming expectations in matching games.¹² The novelty of this approach is that it allows explicitly for differences in perceptions, which then helps players choose rationally between alternative outcomes. The framework provided shows that in matching games, the players’ incentive to coordinate induces them to look for salient points. However, as salience is subject to personal interpretation, the existence of such features is not necessarily uniquely defined. We describe the approach first, then apply it explicitly to our question in section 4.

3.1.1 *The Game of Blockmarking*

The game of blockmarking is played in the following way. Two players (player 1 and player 2) are shown a number of wooden blocks, and each has to secretly mark one. If both players mark the same block, they receive an identical pecuniary prize; otherwise, they receive nothing. There are then three variants of the game, summarized in figure 1.

Figure 1. The Game of Blockmarking



¹²This framework was used and extended by Janssen (2001).

In Blockmarking 1 (B1), the players are given five blocks identical in size, color, shape, and material. In Blockmarking 2 (B2), the same game is repeated, except now one of the five blocks is of a different color (white). In Blockmarking 3 (B3), players are now given twenty blocks, eighteen of which are grey and two of which are white. Furthermore, closer inspection of the blocks allows players to see that the grain of the wood in just one of the grey blocks is wavy (block 17 on the graph). B3 can thus be described either in terms of color (C) or in terms of the grain of the wood (G). As the game is of a matching nature, it is to the players' interest to look for salient features that help achieve tacit coordination. In example B1, there is no clear way of differentiating between the blocks, so one is inclined to simply choose at random. In example B2, however, the difference in color allows players to distinguish between the blocks in such a way that it is always wise to go for the one that is white. The *unique instantiation* of the white block thus provides the two players with a focal point. Similarly, in B3, if color is the distinguishing feature that occurs to the players, they are then inclined to mark one of the white blocks, even though such action does not automatically lead to coordination. However, if players have managed to see that not only color but also the grain of the wood differentiates the blocks, uniqueness can again be guaranteed. The difficulty now, however, is that the grain pattern of the wood is not necessarily identifiable (*conspicuous*) by (*to*) all players.

How does player 1 proceed with forming her choice in B3? Bacharach argues that there is no one mechanical method of answering this question. However, one "can make use of certain general considerations about how agents understand situations and understand each other to understand them, considerations built on the notion of 'normality'" (Bacharach 1993, 262). This notion of "normality" provides a common platform that allows the individual to begin to evaluate her options. It is "normal" therefore to assume that all players distinguish the blocks according to color, but that not all players see the difference in grain. This is the most basic assumption about B3. Next to that, Bacharach also adds that while seeing the grain is by no means certain, there is a definite probability v with which an arbitrary normal person will have seen it. If the player is normal herself, then it is natural to assume that she knows of this probability v and believes it to hold for her (equally

normal) partner.¹³ Bacharach's contribution then is to demonstrate that knowledge of this prior v is sufficient for player 1 to base her choice on. His analysis shows that if this likelihood is big, then it is to her advantage to mark the grey block with the wavy pattern; otherwise, she is better off marking one of the white blocks and facing, at most, a 50 percent chance of matching her partner's choice. In the appendix we describe how this is derived analytically and show that for the action of selecting the wavy-grained block to be optimal in B3, the condition that needs to be satisfied is that $v > \frac{1}{2}$. In general, the balance of reasons favors marking the block with the wavy grain only if v is a large enough number by comparison with $\frac{1}{m}$, where m is the number of white blocks. Bacharach concludes that the relative rarity of the white blocks, captured here by $\frac{1}{2}$, is pulling against the conspicuousness v of the grain pattern, as the less rare the white blocks are (the bigger m is), the more likely the player is to mark the wavy-patterned grey block (provided she has seen it).

4. Variable Universes and Inflation Targets

How does the analogy carry over to monetary policy? As argued earlier, our analysis aims to show how announcing a quantitative target for monetary policy alters the number of options available to the individual. The first observation is that if the central bank does *not* announce an explicit quantitative target, then as far as the individual is concerned, the best she can do is implement the MS action, equivalent to marking a white block at random. The option of following the target is therefore not given to her, and her position in the game is as though she has not seen the difference in grain.

When the central bank does, however, announce an explicit quantitative target, then we interpret Bacharach's notion of "normality" to imply the following three things:¹⁴

¹³Naturally, this probability is not relevant to her decision, or indeed known to her, if she has not seen the difference in grain herself.

¹⁴Note that our description of what the provision of an explicit quantitative target implies draws heavily from the common notion that IT constitutes "a simple yardstick by which to judge policy. Given lexicographic preferences over

- (i) First, the option of following the target is available to everyone. We believe this to be an adequate representation of the fact that inflation targets are widely announced and are therefore seen and understood by everyone. In that respect, it is reasonable to assume that any random individual would understand that the options shown in table 1 under IT are now available to them.
- (ii) Second, not necessarily everyone believes that the target is credible. In other words, not everyone believes that the central bank can actually achieve the target, or even that the notion of “price stability” and the value x^T announced are necessarily equivalent. This is, in our view, an important reflection of the fact that revealing a quantitative objective is not automatically sufficient for attaining it.
- (iii) Finally, and in the spirit of the game of blockmarking, while credibility is by no means certain, there is nevertheless a common perception as to how credible such a target might be. In other words, everyone knows and understands how the target is perceived by others on average, or, alternatively, everyone knows what proportion of people consider the target credible. This frequency, or likelihood, is now our parameter v and, based on Bacharach’s notion of normality, constitutes common knowledge. We would want to investigate what levels of credibility are required before the individual decides to form expectations according to the target (equivalent to the $\frac{1}{m}$ condition above).

Based on the above, we provide the following definition.

DEFINITION 3. *Define v as follows: the proportion of people that believe that the target is credible.*

We can now rely on this likelihood v to evaluate individual i ’s expected disutility $\mathbb{E}[u_i(a_i, \bar{a})]$ of following action a_i for a given collective action $\bar{a}$. This has the advantage that the priors used are now known to her. However, as v is not the actual probability with

inflation and other goals, an inflation target range, and a fixed horizon, inflation targeting becomes very easy to monitor” (Faust and Henderson 2004, 127).

which the collective action is equal to the target ($v \neq q$), we need to proceed in two steps in order to provide a comprehensive framework that accounts for all possible outcomes.¹⁵

CASE 1. First, we consider the case when the collective choice is always the default MS action. The target is therefore never applied, irrespective of whether it is deemed credible. This may happen either because the average perception is that the central bank's target is not credible or because the average individual does not believe that others think the target is credible. The default MS action is then applied on average, and inflation expectations equal x^e . Are there then circumstances in which individual i is still better off following the target? Her expected disutility of following either of her two options is then

$$\begin{aligned}\mathbb{E}[u_i(x_i, x^e)] &= (1 - v)u_i(x_i, x^e) + vu_i(x_i, x^e) \\ &= u_i(x_i, x^e) \\ \mathbb{E}[u_i(x^T, x^e)] &= (1 - v)u_i(x^T, x^e) + vu_i(x^T, x^e) \\ &= u_i(x^T, x^e).\end{aligned}$$

From table 1, these equal, respectively,

$$\begin{aligned}\mathbb{E}[u_i(x_i, x^e)] &= \frac{\alpha + \beta}{(2\alpha + \beta)^2} \\ \mathbb{E}[u_i(x^T, x^e)] &= \sigma_\xi^2 + \frac{\alpha}{(2\alpha + \beta)^2}.\end{aligned}$$

It follows that

$$\begin{aligned}\mathbb{E}[u_i(x^T, x^e)] &< \mathbb{E}[u_i(x_i, x^e)] \quad \text{iff} \\ \sigma_\xi^2 &< \frac{\beta}{(2\alpha + \beta)^2}.\end{aligned}\tag{27}$$

Expression (27) illustrates the condition for which individual i would effectively ignore the fact that everyone else applies the MS

¹⁵Actually, there are three steps, but the third is trivial, as we show later in footnote 16.

default action and still follow the target, even though she is aware she would be the only one. But this is only rational if following the target constitutes her dominant strategy. It is therefore no surprise that (27) is identical to the condition shown earlier under dominance, i.e., (24).

CASE 2. Second, we relax the stringency of the first assumption and assume now that the proportion of people who believe that the target is credible actually apply it. When, then, should individual i choose the target? This is a somewhat less restrictive case than the one above, and expected disutility for player i of pursuing either of her two options is now

$$\begin{aligned}\mathbb{E}\{u_i[x_i, (x^e \text{ or } x^T)]\} &= (1 - v)u_i(x_i, x^e) + vu_i(x_i, x^T), \\ \mathbb{E}\{u_i[x^T, (x^e \text{ or } x^T)]\} &= (1 - v)u_i(x^T, x^e) + vu_i(x^T, x^T)\end{aligned}$$

and, from table 1,

$$\begin{aligned}\mathbb{E}\{u_i[x_i, (x^e \text{ or } x^T)]\} &= (1 - v)\frac{\alpha + \beta}{(2\alpha + \beta)^2} + v\left[\frac{1}{4}\sigma_\xi^2 + \frac{4\alpha + \beta}{(2\alpha + \beta)^2}\right], \\ \mathbb{E}\{u_i[x^T, (x^e \text{ or } x^T)]\} &= (1 - v)\left[\sigma_\xi^2 + \frac{\alpha}{(2\alpha + \beta)^2}\right] + v * \frac{1}{4}\sigma_\xi^2.\end{aligned}$$

It follows that

$$\begin{aligned}\mathbb{E}\{u_i[x^T, (x^e \text{ or } x^T)]\} &< \mathbb{E}\{u_i[x_i, (x^e \text{ or } x^T)]\}, \quad \text{iff} \\ \sigma_\xi^2 &< \frac{\beta + v4\alpha}{(1 - v)(2\alpha + \beta)^2}.\end{aligned}\tag{28}$$

Again, the size of the shock needs to be below a given ratio before the individual applies the target. However, as the condition is less restrictive than before, (since $\frac{\beta}{(2\alpha + \beta)^2} < \frac{\beta + v4\alpha}{(1 - v)(2\alpha + \beta)^2}$), it is also the sufficient condition for individual i to follow that target. Interestingly, we can rewrite this condition in terms of v :

$$\frac{(2\alpha + \beta)^2\sigma_\xi^2 - \beta}{4\alpha + (2\alpha + \beta)^2\sigma_\xi^2} < v.\tag{29}$$

This now shows that the sufficient condition for individual i to follow the target is when a suitably large proportion (in relation to the

shocks and the quality of information available) of the public believe the target to be credible.¹⁶

Having identified this condition, we discuss briefly the role of v in the monetary policy game.¹⁷ Credibility of central bank actions is an aspect that is highly (if not only) dependent on the institution's track record. In other words, we believe that people revise their beliefs about the central bank based on what they observe (past performance). At time t , v_t therefore depends on whether actual inflation in the previous period was equal (or close, by a well-predefined amount) to the target ($x_{t-1} \simeq x^T$). If so, the private sector will increase v_t for the current period (by comparison with v_{t-1}); if not, it will reduce it and hence penalize the bank. This implies automatically that while credibility is endogenous to the system, it is also *predetermined* at the time the decision is made, in the sense that v_t is the "stock" of credibility that the central bank can no longer affect. The central bank's actions at time t will therefore only affect v_{t+1} . The implicit timing of the static game at any period is therefore as follows:

$$v_t|_{x_{t-1}} \longrightarrow \xi_t \longrightarrow a_{i,t} = \begin{cases} x^e & \text{if } v_t < \text{condition} \\ x^T & \text{if } v_t \geq \text{condition} \end{cases} \longrightarrow x|_{\xi_t, \bar{a}_t} \longrightarrow x_t.$$

¹⁶Note that we also need to examine a third case in which the average action is now assumed to always be the target. One could envisage a situation (admittedly unlikely) in which everyone thinks for themselves that the target is not credible but choose nevertheless to follow it, because they believe they are alone in thinking that. It is trivial then to show that following the target is individual i 's preferred strategy.

$$\begin{aligned} E\{u_i[x_i, x^T]\} &= \frac{1}{4}\sigma_\xi^2 + \frac{4\alpha + \beta}{(2\alpha + \beta)^2} \\ E\{u_i[x^T, x^T]\} &= \frac{1}{4}\sigma_\xi^2 \\ E\{u_i[x^T, x^T]\} &< E\{u_i[x_i, x^T]\} \end{aligned}$$

¹⁷The formation of expectations described here is similar in spirit to what Bomfin and Rudebusch (2000) have in that, depending on the level of credibility of the central bank, the private sector shifts its expectations between the target and an alternative—in their case, past inflation; in our case, MS action. Also, while we have the choice between two values (hence discrete), they have the choice within a range of values (hence continuous).

Credibility is assessed first, based on last period's performance; the supply shock materializes; agents then form expectations accordingly; and the central bank reacts, leading to an inflationary outcome for the period equal to x_t . How this credibility is updated, however, is beyond the scope of this paper. Our efforts above were confined to establishing the condition required for an individual to want to switch from x^e to x^T when forming expectations.

4.1 *Inflation Targeting and Individual i 's Options:* *A Discussion*

We argue that the decision by a central bank to explicitly quantify its objective of price stability aims at providing a platform for individuals to coordinate their expectations. From the point of view of the individual, however, this becomes interesting only if certain conditions are met. Our analysis has shown that what matters is, first, the economic conditions that prevail (namely, supply shocks) and, second, the way this quantitative target is perceived by the public.

Figure 2 summarizes the conditions for the variance of the supply shock σ_ξ^2 that need to hold for the inflation target to be individual i 's best option. We discuss these next in greater detail.

4.1.1 *Inflation Targeting as a Dominant Strategy*

We have shown that if $\sigma_\xi^2 < \Psi$ (from figure 2), then following the target becomes the individual's dominant pure-form strategy. For this to be true, it is important that the supply shock is relatively small in relation to the given ratio of the public and private information precision. Figure 3 now plots the value of this ratio Ψ (vertical axis)

Figure 2. Individual i 's Options

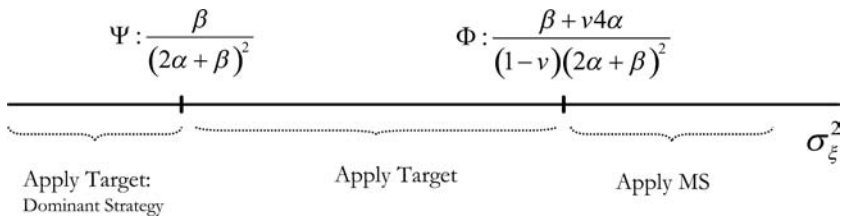
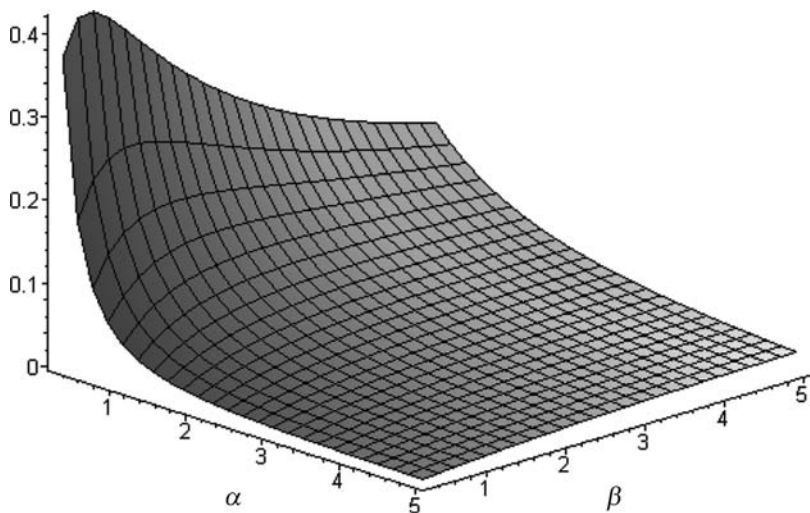


Figure 3. Inflation Targeting as a Dominant Strategy—"Ψ" against α and β



against values of private (α) and public (β) information precision. As the condition is satisfied for values below the surface shown, we can see that supply shocks are required to be very small for this to hold. Indeed, if the economy is hit by large shocks instead, then it is unlikely that this condition is met, implying that the provision of a target does not help agents coordinate at the level intended by the central bank. We find this intuitively appealing because it evaluates the effectiveness of the target publicized within the context of the economic conditions in which it is applied.

Moreover, figure 3 also shows that if public information is very imprecise (i.e., α is low), then the condition is easier to satisfy, *ceteris paribus*, and the provision of an inflation target is indeed helpful. This implies that numerical targets become substitutes for imprecise public information; in the absence of concrete alternative information, the provision of one clear inflation target becomes the only unequivocal piece of information available to all. To carry the analogy to the blockmarking game, if white blocks become more and more numerous, then having seen the block with the wavy grain, one is more likely to mark it.

4.1.2 Expectations Formation as a Matching Game

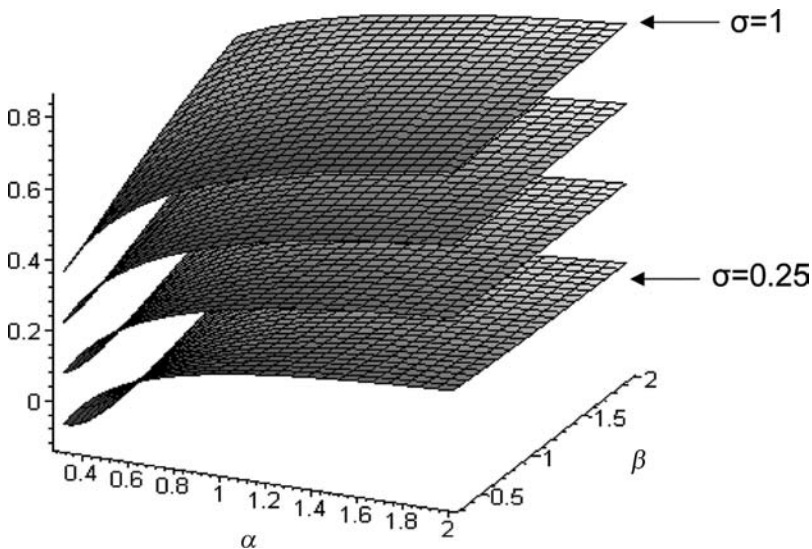
However, if (24) is not satisfied (i.e., $\sigma_\xi^2 > \Psi$), then individual i 's optimal action in pure-form strategies requires “matching” the average action. We differentiate between two cases:

- (i) If $\Psi < \sigma_\xi^2 < \Phi$, then she is still better off fixing her expectations at the central bank's announced target. We show that this condition can be interpreted in terms of v , which suggests that the individual needs to know how credible the target is considered on average, i.e.,

$$K = \frac{(2\alpha + \beta)^2 \sigma_\xi^2 - \beta}{4\alpha + (2\alpha + \beta)^2 \sigma_\xi^2} < v. \quad (30)$$

If this level of confidence is greater than the ratio shown, then it is still to her advantage to form expectations according to the target. Figure 4 now describes how easy it is for this condition to be met in relation to α , β , and the supply shock

Figure 4. The Role of Interpretations—“ K ” against α and β



σ_ξ . On the vertical axis we plot the left-hand side of (30) in terms of α and β and for four different values of the supply shock ($0 < \sigma_\xi < 1$).

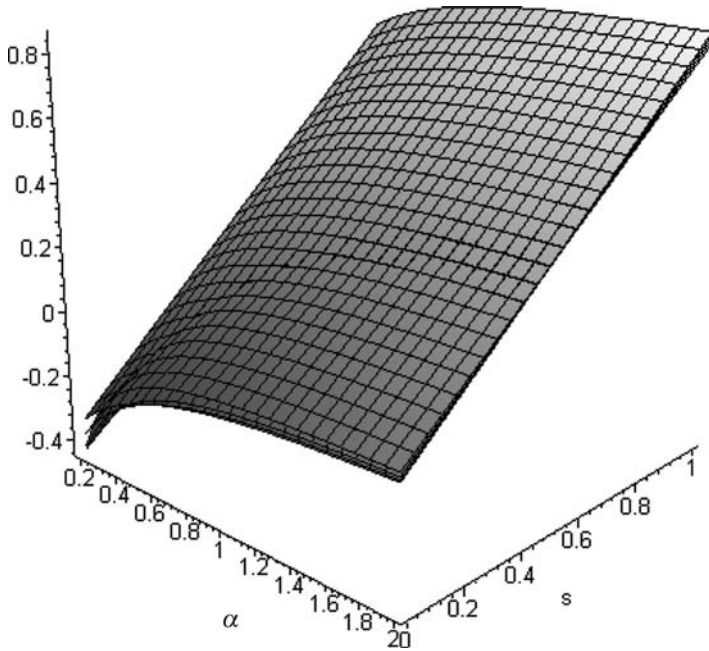
The condition is now satisfied for values that fall above each of the surfaces shown; thus, the higher the plateau, the more difficult it is to meet it. There are two interesting features that arise from figure 4. First, it is the case that as the variance of the shock increases, then (30) becomes more difficult to satisfy. In other words, if larger shocks are expected, then individual i needs an ever-greater degree of confidence that the target is perceived to be credible (i.e., higher v) before she chooses it herself. This is consistent with what is mentioned above—namely, that in the presence of large shocks, inflation targeting is less convincing in its role as a coordinator of expectations. Second, as public information suffers from lack of clarity (i.e., α is low), the provision of a clear and unique quantitative signal helps relax the stringency of the condition. We see this as each of the layers in figure 4 tilt downward for small values of α , thus making the condition easier to satisfy.

- (ii) If, however, $\sigma_\xi^2 > \Phi$, then the individual is expected to do better by applying the MS default action.

In all cases above, the role of private information is de-emphasized in that it does not impose a constraint on either (27) or (30). This is demonstrated in figure 5 for the latter condition, which is now plotted against α and σ_ξ and then for different values of β . We see this as the plates are very closely overlapping, thus indicating that changing private information does not help alter the stringency of the condition.

5. Conclusions

Any private individual forms expectations of inflation based on information that is available to her. Our paper concentrates on the way the central bank's communication strategy might affect these expectations. We begin our analysis by arguing that it is not always possible for a monetary policy authority to assume that it can affect private expectations in such a way that they will match its own intentions. Private individuals rely on information that is available

Figure 5. The Role of Private Information

to them publicly (and thus common to everyone) as well as information that might be unique to them individually. Monetary policy becomes an information game, then, in which private individuals base their decision on a combination of all information available, corrected for their respective degree of (or lack of) precision. As the level of expectations affects the final outcome of inflation, the private sector needs to deduce the central bank's objective, its ability to achieve it, and what everyone else believes. We apply the MS (2002) model to demonstrate that the latter point implies that coordinated expectations are preferable, although not necessarily the guarantee of optimal outcomes. We then use Bacharach's variable-universe approach to demonstrate exactly how people interpret the options available to them given the actions of all other players in the game. Our contribution has therefore been to merge the two models and provide a comprehensive framework for individuals to enumerate their options and thus form expectations.

Based on this, we find that a central bank that announces a quantitative target may, *ceteris paribus*, benefit from helping private-sector expectations coordinate at the level of its objectives. We describe the conditions for which this happens and discover that inflation targeting does indeed achieve coordination—first, when the supply shocks expected are small (in other words, the economy is stable) and, second, when public information fails in all other respects to provide the private sector with clear signals as to what the relevant level of inflation is going to be. It is in this sense that we argue that inflation targets are substitutes for poor public information. Naturally, as we show above, it is not sufficient for any individual to view this quantitative signal; the individual also needs to know that the signal is, on the whole, perceived to be credible. If this holds, then following the signal will constitute her preferred strategy and the central bank will have effectively provided a focal point.

Appendix. Variable-Universe Games (B3)

Bacharach provides a thorough proof to B3 in the appendix to his paper, but the essence of the game faced by the two players individually can be summarized as follows. In solving B3, player 1 is effectively faced with two alternative actions: $M\tilde{h}$, mark a white block at random, or Mw , mark the grey block with the wavy grain. Furthermore, as explained in the main text, the crucial point in this analysis is the likelihood with which player 2 has noticed the grain. Player 1 is thus left with the following two choices when forming her views about player 2. Either she believes that her opponent has seen the grain (and the assumption of normality implies that he will have with probability v), or she does not believe that he has seen the grain (with probability $1 - v$). It is reasonable to assume that if player 2 has indeed noticed the grain, then he will mark that block with some nonzero probability. However, if he has not noticed the grain, then he can never mark a block accordingly. From player 1's perspective, therefore, her expected utility from choosing one of her two actions is the following.

DEFINITION 4. *Both players have an identical set of feasible strategies, $R^+ = \{C, G\}$, and possible actions, $A = \{M\tilde{h}, Mw\}$. Define*

$U_1(x_{1,a(\bullet)}, x_{2,a(\bullet)})$, player 1's utility from following action $x_{1,a(\bullet)}$ and player 2 following action $x_{2,a(\bullet)}$, for $a \in A$, where $a(C) = M\tilde{h}$ and $a(G) = Mw$.

As v is not the probability with which player 2 marks the block with the wavy grain but merely the likelihood with which any "normal" player will have seen it, we need to deal with two cases in order to ensure that our approach accounts for all possible outcomes.

CASE A1. Player 2 always marks a block according to color, either because he has not seen the grain himself or because he believes his partner has not. Then player 1's expected utility is

$$\begin{aligned}\mathbb{E}_1 U(M\tilde{h}, M\tilde{h}) &= (1 - v)U_1(M\tilde{h}, M\tilde{h}) + vU_1(M\tilde{h}, M\tilde{h}) \\ \mathbb{E}_1 U(Mw, M\tilde{h}) &= (1 - v)U_1(Mw, M\tilde{h}) + vU_1(Mw, M\tilde{h}).\end{aligned}$$

Next, we normalize $U(x_1 = x_2) = 1$ and calculate the expected utilities:

$$\begin{aligned}\mathbb{E}_1 U(M\tilde{h}, M\tilde{h}) &= (1 - v)\frac{1}{2}U(x_1 = x_2) + v\frac{1}{2}U(x_1 = x_2) = \frac{1}{2} \\ \mathbb{E}_1 U(Mw, M\tilde{h}) &= (1 - v) * 0 + v * 0 = 0.\end{aligned}$$

This implies that $\mathbb{E}_1 U(M\tilde{h}, M\tilde{h}) > \mathbb{E}_1 U(Mw, M\tilde{h})$, and therefore player 1 has an incentive to match her partner's action by also marking a white block at random.

CASE A2. Player 2 now marks a block based on the grain when he has noticed it. Otherwise, he marks a block according to color. Expected utility for player 1 is now

$$\begin{aligned}\mathbb{E}_1 U[M\tilde{h}, (M\tilde{h} \text{ or } Mw)] &= (1 - v)U_1(M\tilde{h}, M\tilde{h}) + vU_1(M\tilde{h}, Mw) \\ \mathbb{E}_1 U[Mw, (M\tilde{h} \text{ or } Mw)] &= (1 - v)U_1(Mw, M\tilde{h}) + vU_1(Mw, Mw)\end{aligned}$$

and, therefore,

$$\begin{aligned}\mathbb{E}_1 U[M\tilde{h}, (M\tilde{h} \text{ or } Mw)] &= (1 - v)\frac{1}{2}U(x_1 = x_2) + v * 0 = \frac{1 - v}{2} \\ \mathbb{E}_1 U[Mw, (M\tilde{h} \text{ or } Mw)] &= (1 - v) * 0 + vU(x_1 = x_2) = v.\end{aligned}$$

It follows that

$$\mathbb{E}_1 U[Mw, (M\tilde{h} \text{ or } Mw)] > \mathbb{E}_1 U[M\tilde{h}, (M\tilde{h} \text{ or } Mw)], \quad \text{iff } v > \frac{1}{3}.$$

But between the two cases, the necessary and sufficient condition for player 1 to decide to mark a block according to the grain is

$$\begin{aligned} \mathbb{E}_1 U[Mw, (M\tilde{h} \text{ or } Mw)] > \mathbb{E}_1 U(M\tilde{h}, M\tilde{h}) &\iff \\ v > \frac{1}{2}. \end{aligned} \tag{31}$$

In other words, the balance of reasons favors marking the block with the wavy grain only if v is a large-enough number by comparison with $\frac{1}{m}$, where m is the number of white blocks. Bacharach argues therefore that the relative rarity of the white blocks, captured here by $\frac{1}{2}$, is pulling against the conspicuousness v of the grain pattern, as the less rare the white blocks are (the bigger m is), the more likely the player is to mark the wavy-patterned grey block.

The point that is crucial to Bacharach's analysis is the fact that players have particular ways of perceiving the game, such that the framing of the game (universe) available to them individually is not necessarily available to other players as well (variable). Before deciding on a possible action—for example, marking the block with the wavy grain (provided they have seen it)—the player has to form a view as to how likely her counterpart is to have noticed the grain (as well as the color) as a possible distinguishing feature. Evaluating that is necessary before choosing a strategy, implying that having noticed the block with the wavy grain for oneself is not sufficient reason to mark it. Any player therefore needs to assess whether they believe that what is conspicuous to them is also conspicuous to others. Marking the wavy-grained block is the desirable strategy only if the expected value of doing so is greater than the expected value of marking a white block at random.

References

- Amato, J. D., S. Morris, and H. S. Shin. 2002. "Communication and Monetary Policy." *Oxford Review of Economic Policy* 18 (4): 495–503.

- Amato, J. D., and H. S. Shin. 2003. "Public and Private Information in Monetary Policy Models." BIS Working Paper No. 138 (September).
- Angeletos, G.-M., and A. Pavan. 2007. "Efficient Use of Information and Social Value of Information." *Econometrica* 75 (4): 1103–42.
- Bacharach, M. 1993. "Variable Universe Games." In *Frontiers of Game Theory*, ed. K. Binmore, A. Kirman, and P. Tani, 255–75. Cambridge, MA: MIT Press.
- Bernanke, B. S., T. Laubach, F. S. Mishkin, and A. S. Posen. 1999. *Inflation Targeting: Lessons from International Experience*. Princeton, NJ: Princeton University Press.
- Blinder, A. S., C. Goodhart, P. Hildebrand, D. Lipton, and C. Wyplosz. 2001. *How Do Central Banks Talk?* Geneva Report on the World Economy, No. 3, London: Centre for Economic Policy Research/International Centre for Monetary and Banking Studies.
- Bomfin, A. N., and G. D. Rudebusch. 2000. "Opportunistic and Deliberate Disinflation Under Imperfect Credibility." *Journal of Money, Credit, and Banking* 32 (4): 707–21.
- Bryan, M. F., and S. Palmqvist. 2006. "Testing Near-Rationality Using Detailed Survey Data." Riksbank Research Paper Series, No. 20.
- Canzoneri, M. 1985. "Monetary Policy Games and the Role of Private Information." *American Economic Review* 75 (5): 1056–70.
- Casajus, A. 2000. "Focal Points in Framed Strategic Forms." *Games and Economic Behavior* 32 (2): 263–91.
- Clarida, R., J. Galí, and M. Gertler. 1999. "The Science of Monetary Policy: A New Keynesian Perspective." *Journal of Economic Literature* 37 (4): 1661–1707.
- Dixit, A. K., and S. Skeath. 1999. *Games of Strategy*. W. W. Norton and Company.
- Fatás, A., I. Mihov, and A. K. Rose. 2007. "Quantitative Goals for Monetary Policy." *Journal of Money, Credit, and Banking* 39 (5): 1163–76.
- Faust, J., and D. W. Henderson. 2004. "Is Inflation Targeting Best-Practice Monetary Policy?" *Review* (Federal Reserve Bank of St. Louis) 86 (4): 117–43.

- Hughes Hallett, A., and N. Viegi. 2002. "Inflation Targeting as a Coordination Device." *Open Economies Review* 13 (4): 341–62.
- Janssen, M. C. W. 2001. "Rationalizing Focal Points." *Theory and Decision* 50 (2): 119–48.
- Johnson, D. R. 2002. "The Effect of Inflation Targeting on the Behavior of Expected Inflation: Evidence from an 11 Country Panel." *Journal of Monetary Economics* 49 (8): 1521–38.
- Keynes, J. M. 1936. *The General Theory of Employment Interest and Money*. London: Macmillan.
- King, M. 2002. "The Inflation Target 10 Years On." Speech delivered to the London School of Economics, November 19.
- Kuttner, K. N. 2004. "The Role of Policy Rules in Inflation Targeting." *Review* (Federal Reserve Bank of St. Louis) 86 (4): 89–111.
- Leiderman, L., and L. E. O. Svensson, eds. 1995. *Inflation Targets*. London: Centre for Economic Policy Research.
- Leitemo, K. 2005. "Inflation Targeting Rules: History Dependent or Forward-Looking?" Manuscript, Norwegian School of Management, Oslo.
- Levin, A. T., F. M. Natalucci, and J. M. Piger. 2004. "The Macroeconomic Effects of Inflation Targeting." *Review* (Federal Reserve Bank of St. Louis) 86 (4): 51–80.
- Mishkin, F. S., and K. Schmidt-Hebbel. 2001. "One Decade of Inflation Targeting in the World: What Do We Know and What Do We Need to Know?" In *Inflation Targeting: Design, Performance, Challenges*, ed. N. Loayza and R. Soto, 117–219. Santiago: Central Bank of Chile.
- . 2007. "Does Inflation Targeting Make a Difference?" NBER Working Paper No. 12876 (January).
- Morris, S., and H. S. Shin. 2002a. "Global Games: Theory and Applications." In *Advances in Economics and Econometrics*, the Eighth World Congress, ed. M. Dewatripont, L. Hansen, and S. Turnovsky, 56–114. Cambridge University Press.
- . 2002b. "Social Value of Public Information." *American Economic Review* 92 (5): 1521–34.
- Orphanides, A., and J. C. Williams. 2005. "The Decline of Activist Monetary Policy: Natural Rate Misperceptions, Learning, and Expectations." *Journal of Economic Dynamics and Control* 29 (11): 1927–50.

- Paloviita, M., and M. Virén. 2005. "The Role of Expectations in the Inflation Process in the Euro Area." Bank of Finland Research Discussion Paper No. 6.
- Phelps, E. S. 1983. "The Trouble with 'Rational Expectations' and the Problem of Inflation Stabilization." In *Individual Forecasting and Aggregate Outcomes*, ed. R. Frydman and E. S. Phelps, 31–40. New York: Cambridge University Press.
- Rudebusch, G. D., and J. C. Williams. 2007. "Revealing the Secrets of the Temple: The Value of Publishing Central Bank Interest Rate Projections." Forthcoming in *Asset Prices and Monetary Policy*, ed. J. Y. Campbell. Chicago: University of Chicago Press.
- Sugden, R. 1995. "A Theory of Focal Points." *The Economic Journal* May (105): 533–50.
- Svensson, L. E. O. 1999. "Inflation Targeting as a Monetary Policy Rule." *Journal of Monetary Economics* 43: 607–54.
- . 2003. "What Is Wrong with Taylor Rules? Using Judgment in Monetary Policy through Targeting Rules." *Journal of Economic Literature* 41 (2): 426–77.
- Wilson, R. K., and C. M. Rhodes. 1997. "Leadership and Credibility in N-Person Coordination Games." *Journal of Conflict Resolution* 41 (6): 767–91.
- Woodford, M. 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton, NJ: Princeton University Press.

Transaction Pricing and the Adoption of Electronic Payments: A Cross-Country Comparison*

Wilko Bolt,^a David Humphrey,^b and Roland Uittenbogaard^a

^aDe Nederlandsche Bank

^bFlorida State University

After safety, the efficiency of a nation's payment system is a primary concern of central banks. Since electronic payments are typically cheaper than paper-based or cash payments, pricing these transactions should speed up the shift to electronics. But by how much? Norway explicitly priced point-of-sale and bill-payment transactions and rapidly shifted to electronic payments, while the Netherlands experienced a similar shift without pricing. Controlling for terminal availability and differences between countries, direct pricing accelerated the shift to electronics by about 20 percent. The quid pro quo was the elimination of bank-float revenues.

JEL Codes: D12, G21.

1. Introduction

The average bank cost of an electronic payment is one-third to one-half that of its paper-based equivalent or cash (cf. Humphrey et al. 2006). A merchant's average cost of accepting a bill payment electronically over a giro network or at the point of sale (POS) is also lower (credit cards excepted). Since the resource cost of a country's payment system may account for 1–2 percent of its GDP, it is clear

*The authors would like to thank Hans Brits, Nicole Jonker, Rein Kieviet, Raymond Kleijmeer, Andrew Stone, Carlo Winder, and two anonymous referees for valuable comments and suggestions. The views expressed in this article are those of the authors alone and do not necessarily represent those of De Nederlandsche Bank, the European System of Central Banks, or the Federal Reserve Bank of Philadelphia. Author e-mails: Bolt: w.bolt@dnb.nl; Humphrey: dhumphrey@cob.fsu.edu; Uittenbogaard: r.a.uittenbogaard@dnb.nl.

that shifting from paper to electronic payments can confer social benefits. Importantly, the discounted value of these cost benefits will be larger the more rapidly this shift occurs. Additional effects also exist and are of concern to central banks since the replacement of cash by electronic payments can alter the monetary aggregates (Duca and VanHoose 2004), reduce government seigniorage revenues (Humphrey, Kaloudis, and Øwre 2004), and make tax evasion and illegal transactions more difficult to hide.

There is overwhelming evidence that consumers respond to price incentives but very little evidence on how strong this response may be in the payments area. Although consumers are used to responding to price incentives, they tend not to welcome the opportunity to trade off perceived payment preferences with relative prices when their payment use has commonly been viewed as being “free.” While businesses often pay directly for the payment services they use via explicit transaction fees or compensating balances, consumers have traditionally paid implicitly through lost float or lower (or no) interest on transaction balances. However, consumer surveys indicate that certain implicit costs (e.g., availability, convenience, and security) can affect payment choice (Borzekowski and Kiser 2006; Klee 2006), and models have been developed to discriminate between pecuniary and behavioral reasons for credit card versus debit card use (Zinman 2005). These and other motivations for choosing different payment instruments were outlined in a recent survey, but their relative importance has not been determined (Scholnick et al. 2007). As per-transaction pricing of consumer payments is rare in Europe and the United States, these analyses cannot address what the impact of explicit pricing would be. Such pricing is rare since, reportedly, banks fear a loss of deposit market share if they are the first (and only) bank to implement it, while antitrust authorities would be suspicious of industry efforts to coordinate the implementation of per-transaction prices to minimize changes in relative market shares.

One country—Norway—has overcome these difficulties by coordinating only the timing of when direct pricing of consumer payments would start—not the level of prices to be charged, which could in fact be zero. The quid pro quo was a phasing out of banks’ practice of recouping payment costs through payment float—debiting consumer accounts prior to a value date for bill payments or delaying

funds availability for credits to accounts—which made it appear that payment use was “free” because the monetary cost to users was indirect and implicit. The goal was to make payment costs more explicit so consumers could match better the benefits and costs of different payment instruments, a response expected to lower the social cost of their payment system (Enge and Øwre 2006).

Our purpose is to determine the effect of differential transaction-based pricing of payment instruments on the adoption rate of electronic payments. This is done by comparing the shift to electronic payments in two countries—one that has transaction pricing (Norway) and one that does not (the Netherlands). Transaction-based prices are key since they directly affect consumers’ decisions about payment use, whereas implicit prices and fixed fees can have limited behavioral effects since these costs do not vary with usage and, when imposed, are typically stable over time. Nonprice effects—such as availability, convenience, and security—also influence payment use, and our attempt to control for these effects utilizes a two-country model.

Data on payment instrument use for many developed countries is available annually in various Bank for International Settlements and European Central Bank documents, as well as from payment statistics by national central banks. As these time series rarely exceed fifteen years, a parsimonious model specification is necessary. A comparable time series of actual payment instrument prices on a broad range of payment instruments is available only for Norway. We contrast the rapid adoption of electronic payments in Norway over 1990–2004 with the experience of the Netherlands, which also rapidly adopted electronic payments but did not impose per-transaction prices on consumers. By applying a system estimation to our model, we improve the degrees of freedom and increase the efficiency of our estimators.

If the incremental effect of direct pricing is large, holding constant other intracountry and cross-country influences affecting the adoption of electronic payments, then the potential social benefit can also be large. This suggests that antitrust concerns raised by possible bank coordination of the implementation of prices (but not the level) could be offset by subsequent social benefits. We find that while pricing has speeded up the shift to electronic payments, the shift was only about 20 percent faster in Norway with pricing than in

the Netherlands without pricing. Our analysis extends earlier work on payment pricing (Humphrey, Kim, and Vale 2001) by separating terminal-availability influences from price effects for point-of-sale transactions and by doubling the time period covered to capture more price variation. Newly available data also allows us to analyze the paper/electronic trade-off for giro-based bill payments as well as issues associated with nonstationarity in aggregate trended payment data. Finally, we control for unspecified and hard-to-measure nonprice influences to obtain potentially more accurate price elasticities.

In what follows, section 2 illustrates how the composition of payments has evolved in Norway and the Netherlands along with the levels of relative paper and electronic prices in Norway. Explicit transaction prices for consumers are zero in the Netherlands. Our focus is on the substitution of debit cards for cash (or cash and checks) at the point of sale along with the substitution of electronic giro payments for paper-initiated giro transactions.

In section 3, a parsimonious “country-difference” model is specified to separate the effect of pricing on point-of-sale debit card use and ATM (automated teller machine) cash withdrawals from differences in terminal availability and real personal consumption in our two countries. A similar model relies primarily on prices for the substitution of electronic versus paper-initiated giro payments (Internet connections far exceed giro use and are not a constraint, and, in any case, an electronic or paper giro can be initiated by phone or in person). No good data exists on payment use prior to the start of pricing (1986) in Norway, so nonprice attributes of different payment instruments can affect measured price elasticities if standard analysis is applied using only one country (as was the case in Humphrey, Kim, and Vale 2001). The Netherlands, which did not price, is used to “hold constant” nonprice attributes of the different payment instruments to obtain potentially more accurate price effects. Put differently, we seek to “subtract” the shift to electronic payments that presumably would have occurred without pricing—due to nonprice attributes, terminal availability, and per-person consumption levels—from the shift that is observed with pricing plus these three influences on use.

Our set of four equations is estimated in section 4 in a seemingly unrelated regression framework to improve efficiency. The effects of

prices on payment composition, including the implied price elasticities, are presented here. Different models are estimated to judge the robustness of the price effect under alternative specifications, such as different lagged relationships, first differences, and error correction. A summary of our results is contained in section 5, along with an estimate of the bank cost savings associated with the shift to electronic payments.

2. Payment Composition, Pricing, and Other Influences on Payment Instrument Use

2.1 Payment Composition

Both Norway and the Netherlands experienced a relatively rapid change in their payment composition for point-of-sale and bill-payment transactions over 1990–2004. Point-of-sale instruments are now almost solely debit cards and cash, but in the early 1990s, checks were also important.¹ As seen in table 1, the number of debit card transactions per person per year in Norway rose from 5 to 146 over our fifteen-year period, growing 25 percent per year.² The Netherlands started from a smaller base of one transaction per person per year but rose to seventy-seven, a 33 percent annual growth.³

¹Unlike in the United States, credit card use is low in Europe (and miniscule in the two countries we examine).

²Oil company terminals and cards were introduced in the 1980s as a substitute for cash at gas stations. Although these terminals also accepted bank debit cards, oil company cards could not be used elsewhere and were not priced. The Norwegian payment statistics do not include oil company transactions as debit card purchases (Norges Bank 2000, 33) and neither do we. Oil company terminals are included in our series of debit card terminals, however, since they accept debit cards for payment.

³Checks written per person in Norway went from twelve per person annually in 1990 to less than one in 2004. In the Netherlands, they went from seventeen to zero. Credit card transactions per person in both countries were less than one in 1990 and only three per person (the Netherlands) to five (Norway) in 2004 (or about 3 percent of card use in each country). The dominance of debit cards over credit cards is probably due to the fact that banks—not the credit card companies—through a joint venture were the first to introduce EFTPOS (electronic funds transfer at point of sale) directly from deposit accounts and have POS terminals connected to the bank network installed in shops. The banks' purpose was to replace checks and cash with cheaper electronic cards at the point of sale.

Table 1. Payment Instrument Use Per Person in Norway and the Netherlands (1990–2004)

	1990	1993	1996	1999	2002	2004	Growth Rate
Debit Card Transactions							
Norway	5	16	36	71	113	146	25%
Netherlands	1	4	24	44	66	77	33%
ATM Cash Withdrawals							
Norway	14	17	22	24	23	22	3%
Netherlands	8	20	26	28	30	28	9%
Electronic Giro (Credit Transfers + Direct Debits)							
Norway	15	18	29	46	65	78	12%
Netherlands	44	54	70	93	116	124	7%
Paper Giro (Credit Transfers)							
Norway	53	44	52	38	24	18	−7%
Netherlands	34	32	33	27	21	18	−5%
Source: DNB statistics, www.dnb.nl ; Norges Bank Annual Report on Payment Systems, www.norgesbank.nl .							

Part of this difference is due to the fact that Norway started 1990 with more debit card terminals in place than the Netherlands and so was at a higher point of usage on their logistic growth curve. In 2004, the average amount of a debit card transaction was about €55 in Norway and €44 in the Netherlands.

No country has time-series data on the number of cash transactions, although a few (markedly) different estimates exist for some countries at different points in time. These estimates differ primarily because of the difficulty of estimating very small-value cash transactions in which coins are often used and for which stored value cards—which so far have very limited acceptance—are the only real substitute.⁴ We use the number of cash withdrawals at ATMs as

⁴Brits and Winder (2005) provide an estimate of cash use in the Netherlands in 2002. Cash accounted for 85 percent of POS transactions and 56 percent of

our indicator of cash use in transactions.⁵ This quantity measure seems appropriate since the average real value of an ATM withdrawal rose by less than 1 percent annually in both countries.⁶ While each cash withdrawal (€138 on average in Norway and €107 in the Netherlands in 2004) funds multiple actual cash transactions, the act of withdrawing cash is priced in Norway, while its use at the point of sale is not. Thus, we compare debit card and cash use at the point where both are actually priced and consumer choice is exercised.⁷

Analysis of the U.S. market suggests that the price depositors pay when withdrawing cash from a foreign ATM (an ATM owned by a different bank) affects ATM deployment decisions (to earn revenue) and tends to tie depositors to banks with larger ATM networks (Hannan 2005; Knittel and Stango 2004). These effects are not important in the Netherlands, as consumers are not charged for ATM usage from any bank, and they play a minor role in Norway where there is only a foreign ATM fee set by the depositor's bank but no surcharge by the ATM owner. Thus, the contemporaneous determination of ATM price and terminal deployment is weak, and, in any event, we lag terminal availability in our model specification to mitigate this possible endogeneity problem.

sales value, while debit cards accounted for 13 percent of transactions and 40 percent of sales. The average value of a cash transaction was around €10 but over €47 for debit cards in 2002.

⁵Only recent time-series data exist on the number or value of cash withdrawals over the counter at financial institutions or through "cash-back" opportunities at supermarkets. ATM cash withdrawals are the only consistent data that exist for our time period.

⁶In 2004 prices, the average real ATM withdrawal in Norway rose gradually from €127 in 1991 to €138 in 2004, while in the Netherlands it rose from €94 in 1991 to €107 in 2004. These figures imply annual real-growth rates of roughly 0.6 percent and 0.9 percent for, respectively, Norway and the Netherlands. There is thus little reason to specify both the number and the value of ATM cash withdrawals.

⁷In reality, consumer payment choice is more complex. First, at an ATM there is the choice of whether to withdraw or not; then second, at the point of sale, there is the choice of whether to use "free" cash or a priced debit card. We leave this "two-stage decision" issue aside in our analysis. Our view is that the use of cash at the point of sale will be influenced by the cost of consumers' replenishing their inventory of cash via an ATM (or other sources).

In both Norway and the Netherlands, debit card use expanded at a rapid rate, while growth of ATM cash withdrawals was much smaller. As shown below (and in table 2), the average price of an ATM withdrawal rose relative to a debit card transaction in Norway, but these two prices were both zero in the Netherlands. If relative price was the only influence on relative use, we would expect a slower growth for ATM withdrawals in Norway (where the ATM price was, after 1996, higher than debit cards) than in the Netherlands (where there is no difference in relative prices). We see indications of this for ATMs in table 1 (as transaction growth in Norway is slower than in the Netherlands), but we do not see it for debit card use (where relative prices would favor more rapid growth in Norway over the Netherlands).

In order to reflect the substantially lower cost associated with electronic bill payments, employee disbursements, and interbusiness transactions over giro networks, the price of an electronic giro payment in Norway was less than a paper-initiated giro transaction (either delivered in the mail or passed over the counter at a bank or postal office). Since giro prices were zero in the Netherlands, one would expect to see a more rapid growth of electronic giro transactions and slower growth (or greater reduction) of paper giro transactions in Norway than in the Netherlands. Both of these expectations are realized in table 1. Per-person use of electronic giro payments in Norway rose from 15 transactions to 78 over 1990–2004 (growing 12 percent a year) while only rising from 44 transactions to 124 over the same period in the Netherlands (growing 7 percent annually).⁸ At the same time, paper giro use fell in both countries but from a higher level and at a greater rate in Norway. Indeed, by 2004 individuals in both countries initiated only 18 paper giro transactions per year.

2.2 Payment Prices in Norway

The average—and sometimes weighted-average—per-transaction prices being charged for different payment instruments in Norway are illustrated in table 2. Since there are no per-transaction fees in

⁸By 2004, direct debits accounted for 10 percent of electronic giro payments in Norway but 56 percent in the Netherlands. This is the main reason why electronic giro payments per person in the Netherlands are higher than in Norway.

Table 2. Average Per-Transaction Prices for Different Payment Instruments in Norway

Prices in Euros	1990	1993	1996	1999	2002	2004	Growth Rate
Debit Card Price							
Norway	.18	.23	.25	.26	.28	.26	2%
ATM Cash Withdrawal Price							
Norway	.05	.18	.24	.29	.39	.40	14%
Relative Price: Debit Card/ATM Cash Withdrawal							
	3.60	1.28	1.04	.90	.72	.65	−11%
Electronic Giro Price							
Norway	.10	.18	.22	.23	.31	.27	7%
Paper Giro Price							
Norway	.35	.62	1.18	1.86	2.65	2.76	15%
Relative Price: Electronic Giro/Paper Giro							
	.29	.29	.20	.12	.11	.09	−8%
Source: DNB statistics, www.dnb.nl ; Norges Bank Annual Report on Payment Systems, www.norgesbank.nl .							

the Netherlands, the relative prices that Norwegian consumers face also reflect the difference in prices faced between Norway and the Netherlands. This is the price effect we wish to separate from other influences on payment choice in these two countries.⁹

⁹Consumers in the Netherlands do face a fixed annual fee for maintaining an account (about €6) that allows the holder to use a debit card and the bank’s ATM network to withdraw cash, while Internet banking typically has a one-time startup fee (around €15). Because these fees are fixed, consumers “see” a zero price per additional transaction and respond accordingly. If the fixed fees, which did not vary much over time, were added as additional variables in our (double log) model, basically only the intercept and not the price elasticities would be affected. Additionally, while the average value of a debit card transaction was between €40 and €50, about 3 percent of all Dutch debit card transactions

The weighted-average per-transaction price of a cash withdrawal in Norway was in 1996, and earlier, less than that for debit cards.¹⁰ This was because a cash withdrawal at one's own bank was free during business hours and prices applied only to withdrawals after business hours or at another bank's ATM. While debit cards started out in 1990 with a price that was more than three times higher than the weighted average of different ATM prices (table 2, row 3), it ended up being only 65 percent of the cash withdrawal price in 2004. Thus, only after 1996 did the absolute price of a debit card favor its use over cash when EFTPOS terminals were available.¹¹ But even before 1996, there was an indirect inducement to use debit cards in Norway when it became possible in late 1992 to obtain "cash back" from a debit card transaction at the point of sale.¹² This avoided the extra cost and inconvenience of having to use an ATM to withdraw cash, since small amounts of cash could be obtained at no additional cost when making purchases at the local market.

There was a stronger relative price inducement to use an electronic rather than a paper-initiated giro transaction for consumer bill payments. In 1990, the price of an electronic giro transaction was only 29 percent as high as a paper giro payment, but by 2004

were subject to a merchant surcharge of about 15 eurocents (De Nederlandse Bank 2004) when the transaction value was less than €10–€12. Some merchants wished to discourage use of debit cards for low-value transactions since accepting cash is cheaper due to bank fees paid by merchants for debit card transactions.

¹⁰This observation only holds on a per-transaction basis. On average, one ATM withdrawal could fund roughly two to three debit card transactions. However, since this difference in "transaction domain" between both instruments is relatively stable over time, it should only affect the intercept in our model in logs.

¹¹The relative debit card/ATM price changes shown reflect banks' initial efforts to induce depositors to shift cash withdrawals from branch offices to cheaper ATMs and then later from cash use to even cheaper debit card transactions.

¹²Although cash-back transactions and cash at the counter at one's own bank are also sources for obtaining cash for free in Norway—and implicitly lower the effective price for obtaining cash compared to our use of the weighted average of free and priced ATM access—these data are available only for recent years and therefore could not be directly included in the analysis. The alternative of including a cash-back dummy, since it was collinear with the debit card price variable already in the model, yielded anomalous results, suggesting that these two effects cannot be reliably separated. Thus, the debit card price elasticity reported below is best considered as a combination of price and cash-back effects in Norway.

this had fallen to only 9 percent of the paper price. In the beginning, electronic giro payments were initiated via telephone, but this was later overtaken by the spread of Internet banking. This applies to credit transfers where the consumer retains control in initiating a payment, as opposed to a direct debit where the receiver of the credit initiates the debit to the consumer's account under a pre-arranged contractual agreement. Billers often give a slight discount to customers who pay by direct debit, thus creating a slight price advantage over a credit transfer. However, regardless of which party initiates the payment, both are counted as a single electronic giro transaction.

It is important to note that the prices charged in Norway do not cover the full bank cost of making a payment (cf. Flatraaker and Robinson 1995; Gresvik and Øwre 2003). In 1988, transaction prices covered only around 25 percent of the banks' payment cost, but this coverage had risen to around 70 percent in 2001.¹³ As well, in both countries banks initially made some effort to inform customers of the advantages of using lower-cost electronic payments whether or not the transactions were directly priced (but we cannot explicitly account for this in our model).

2.3 Terminal Availability and Levels of Consumption

While relative prices provide an inducement to use electronic payments at the point of sale, this can be accomplished only if a merchant has an EFTPOS terminal that can be used. This observation points to the two-sided nature of the payment market, which influences the adoption rate of new payment instruments. In particular, the market for electronic payment services is considered a two-sided market in the sense that both consumers and merchants are needed simultaneously to demand and "consume" card payments. Suppliers of payment card services (or so-called "platforms") can effectively cross-subsidize between merchants and consumers through differential pricing to stimulate this demand. In two-sided markets, typically only one side is charged on a per-transaction basis, while the

¹³The relationship between fees and underlying costs is different in Sweden, with surplus bank revenues from card transactions cross-subsidizing the expense of providing cash, distorting resource allocation (Sveriges Riksbank 2004, with more detail in Guibourg and Segendorff 2007).

other side obtains the service (almost) for free in order to generate greater demand.¹⁴ Indeed, merchants value a wide diffusion of payment cards among consumers, while consumers benefit from high terminal density at retail locations that accept their cards. In our analysis, payment card and ATM terminal density are included to take this two-sided effect into account in explaining relative payment card usage.

Table 3 shows the number of EFTPOS terminals in place in Norway and the Netherlands over 1990–2004 per one million of

Table 3. Terminal Availability, Real Consumption, and Demographic Influences on Payment Instrument Use

	1990	1993	1996	1999	2002	2004	Growth Rate
Debit Card EFTPOS Terminals (per mil population)							
Norway	2,487	6,324	8,932	13,214	17,723	21,091	15%
Netherlands	148	1,600	6,170	9,176	10,941	11,967	34%
ATM Terminals (per mil population)							
Norway	419	396	426	451	484	473	0.8%
Netherlands	180	291	395	421	465	468	6.6%
Real Per Capita Personal Consumption (in 1,000)							
Norway	11.9	12.1	13.9	15.1	17.8	16.7	2.3%
Netherlands	9.2	9.3	9.9	10.9	11.4	11.3	1.4%
Share of Young Adults in Population							
Norway	8.0	7.8	7.2	6.4	6.0	6.0	−1.9%
Netherlands	8.5	8.2	7.0	6.1	6.0	6.0	−2.3%
Source: DNB statistics, www.dnb.nl ; National Accounts; Dutch CBS; Norges Bank Annual Report on Payment Systems, www.norgesbank.nl ; IFS.							

¹⁴In Norway, the consumer side is directly charged for its use of payment instruments, while in the Netherlands, the retailer side of the market pays per transaction. Bolt and Tieman (2004) provide an explanation for these widely observed completely skewed pricing strategies in two-sided markets. See Rochet and Tirole (2003) for a rigorous analysis of two-sided markets and competition.

population (which controls for differences in population size).¹⁵ As shown in the first two rows, Norway had almost twice as many debit card terminals as the Netherlands in 2004, and this difference was far more extreme in earlier periods. While the growth of EFTPOS terminals has been more than twice as rapid in the Netherlands, it still has a long way to go to provide the same density of terminal access as Norway. By this measure alone, it really would not be possible—regardless of any price incentive—for consumers in the Netherlands to use debit cards with the same intensity per person as they do in Norway. As noted earlier, there is no price incentive to use debit cards in the Netherlands, so there are two reasons—no price incentive and fewer EFTPOS terminals per person—to expect that the Netherlands would use debit cards less intensively than Norway. Even so, as shown below, it is difficult to separate the effect of prices from terminal availability on debit card and ATM use.

The same “separation problem” exists for cash withdrawals at ATMs. Norway prices ATM withdrawals, while the Netherlands does not, and for the entire period Norway also provided a greater density of ATMs to withdraw cash from (table 3, row 3). Separating the price effect from the terminal effect for ATM cash withdrawals may be somewhat easier here since by 2004 both countries had almost the same ATM density but withdrawals were priced only in Norway and, compared to the Netherlands after 1993, per-person use in Norway was correspondingly less (table 1, row 3).¹⁶

Inferences on the relative importance of pricing may be more accurate if two other possible, but small, influences on payment choice are considered. One concerns differences in the level of real

¹⁵In 2004, the population in the Netherlands was 16.3 million; in Norway it was 4.6 million.

¹⁶As Norway is roughly nine times larger than the Netherlands, differences in population density may compromise the usefulness of our availability measure of ATM and EFTPOS terminals. However, both countries are highly urbanized, which is probably the most important driver for installing terminals. In Norway, the five largest cities account for about 25 percent of total population but only 1 percent of total geographic area (see Norway statistics, www.ssb.no). Less extreme, in the Netherlands, the ten largest cities make up roughly 20 percent of Dutch population, with 3.5 percent of the geographic area (see CBS statistics, www.cbs.nl). Since this difference in densities is effectively a constant over fifteen years, in our log-difference equation its impact would affect only the intercept and not the slope parameter, which is our terminal elasticity.

per capita personal consumption between the two countries, since higher levels of real consumption tend to be associated with larger numbers of transactions.¹⁷ A second influence concerns the possibility that changes in the number of young adults in both countries may affect differences in new payment adoption rates. Consumer surveys indicate that young adults and higher-income individuals adopt new payment arrangements more rapidly than others, even without pricing. But direct pricing could well affect the adoption rates of those with greater habit persistence, those with a lower opportunity cost, or those who do not value much the added convenience or security that electronic payments can offer.

The level and variation of both per capita consumption and the share of young adults in the population over time are illustrated in the bottom half of table 3. Real per capita consumption in Norway was 29 percent greater than that in the Netherlands in 1990 but rose to be 48 percent higher in 2004. This difference should be associated with a rising number of all types of transactions in Norway relative to the Netherlands. There are smaller differences between these two countries in the shares of young adults—new entrants into the labor force aged twenty to twenty-four. Indeed, these shares are falling in both countries.¹⁸

3. A Country-Difference Model of Payment Choice

Differences between Norway and the Netherlands are used to try to explain per capita use of debit cards, ATM cash withdrawals, and electronic and paper giro payments. As outlined above, the main influences on payment use and composition are differences in the number of EFTPOS and ATM terminals per million population, the prices being charged in Norway (positive) and the Netherlands (zero), and differences in the level of real per capita consumption.

¹⁷All monetary values for Norway (prices as well as real consumption) have been translated from Norwegian kroner into euros using a purchasing power parity exchange rate. Also, real per capita consumption in Norway includes oil revenues only indirectly, as some of this revenue is used to finance government expenditures, which likely reduces taxes from what they otherwise would be, permitting real consumption to be larger.

¹⁸Demographic variables are typically extremely smooth series. In implementation, this created convergence problems in our system estimation and the population share variable was excluded.

Our time period is short (only fifteen years), as time-series data on payment instrument use have only recently been deemed important enough to be routinely collected at the country level by government agencies. While some time-series data on some payment types do exist for longer periods in some countries, this information is not comprehensive, nor are payment instrument prices available, since very few types of payment services are directly priced. Norway is the exception that allows us to undertake this analysis. These data constraints impose a parsimonious specification on our explanatory four-equation model:

$$\begin{aligned}
 CARD_t = & \alpha_1 + \alpha_2 CARDTERMINAL_{t-1} + \alpha_3 CARDPRICE_t \\
 & + 1/2(\alpha_{22} CARDTERMINAL_{t-1}^2 \\
 & + \alpha_{33} CARDPRICE_t^2) \\
 & + \alpha_{23} CARDTERMINAL_{t-1} * CARDPRICE_t \\
 & + \alpha_4 CONSUMPTION_t + \varepsilon_{1t}
 \end{aligned} \tag{1}$$

$$\begin{aligned}
 ATM_t = & \beta_1 + \beta_2 ATMTERMINAL_{t-1} + \beta_3 CARDPRICE_t \\
 & + 1/2(\beta_{22} ATMTERMINAL_{t-1}^2 \\
 & + \beta_{33} CARDPRICE_t^2) \\
 & + \beta_{23} ATMTERMINAL_{t-1} * CARDPRICE_t \\
 & + \beta_4 CONSUMPTION_t + \varepsilon_{2t}
 \end{aligned} \tag{2}$$

$$\begin{aligned}
 EGIRO_t = & \gamma_1 + \gamma_2 EGIROPRICE_t + 1/2(\gamma_{22} EGIROPRICE_t^2) \\
 & + \gamma_3 CONSUMPTION_t + \varepsilon_{3t}
 \end{aligned} \tag{3}$$

$$\begin{aligned}
 PGIRO_t = & \delta_1 + \delta_2 EGIROPRICE_t + 1/2(\delta_{22} EGIROPRICE_t^2) \\
 & + \delta_3 CONSUMPTION_t + \varepsilon_{4t}.
 \end{aligned} \tag{4}$$

In the variable definitions below, NOR indicates Norway and NL indicates the Netherlands. Differences between these countries are expressed in index form:¹⁹

¹⁹In many cases, the log of the absolute difference in our variables between countries was negative (or changed from positive to negative), so all variables are expressed as the log of the ratio or index of the difference between countries.

$$\begin{aligned}
CARD &= \ln (\text{NOR debit card use/NL debit card use}), \text{ on a per-person basis;} \\
CARDTERMINAL &= \ln (\text{NOR card terminals/NL card terminals}), \text{ per million population;} \\
CARDPRICE &= \ln (\text{NOR card price/NOR ATM price}); \\
CONSUMPTION &= \ln (\text{NOR personal consumption/NL personal consumption}), \text{ real per capita;} \\
ATM &= \ln (\text{NOR ATM cash withdrawals/NL ATM cash withdrawals}), \text{ per person;} \\
ATMTERMINAL &= \ln (\text{NOR ATM terminals/NL ATM terminals}), \text{ per million population;} \\
EGIRO &= \ln (\text{NOR electronic giro use/NL electronic giro use}), \text{ per person;} \\
EGIROPRICE &= \ln (\text{NOR electronic giro price/NOR paper giro price}); \\
PGIRO &= \ln (\text{NOR paper giro use/NL paper giro use}), \text{ per person.}
\end{aligned}$$

Since debit card and ATM terminals have to be in place before consumers can use them—and, even when in place, typically have a lag before they are used at a significant level—these two terminal variables are lagged by one year in the model to give a closer correspondence between the exogenous availability of new terminals and their possible effect on use. Prices, of course, also have to be known before they can affect payment choice. The lag here is likely much shorter, and prices are specified as exogenous and contemporaneous.²⁰ Note that in the absence of Dutch prices for payment instruments, we implicitly take the relative price for the Netherlands as being constant, which will then only affect the intercept in our model.

Looking at the data, it appears that the two countries differ in when they introduced electronic payment instruments. In the Netherlands, usage of the electronic giro was on a higher level than in Norway in 1990, whereas Norway had a higher density of ATM and EFTPOS terminals. This “starting value” problem is taken into account in our logarithmic specification through the intercept, which

²⁰The effects of different lag arrangements on the results are noted in section 4.

is not restricted to a value of 1 (which would imply equal starting values for 1990).

4. Estimation Results and the Effect of Price on Payment Instrument Use

The system of equations (1)–(4) was estimated in a seemingly unrelated regression framework to allow for the possible correlation between errors in locally identifying debit card use with those for ATM cash withdrawals and similarly for electronic and paper giro use. With fifteen observations per equation, there are thirty-eight degrees of freedom ($d.f. = 4 * 15 - 22$). As shown in the appendix, the explanatory power of the model was high (the respective adjusted R^2 s were .96, .98, .84, and .75 from the system estimation). As the variables are not $I(0)$ and thus our levels estimation may be unbalanced, yielding spurious results, we first checked for the presence of any residual autocorrelation. Fortunately, system residual portmanteau (Ljung-Box) Q-test statistics (adjusted for small sample) indicate that the autocorrelations of the residuals are not statistically significant. As well, the Durbin-Watson values are fairly reasonable for our four levels equations (respectively, 2.12, 2.80, 1.86, and 2.03). These autocorrelation tests suggest that the variables of our four equations are likely to be cointegrated.

To formally test for cointegration, we applied augmented Dickey-Fuller (ADF) tests to the residuals of the levels equations (1)–(4), using critical values computed by Phillips and Ouliaris (1990). The test statistics show that we reject the null of no cointegration at the 5 percent level for the ATM and the electronic giro equation, and at a 10 percent level for the debit card and paper giro equation (see the last column in table 5, discussed later).²¹ The signs of the estimated parameters appear to be reasonable and expected from theory, so the degree of spurious correlation, if any, is likely to be small. Moreover, as shown in the next subsection, the residuals of our levels equations—measuring the deviations of a “long-run”

²¹By definition, cointegration requires that the variables be integrated of the same order. Keeping in mind our small sample, applying ADF tests—using MacKinnon (1996) critical values—indicated that for fourteen out of seventeen variables, the null of a unit root is not rejected.

Table 4. Price and Terminal Elasticities Under Different Model Specifications

	Lagged Terminals	Lagged Terminals and Prices	No Lags	Separate Prices	
				Own	Substitute
Debit Card					
Price Effect	−.22**	−.48**	−.06	−.19	−.03
Terminal Effect	.53**	.57**	.94**		.49**
ATM Cash Withdrawal					
Price Effect	.23**	.31**	.29**	−.86**	.69**
Terminal Effect	−.16**	−.49**	−.35**		−.35**
Electronic Giro					
Price Effect	−.46**	−.53**	−.44**	.21*	.10*
Paper Giro					
Price Effect	.27**	.25**	.33**	−.03	.03
Note: Starred (**,*) values indicate significance levels of 1 percent and 5 percent, respectively.					

relationship—significantly affect the short-run dynamics of the variables. Hence, we feel that these findings support the results of our preferred model using levels data.²²

The derivatives of equations (1)–(4), first with respect to relative prices and then with respect to lagged terminal availability, are shown in column 1 of table 4 (with parameters and other statistics shown in the appendix). A 10 percent reduction in the price of debit cards relative to an ATM cash withdrawal is associated with a 2.2 percent rise in the relative use of cards in Norway compared

²²Naturally, given the small sample size, unit-root and cointegration tests have reduced power, and strict interpretation of the estimated elasticities should be made with caution.

Table 5. Price and Terminal Elasticities: Data in Levels and First Differences

	Levels Data	Error Cor. Residuals	Error Cor. 1st Diff. & Levels ^a	1st Diff. Data ^a	ADF Test
Debit Card					−3.82 ^{° b}
Price Effect	−.22**		3.25	.47	
Terminal Effect	.53**		.82	.75**	
Feedback Parameter		−0.93*	−.07		
ATM Cash Withdrawal					−5.55*
Price Effect	.23**		.32*	.06	
Terminal Effect	−.16**		−.39**	.47**	
Feedback Parameter		−0.84*	−.83**		
Electronic Giro					−4.50*
Price Effect	−.46**		−.60**	.37**	
Feedback Parameter		−0.71*	−.24		
Paper Giro					−3.66 [°]
Price Effect	.27**		.49**	.11	
Feedback Parameter		−0.85*	−.45*		
Note: Starred (**, *) values indicate significance levels of 1 percent and 5 percent, respectively; circled (°) values indicate a 10 percent level. ^a Debit cards and ATMs formed one system estimation, while electronic and paper giros formed another in the first-differenced and error-correction models. ^b Test based on one significant explanatory variable.					

with the Netherlands (which has a zero explicit price for both cards and ATMs). At the same time, a 10 percent increase in debit card terminals in Norway relative to the Netherlands is associated with a 5.3 percent rise in debit card use in Norway relative to the Netherlands. As seen in the table, lagging both terminals and prices by one period doubles the strength of the price response (from −.22 to −.48)

but does not alter the terminal elasticity. Assuming no lags, however, increases considerably the apparent responsiveness of debit card use to changes in terminal availability—making it almost one-to-one in percentage terms—but the trade-off is that it generates a price elasticity insignificantly different from zero.

Over 1990–2004, the price of ATMs in Norway rose relative to that of debit cards. The price elasticity suggests that a 10 percent rise in the relative price of ATMs is associated with a 2.3 percent decrease in relative use.²³ Numerically, this is very similar to the result for the debit card equation, where a 10 percent reduction in the relative price of debit cards gives a 2.2 percent rise in relative use.²⁴ The ATM terminal elasticity, however, has an unexpected sign and is negative at its mean. When evaluated yearly, the terminal elasticity is positive over 1990–94, but the negative relationship for the remaining years dominates, giving a negative mean. Looking more closely at ATM use and terminal availability by year (not shown) suggests that the source of the negative elasticity is that per-person ATM use in Norway reaches a peak in 1998 and then falls, while ATM availability in Norway reaches a peak five years later in 2003. Similarly, ATM use in the Netherlands peaks in 2001, but terminals continue to expand. The apparent explanation for the negative ATM terminal elasticity is that ATM use has reached saturation (due in part to the price disincentive), while terminals are still being added (allowing banks to substitute ATMs for expensive branches, which fell absolutely in the two countries), giving the result that terminals are expanding while use is falling.

The estimated price effects for electronic and paper giro payments conform to expectations since, when the relative price of electronic giro transactions falls 10 percent, relative use of this instrument in Norway rises 4.6 percent compared with the Netherlands. Similarly, a 10 percent increase in the relative price of paper giro

²³Since the price ratio used in the ATM equation is the same as that used in the debit card equation— $\ln(\text{Norway debit card price}/\text{Norway ATM price})$ —the negative debit card price elasticity would become a positive elasticity in the ATM equation.

²⁴The ATM price effect is larger when both terminals and prices are lagged in the model. Using bank-level data for Spain, Scholnick et al. (2007) also find that debit cards substitute for ATM cash withdrawals.

payments is associated with a 2.7 percent reduction in relative use between the two countries.²⁵ Once an individual switched to making an electronic giro payment, almost all of their giro transactions were electronic, and electronic payment volume grew by inducing more and more individuals to switch. In contrast, the substitution of debit cards for ATMs was twofold since it involved individuals shifting a progressively larger share of their point-of-sale transactions from cash to cards over time as more terminals became available and relative prices changed as well as inducing more and more individuals to adopt and use cards.

To illustrate the robustness of our results, our preferred model in equations (1)–(4) was respecified so that direct debits, which comprise 10 percent of electronic giro payments in Norway but 56 percent in the Netherlands, were deleted from the electronic giro use (basically leaving only credit transfers). This had almost no effect on the price results shown in column 1 of table 4. Equations (1)–(4) were respecified again to include checks with ATMs so that both can substitute with debit cards. Checks were important in the early 1990s, had a high price, and their use effectively fell to zero by 2004. Nothing of substance was changed except that the debit card price elasticity lost significance.²⁶

Real per capita personal consumption was markedly higher in Norway and growing faster than in the Netherlands. We expected that this would have a significantly positive effect on expanding relative electronic payment use in Norway. However, the effect of real per capita personal consumption on payment use was insignificant in all four equations.

Just as an exercise, equations (1)–(4) were simplified by deleting the squared terminal, squared price, and terminal-price interaction variables. Then the remaining price ratio in each equation (e.g., debit card price/ATM price and electronic giro price/paper giro price) was reexpressed as the log of separate own and substitute price variables

²⁵Since the same price ratio is used in both the electronic and paper giro equations— $\ln(\text{Norway electronic giro price}/\text{Norway paper giro price})$ —the negative electronic giro price elasticity would become a positive elasticity in the paper giro equation.

²⁶The price and terminal elasticities were only slightly changed if, instead of estimating equations (1)–(4) as a single system, system estimation was applied to (1) and (2) and then separately to (3) and (4).

for each equation. The resulting own and substitute price elasticities, along with the reestimated terminal effect, are shown in the last two columns of table 4. Our preferred model (in column 1) is specified in ratio form, due to our limited sample, but it is of interest to see the implied own and cross-price elasticities that result from estimating each price elasticity separately. All but one own price elasticity is negative, and three of the four cross-price elasticities are positive (as would be expected for a substitute payment instrument). However, considering that only one negative own elasticity and two positive cross-elasticities were significant, it seems that the price effects are not very strong.

4.1 *Cointegration and Error Correction*

In the previous section we formally tested whether the variables of our levels equations were cointegrated. These tests indicated that the residuals were stationary and that we could reject the null hypothesis of no cointegration. To assess how deviations from the long-run equilibrium—as captured by the movements of the residuals—affect the short-run dynamics of the variables, we also estimated an error-correction (system) model. Since the models in levels, in first differences, or in error-correction form are all nested within an “autoregressive distributed lag” framework, it allows us to test which model fits the data best.

To illustrate, consider the following extension of equation (1), written in an autoregressive distributed lag regression format by adding lagged endogenous and exogenous variables:²⁷

$$\begin{aligned} CARD_t = & \alpha + \gamma CARD_{t-1} + \delta_1 CARDPRICE_t + \delta_2 CARDPRICE_{t-1} \\ & + \beta_1 CARDTERMINAL_{t-1} + \beta_2 CARDTERMINAL_{t-2} \\ & + u_t. \end{aligned} \tag{5}$$

Without affecting its ability to explain the data or changing the least-squares estimates of the parameters of interest, (5) may be

²⁷Note that compared with equation (1), the squared variables, interaction terms, and consumption have been excluded. These additional variables could be included without affecting our illustration.

rewritten in error-correction form:

$$\begin{aligned}\Delta CARD_t &= \alpha + \beta_1 \Delta CARDTERMINAL_{t-1} + \delta_1 \Delta CARDPRICE_t \\ &\quad + (\gamma - 1)(CARD_{t-1} - \alpha_{22} CARDTERMINAL_{t-2} \\ &\quad - \alpha_{33} CARDPRICE_{t-1}) + u_t,\end{aligned}\tag{6}$$

where $\alpha_{22} = \frac{\beta_1 + \beta_2}{\gamma - 1}$ and $\alpha_{33} = \frac{\delta_1 + \delta_2}{\gamma - 1}$ denote the long-run elasticities (equivalent to the elasticities in levels equation (1)). In (6) we have an equilibrium relationship describing the short-run dynamics,

$$\begin{aligned}\Delta CARD_t &= \alpha + \beta_1 \Delta CARDTERMINAL_{t-1} + \delta_1 \Delta CARDPRICE_t \\ &\quad + u_t,\end{aligned}$$

and an equilibrium error,

$$CARD_{t-1} - \alpha_{22} CARDTERMINAL_{t-2} - \alpha_{33} CARDPRICE_{t-1},$$

which measures the deviation from the long-run relationship between the variables $CARD_{t-1}$, $CARDTERMINAL_{t-2}$, and $CARDPRICE_{t-1}$. Consequently, the feedback parameter $\gamma - 1$ can be interpreted as the proportion of the resulting disequilibrium that is reflected in the movement of $CARD_t$ in one period. If the parameter $\gamma_1 - 1$ is negative and significantly different from zero, the model in error-correction format cannot be rejected.²⁸ In this case, long-run equilibrium deviations have a significant impact on the short-run dynamics, which disqualifies the model in first differences. On the other hand, insignificance of the adjustment parameter would favor a first-difference model and implies a “disconnect” between the short run and long run. This disconnect would then cast doubt as well on the empirical relevance of the long-run relationship (even when the variables are cointegrated).

Because our variables in levels are cointegrated, direct estimation of equations (1)–(4) yields “super-consistent” estimators of the (cointegrating) long-run elasticities. Under cointegration, the residuals from the levels equations can be used to estimate the

²⁸As a stability condition, the feedback parameter $\gamma - 1$ needs to be between 0 and -1 .

error-correction model. Alternatively, one can estimate equation (6) directly, using first differences and lagged level variables, but this will reduce our already limited degrees of freedom even further. The error-correction estimation results of all four equations are shown in table 5, along with the price and terminal elasticities for our preferred levels model from table 4.

The second column shows the adjustment parameters using the stationary residuals of the levels equations (1)–(4). All parameters have the right sign and magnitude, and are significant at the 5 percent level, indicating that the short-run dynamics is indeed significantly influenced by deviations from the long-run relationship.²⁹ The first-difference results are shown in column 4 of table 5, but the model is rejected. While the terminal elasticities have the expected sign and are significant, this is at the expense of weak results for the price elasticities. Compared with using residuals in the error-correction estimation, we obtain weaker results when we apply direct estimation of the error-correction equations (see column 3). Here the debit card price elasticity is no longer significant, but the other price elasticities have the expected sign and are significant (even with a reduction in degrees of freedom), although in two cases the feedback parameter was not significant at the 5 percent level.

Given our data limitations, the outcomes of the cointegration tests, and the performance of the error-correction model using stationary residuals, these results weakly suggest that the (cointegrating) price elasticities using levels data in equations (1)–(4)—our preferred model—are robust and can be relied upon as long-run estimates.

4.2 Estimation of Electronic for Paper Substitution in Norway

The effect of pricing on payment instrument use is also estimated for Norway alone. This approach should give similar results

²⁹The residual properties of the error-correction estimation using residuals are fairly reasonable with (adjusted) Ljung-Box stats: $Q(1) = 24.1$ ($p = 0.09$), $Q(2) = 43.1$ ($p = 0.09$), $Q(3) = 62.8$ ($p = 0.07$), and $Q(10) = 163.5$ ($p = 0.41$).

to our country-difference model if nonprice characteristics that affect payment use in a country are weak. The specification is linear and simpler than our country-difference model (due to degrees-of-freedom considerations), and all the data are for Norway:

$$\begin{aligned} CARDATM_t = & \alpha_1 + \alpha_2 CARDATMTERMINAL_{t-1} \\ & + \alpha_3 CARDATMPRICE_t + \alpha_4 CONSUMPTION_t \\ & + \varepsilon_{1t} \end{aligned} \quad (7)$$

$$\begin{aligned} ELEPAPER_t = & \beta_1 + \beta_2 ELEPAPERPRICE_t + \beta_3 CONSUMPTION_t \\ & + \varepsilon_{2t}, \end{aligned} \quad (8)$$

where

$$\begin{aligned} CARDATM &= \ln (\text{debit card use/ATM use}), \text{ on a} \\ &\quad \text{per-person basis;} \\ CARDATMTERMINAL &= \ln (\text{card terminals/ATM terminals}), \\ &\quad \text{per million population;} \\ CARDATMPRICE &= \ln (\text{debit card price/ATM price}); \\ CONSUMPTION &= \ln (\text{personal consumption}), \text{ real per} \\ &\quad \text{capita;} \\ ELEPAPER &= \ln (\text{electronic giro use/paper giro use}), \\ &\quad \text{per person;} \\ ELEPAPERPRICE &= \ln (\text{electronic giro price/paper giro} \\ &\quad \text{price}). \end{aligned}$$

Equations (7) and (8) were estimated in a systems equation framework (with 23 degrees of freedom, d.f. = $2 * 15 - 7$). As shown in table 6, the elasticity of substitution between debit cards and cash was $-.20$ using levels data and $-.31$ in first differences. A 10 percent rise in the relative price of an ATM cash withdrawal (which reduces the ratio of debit card to ATM prices) is associated with a small (2.0 percent or 3.1 percent) rise in the ratio of debit card to ATM use. If this parameter was -1.0 , then the expenditure shares of debit cards and ATMs would be unchanged, since a 10 percent

Table 6. Price and Terminal Substitution for Norway

	Levels Data	First-Differenced Data
Debit Card/ATM Substitution		
Price Effect	-.20*	-.31**
Terminal Effect	.54**	.06
Electronic/Paper Giro Substitution		
Price Effect	.54**	.13
Note: Starred (**, *) values indicate significance levels of 1 percent and 5 percent, respectively.		

relative rise in the ATM price would be exactly offset by a 10 percent decrease in relative ATM use. Since the parameter is less than 1 (in absolute value), the expenditure share of ATMs rises as the price-induced substitution is less responsive than in, say, a traditional Cobb-Douglas framework where the elasticity of input substitution to a price change is 1.0. The elasticity of terminal availability on debit card and ATM use is .54, which indicates that a 10 percent relative rise in debit card terminals leads to a 5.4 percent relative rise in debit card use.

The substitution elasticity between electronic and paper giro transactions initially had the wrong sign (at .54) and was significant using levels data in a system estimation. Use of first-differenced data did not alter this sign, but dropping real personal consumption from (7) and (8), which had a large and significant effect, did (giving a significant -1.98 value for electronic/paper giro substitution). Overall, the card/ATM price and terminal elasticity results were similar between our two-country and single-country applications. This occurs because terminal availability appears to be a good instrument for the combined effect of non-price influences. In contrast, for giro transactions where such a proxy is not available, the two-country giro price elasticity indicated a much smaller degree of price substitution than the single-country specification, suggesting that nonprice effects in a single-country model overstate the bill-payment price elasticity.

4.3 *The Effect of Pricing on Payment Use*

Our overall conclusion is that the availability of card terminals is a good proxy for the net effect of nonprice influences (such as convenience) on card use. And although nonprice influences on card use appear to have a stronger effect than does per-transaction pricing, the shift to electronic payments is speeded up when pricing is present.³⁰ If both prices and terminals are expanded at similar percentage rates, then the adoption of electronic payments could have been speeded up by perhaps 40 percent compared with not having per-transaction pricing.³¹ As seen in tables 2 and 3, however, debit card terminals changed at a much greater rate than did the price of ATMs or the relative prices of cards to ATMs, indicating that in this instance a potential speedup of 40 percent is too high and was not realized.

More precisely, for Norway, the average annual growth in debit card terminal density equaled +15 percent, whereas the growth in card price relative to ATMs was -11 percent. Given the estimated elasticities in column 1 of table 4, this would predict a relative rise of debit card use over ATMs of $15\% \times 0.53 + 11\% \times 0.22 = 10.4\%$ from the terminal and price effects alone. Without any price inducements, this increase in usage would be $15\% \times 0.53 = 8.0\%$, suggesting that the substitution process has been speeded up by approximately $2.4/10.4 = 23\%$, although the realized contribution of pricing to debit card adoption was 2.4 percent a year.³²

Electronic giro payments do not have a terminal constraint, and the influence of consumption growth on payment use is not significant, so only the effect of pricing is measured in the single-country case, while nonprice influences are incorporated in the two-country estimate. The growth of electronic giro relative to paper giro prices

³⁰Dutch survey results confirm the relative importance of terminal availability for payment instrument usage and stress also the nonprice attributes of payment instruments (see Jonker 2007).

³¹This estimate is derived from the ratio of the price elasticity in our preferred model (-22 percent) in column 1 of table 4 to the terminal elasticity (53 percent), which equals .42.

³²The same calculation using price (-.20) and terminal (.54) elasticities for Norway alone from table 6 gives a 20 percent speedup for debit card use (with a contribution of 2.2 percent annually).

was -8 percent, while the price elasticity in table 4 was $.46$, suggesting that the realized contribution of pricing to the adoption of electronic giro payments was $8\% \times 0.46 = 3.7\%$ annually. Thus, in terms of both the size of the estimated price elasticities and their realized impact on adoption rates, the effect of pricing on the shift to electronic payments is greater for giro transactions than for debit cards.

5. Summary and Conclusions

Electronic payment instruments (credit cards excepted) are considerably cheaper than their paper-based alternatives, including cash. Banks and merchants are interested in shifting users to electronic payments to save costs, as are some government policymakers who seek to improve the cost efficiency of their nation's payments system. Historically, banks have recouped their payment costs through (i) interest earned on payment float (from delaying availability of funds credited to accounts and debiting accounts prior to bill-payment value dates), (ii) maintaining a spread between market rates and the rate paid on deposits, and (iii) charging flat monthly fees or imposing balance requirements. In contrast to business users, consumers face very few payment services that are priced on a per-transaction basis and so have little price incentive to choose the lowest-cost instrument either at the point of sale or for bill payments.

Banks are well aware that transaction pricing can speed up the shift to electronic payments, but they are reluctant to lose deposit market share by being the first (and perhaps only) bank to implement explicit prices differentiated according to underlying costs. While this problem is mitigated if most (or all) banks implement pricing at about the same time, antitrust authorities are unlikely to view such coordination as being in the public interest unless the social benefits from pricing are significant and the quid pro quo is a compensating reduction in payment float, a higher interest rate paid on deposits, or a reduction in flat fees or balance requirements. Indeed, float reduction was the trade-off when banks coordinated the timing of when they would implement pricing in Norway (there was no coordination in the prices to be charged, and initially some were zero).

We use the experience of Norway (which priced its payment services) and the Netherlands (which did not) over 1990–2004 to try to determine what the incremental effect of transaction pricing may be on the adoption of debit cards versus withdrawing cash from an ATM and on the adoption of electronic giro transactions (credit transfers and direct debits) versus paper giros. Specifically, we compare per-person payment instrument use in Norway in response to the prices being charged, the availability of terminals, and the level of real consumption with the experience of the Netherlands, which also adopted electronic payments but did not price. Our four-equation country-difference model spanned fifteen years—the limit of the available data—and during this time the share of electronic payments rose by some 60 percentage points, from around the mid-twenties to the mid-eighties, which in most cases easily covered the inflection point where the share of electronic payments switches from rising at an increasing rate to rising at a decreasing rate.

Our model is estimated in a systems-equation framework using levels data, and robustness is illustrated by estimating models in a first-difference and error-correction framework. Price and terminal elasticities derived from these models form the basis for our conclusions and indicate the incremental effect of pricing on the adoption rate of electronic payments. The similarity of our card/ATM price and terminal elasticity results between our two-country and single-country applications suggests that terminal availability is itself a good proxy for hard-to-specify/hard-to-measure nonprice attributes of card use at the point of sale. In contrast, for giro bill-payment transactions where such a proxy was not available, the two-country giro price elasticity indicated a much smaller degree of price substitution than the single-country specification. This indicates that unspecified nonprice effects in a single-country framework can overstate the “true” giro price elasticity value.

The effects of pricing differ depending on which instruments are being considered. Overall, pricing has a smaller effect on shifting consumers from ATM cash withdrawals to debit card use than it does in shifting use from paper to electronic giro transactions. The reason for this difference seems to be that consumers value the non-price benefits associated with debit card use (convenience, security) such that the availability of terminals needed for debit card transactions has a stronger effect on debit card use than prices (as evidenced

by the fact that the debit card price elasticity is smaller than the terminal elasticity). Debit cards also substitute for costly checks, and the high price on these instruments in Norway was associated with their virtual elimination, although the same thing happened in the Netherlands, which did not price. While terminal availability has a stronger effect on debit card use than does pricing, the shift to cards can be speeded up when explicit pricing is combined with terminal availability. Using our estimated elasticities and the actual changes in prices and terminals, the predicted relative rise of debit card use over ATMs was 8 percent from terminal effects alone but rose to 10.4 percent with pricing, an increase of over 20 percent.

The effect of pricing on electronic giro use was greater than it was for debit cards since the electronic giro price elasticity is larger and the percent change in price experienced was greater. Reasons for this difference are the above-mentioned nonprice convenience and security attributes of debit cards, which promoted use, along with the fact that for one-third of our time period the absolute price of a debit card transaction was actually higher than the weighted-average price of an ATM cash withdrawal. In contrast, the price of an electronic giro was always absolutely lower than the paper giro price. Even though the relative prices of debit cards and electronic giros were both falling over the entire period, the higher absolute price of a debit card transaction versus an ATM transaction would be expected to dull the overall price response being measured for the entire period, as there is no strong reason to believe that the price response is symmetric (and symmetry was not imposed in our model) since the nonprice attributes of debit cards and ATMs are different. Thus, if pricing is implemented, it will likely be more successful if the absolute price of the less expensive instrument is always absolutely lower than the price of the more expensive instrument.³³ Also, the fact that terminal elasticities are an important component of the substitution process for cards suggests that non-priced attributes—such as convenience and security—play a greater

³³This was not done in Norway, perhaps because dispensing cash via an ATM was already less expensive than dispensing it through a branch office (assuming the rise in dispensing frequency at an ATM does not rise enough to offset this advantage).

role for cards versus cash than for electronic versus paper giro transactions.

Given the large resource cost of a country's payments system, it is obvious that shifting from paper to electronic payments can confer social benefits. Although not shown here, we estimated that the shift from ATM cash withdrawals and checks to debit cards plus the savings from shifting from paper to electronic giro transactions—if it happened without a lag—may save €0.7 billion in bank costs for Norway (.35 percent of GDP in 2004) and €2.9 billion for the Netherlands (.61 percent of GDP).³⁴ On a discounted basis over time, shifting from 90 percent paper-based instruments and cash to 90 percent electronic and card instruments could save about €2,300 per person in each country.³⁵ Merchant cost savings, for which little information exists, would increase these savings estimates.

As both Norway and the Netherlands are on their way to realizing the full potential gains from electronic payments, the issue of pricing or not pricing is seemingly more a policy topic for developed countries that are not as far along in the substitution process or for most developing countries that are in the initial stages of thinking about how to improve the efficiency of their payments system. The social benefits of electronic payments are quite large and may convince antitrust authorities to allow the coordination of the timing of the implementation of pricing (but not, of course, the prices to be charged) to speed up this transition. But pricing could become a reality even in countries that have largely shifted to electronic payments since, with low or falling interest rate margins, this would facilitate the recoupment of bank payment costs compared to continued reliance on indirect methods (large loan-deposit interest rate spread, delayed funds availability, minimum balance requirements, or monthly fixed fees).

³⁴The savings are absolutely higher for the Netherlands primarily because its population (16.3 million) is much larger than that of Norway (4.6 million), but GDP per capita is lower.

³⁵More information on the cost-savings estimate derived from payment-cost data and fitted logistic S-curves of a changing share of payment use is contained in a working paper (Bolt, Humphrey, and Uittenbogaard 2005).

Appendix

Parameter values and other statistics are reported below for equations (1)–(4) and generated the price and terminal elasticities shown in column 1 of table 4. Standard errors were computed from a heteroskedastic-consistent matrix (Robust-White). The nonlinear estimation procedure was LSQ in TSP (version 5.0). More information on data sources can be found in Bolt, Humphrey, and Uittenbogaard (2005).

Variable	Parameter	Estimation	T-statistic
Constant	α_1	.497961	2.773980
DIFDCARDTML	α_2	.373649	2.558270
DIFDCARDPRICE	α_3	.055702	.180209
DIFDCARDTML2	α_{22}	.167924	1.244530
DIFDCARDPRICE2	α_{33}	.870318	1.219510
DCARDTMLPRICE	α_{23}	−.308628	−.922949
DIFPCONS	α_4	−.571509	−1.137260
Constant	β_1	.035974	.554554
DIFATMTML	β_2	−1.675620	−8.560530
DIFDCARDPRICE	β_3	1.003880	5.852760
DIFATMTML2	β_{22}	5.935940	7.781290
DIFDCARDPRICE2	β_{33}	2.332310	3.406300
ATMTMLPRICE	β_{23}	−3.151620	−4.504500
DIFPCONS	β_4	−.138567	−.766170
Constant	γ_1	−.519080	−1.212480
DIFEGIROPRICE	γ_2	1.068860	2.151150
DIFEGIROPRICE2	γ_{22}	.881313	3.062310
DIFPCONS	γ_3	.544589	1.135200
Constant	δ_1	−.606181	−1.201530
DIFEGIROPRICE	δ_2	−1.694780	−2.901820
DIFEGIROPRICE2	δ_{22}	−1.135840	−3.307110
DIFPCONS	δ_3	−.678980	−1.347590

Debit card equation (1):

adj. R^2 = .960, S.E. of regression = 0.010, D-W = 2.12.

ATM cash withdrawal equation (2):

adj. R^2 = .975, S.E. of regression = 0.035, D-W = 2.80.

Electronic giro equation (3):

adj. $R^2 = .836$, S.E. of regression = 0.087, D-W = 1.86.

Paper giro equation (4):

adj. $R^2 = .746$, S.E. of regression = 0.087, D-W = 2.03.

System residual portmanteau (adjusted) test for autocorrelations:

$Q(1) = 20.92$ ($p = 0.18$), $Q(2) = 33.41$ ($p = 0.40$), $Q(5) = 79.45$ ($p = 0.50$), $Q(10) = 163.44$ ($p = 0.41$).

Log-likelihood: 99.73; Number of observations: 60.

References

- Bolt, W., D. B. Humphrey, and R. Uittenbogaard. 2005. "The Effect of Transaction Pricing on the Adoption of Electronic Payments: A Cross-Country Comparison." Working Paper No. 71, De Nederlandsche Bank.
- Bolt, W., and A. F. Tieman. 2004. "Skewed Pricing in Two-Sided Markets: An IO Approach." Forthcoming in *International Journal of Industrial Organization*.
- Borzekowski, R., and E. K. Kiser. 2006. "The Choice at the Checkout: Quantifying Demand Across Payment Instruments." Forthcoming in *International Journal of Industrial Organization*.
- Brits, H., and C. Winder. 2005. "Payments Are No Free Lunch." *Occasional Studies* (De Nederlandsche Bank) 3 (2).
- De Nederlandsche Bank. 2004. "The Cost of Payments." *Quarterly Bulletin* March: 57–64.
- Duca, J. V., and D. D. VanHoose. 2004. "Recent Developments in Understanding the Demand for Money." *Journal of Economics and Business* 56 (4): 247–72.
- Enge, A., and G. Øwre. 2006. "A Retrospective on the Introduction of Prices in the Norwegian Payment System." *Economic Bulletin* (Norges Bank) 77: 162–72.
- Flatraaker, D.-I., and P. Robinson. 1995. "Income, Costs, and Pricing in the Payment System." *Economic Bulletin* (Norges Bank) 66: 321–32.
- Gresvik, O., and G. Øwre. 2003. "Costs and Income in the Norwegian Payment System 2001. An Application of the Activity

- Based Costing Framework.” Working Paper No. 2003/8, Norges Bank.
- Guibourg, G., and B. Segendorff. 2007. “A Note on the Price and Cost Structure of Retail Payment Services in the Swedish Banking Sector 2002.” *Journal of Banking and Finance* 31 (9): 2817–27.
- Hannan, T. 2005. “ATM Surcharge Bans and Bank Market Structure: The Case of Iowa and Its Neighbors.” Working Paper No. 46, Board of Governors of the Federal Reserve System.
- Humphrey, D., A. Kaloudis, and G. Øvre. 2004. “The Future of Cash: Falling Legal Use and Implications for Government Policy.” *Journal of International Financial Markets, Institutions and Money* 14 (3): 221–33.
- Humphrey, D. B., M. Kim, and B. Vale. 2001. “Realizing the Gains from Electronic Payments: Costs, Pricing and Payment Choice.” *Journal of Money, Credit, and Banking* 33 (2): 216–34.
- Humphrey, D., M. Willeson, G. Bergendahl, and T. Lindblom. 2006. “Benefits from a Changing Payment Technology in European Banking.” *Journal of Banking and Finance* 30 (6): 1631–52.
- Jonker, N. 2007. “Payment Instruments as Perceived by Consumers—Results from a Household Survey.” *De Economist* 155 (3): 271–303.
- Klee, E. 2006. “Paper or Plastic? The Effect of Time on the Use of Check and Debit Cards at Grocery Stores.” Working Paper No. 2, Board of Governors of the Federal Reserve System.
- Knittel, C. R., and V. Stango. 2004. “Incompatibility, Product Attributes and Consumer Welfare: Evidence from ATMs.” NBER Working Paper 10962.
- MacKinnon, J. G. 1996. “Numerical Distribution Functions for Unit Root and Cointegration Tests.” *Journal of Applied Econometrics* 11 (6): 601–18.
- Norges Bank. 2000. *Annual Report on Payment Systems*. Norway: Norges Bank.
- Phillips, P. C. B., and S. Ouliaris. 1990. “Asymptotic Properties of Residual Based Tests for Cointegration.” *Econometrica* 58 (1): 165–93.
- Rochet, J.-C., and J. Tirole. 2003. “Platform Competition in Two-Sided Markets.” *Journal of the European Economic Association* 1 (4): 990–1029.

- Scholnick, B., N. Massoud, A. Saunders, S. Carbo, and F. Rodriguez. 2007. "The Economics of Credit Cards, Debit Cards and ATMs: A Survey and Some New Evidence." Forthcoming in *Journal of Banking and Finance*.
- Sveriges Riksbank. 2004. "Prices and Costs in the Swedish Payment System." *Financial Stability Report 2*: 65–73.
- Zinman, J. 2005. "Debit or Credit?" Working Paper, Dartmouth College.

Using Securities Market Information for Bank Supervisory Monitoring*

John Krainer and Jose A. Lopez

Economic Research Department, Federal Reserve Bank of San Francisco

U.S. bank supervisors conduct comprehensive inspections of bank holding companies and assign them a supervisory rating, known as a BOPEC rating prior to 2005, meant to summarize their overall condition. We develop an empirical model of these BOPEC ratings that combines supervisory and securities market information. Securities market variables, such as stock returns and bond yield spreads, improve the model's in-sample fit. Debt market variables provide more information on supervisory ratings for banks closer to default, while equity market variables provide useful information on ratings for banks further from default. The out-of-sample accuracy of the model with securities market variables is little different from that of a model based on supervisory variables alone. However, the model with securities market information identifies additional ratings downgrades, which are of particular importance to bank supervisors who are concerned with systemic risk and contagion.

JEL Codes: G14, G21.

*The views expressed here are those of the authors and not necessarily those of the Federal Reserve Bank of San Francisco or the Federal Reserve System. We thank Rob Bliss for generously sharing his bank holding company debt database with us. We thank Hyun Song Shin (the editor), an anonymous referee, Fred Furlong, Reint Gropp, and seminar participants at the Bank of England, the Basel Committee's Research Task Force on Banking Supervision, and the Chicago Conference on Bank Competition and Structure for helpful suggestions. We thank Judy Peng Fera, Ryan Stever, and Chishen Wei for research assistance. E-mails: Krainer: john.krainer@sf.frb.org; Lopez: jose.a.lopez@sf.frb.org.

1. Introduction

The most comprehensive tool for banking supervision in the United States is the on-site inspection, where a team of supervisors visits an institution and analyzes its operations in detail. For bank holding companies (BHCs), Federal Reserve examiners assign the institution a rating, known as a BOPEC rating up through 2004, at the conclusion of an inspection. The rating summarizes the examiners' opinion of the BHC's overall financial condition. On-site inspections are conducted on roughly an annual basis. Between BHC on-site inspections, examiners engage in off-site monitoring, which largely consists of analyzing quarterly reports from the institutions in question. However, BHCs that have outstanding public securities are also being monitored by their equity shareholders and debtholders. Since their assessments of BHCs have been found to be timely and accurate in previous studies, the aim of this paper is to investigate the effectiveness of BHC securities market data, both from the equity and debt markets, in off-site monitoring models of BOPEC ratings.¹

Securities market prices should, in an ideal world, tell supervisors all they need to know about a BHC's condition and its likelihood of failure. In practice, however, there are various frictions that make our questions worthy of empirical research. First, perceptions of possible government support for a struggling BHC, and the depository safety net in general, might reduce investor incentives to monitor. This would potentially decrease the sensitivity of security prices to changes in BHC conditions. Second, since banks specialize in solving problems of asymmetric information, the loans they hold as assets may be difficult for outside investors to value. This problem, like the first, might make security prices less sensitive to changes in asset value. Finally, supervisors have access to information that BHCs are not normally required to disclose to investors, raising questions of whether securities prices can tell supervisors anything they do not already know.

¹Note that in this paper we focus on supervisory ratings and not defaults, another key supervisory concern. There exists an extensive literature on bank default dating back to Meyer and Pifer (1970), Sinkey (1975), and Pettway and Sinkey (1980).

In this paper, we seek to accomplish three main objectives. First, we investigate the potential contributions of both equity and debt market information to the supervisory monitoring of BHCs using an off-site monitoring model for BOPEC ratings. We measure the contributions of these variables relative to a model based on supervisory data alone.²

The second objective of our study is to investigate the cases when information from the equity or debt markets is relatively more useful. We find an asymmetric contribution from these variables for predicting BOPEC rating changes that depends on how close a BHC's assets are to its default point. That is, in a model of BOPEC changes, the magnitude of the coefficients on equity variables is larger for BHCs further from default, while the coefficients on the debt variables are larger for BHCs close to default.

Finally, our third objective is to conduct a set of realistic forecasting exercises aimed at assessing the value of securities market data to bank supervisors over the last banking cycle. We find little evidence of improved forecast performance after incorporating securities market information into the off-site monitoring model. The performance of BOPEC forecasts based on supervisory data alone is not statistically different from that of BOPEC forecasts generated by the model augmented with securities market data. However, we find that while the forecasts are not different in a statistical sense, they are different in an economic sense. The forecasts based on the model incorporating securities market information identify additional BOPEC rating changes, especially downgrades, of publicly traded BHCs that were not identified by a benchmark model based just on supervisory data. Given a supervisory objective function that values early warnings of potential downgrades, this identification of additional correct BOPEC changes could outweigh the cost of additional false signals.

An extensive academic literature regarding the complementarity of supervisory and financial market monitoring of BHCs and their subsidiary banks already exists; see Flannery (1998) for a survey. The majority of the recent studies in the literature have

²Throughout the paper, we use the term "supervisory data" to mean data generated by supervisors as part of the BHC quarterly reporting process or as part of the supervisory process.

focused on the various uses of bond market data for monitoring. For example, Flannery and Sorescu (1996) and Covitz, Hancock, and Kwast (2004a, 2004b) study how the sensitivity of bank bond prices changed with the passage of the FDIC Improvement Act, an event that arguably lessened the probability that bondholders would be bailed out in case of a bank failure. DeYoung et al. (2001) find that bank subordinated debt prices do not immediately reflect information generated by supervisors, but that this information does enter prices over several quarters. Evanoff and Wall (2002) find that subordinated debt spreads do as well or better than standard capital ratios at explaining supervisory ratings.

Relatively less attention has been paid to the value of equity market data for supervisory monitoring; for surveys of this work, see Krainer and Lopez (2003, 2004). Even less research has focused on combining equity and debt market data, although there are two notable exceptions. Berger, Davies, and Flannery (2000) examine the timeliness and accuracy of supervisory and market assessments of the condition of large BHCs. They find that, after accounting for market assessments, balance sheet variables do not contribute substantially to the modeling of future indicators of BHC performance, such as changes in nonperforming loans, book equity capital, and return on assets. Their findings suggest that supervisors, bond market participants, and equity market participants produce complementary but different information on BHC performance.

Gropp, Vesala, and Vulpes (2004, 2006) examine the ability of equity market variables and subordinated bond spreads for European banks to signal changes in their agency ratings.³ Using ordered logit models at several horizons and a proportional hazard model, they find that both equity-based measures of distance to default and subordinated debt spreads are useful for detecting changes in bank ratings. Interestingly, they find that the distance-to-default measure performs less well closer to default and that subordinated debt spreads seem to have signal value only close to default. The authors argue that their empirical results provide support for the use of securities market information in supervisors' early-warning models.

³See also Gropp and Richards (2001).

Our paper differs from the prior studies in several important ways. As in Berger, Davies, and Flannery (2000), we use data on U.S. bank holding company condition, but our analysis focuses squarely on supervisory ratings as the key measure of BHC condition. Thus, our analysis addresses the benefits of using market information to predict variables that supervisors care most about. Additionally, we extend their in-sample analysis to an out-of-sample forecasting framework that mimics the way supervisors might actually use securities market data as a supervisory tool. Finally, though we find a limited use of securities market information for forecasting BHC condition, we establish the conditions at the BHC under which equity and debt market signals can be useful. Our results provide clear empirical reasoning for the Berger, Davies, and Flannery (2000) finding that supervisory and debt-related assessments of BHCs are related. Thus, our results are a refinement of their finding that supervisory and equity market indicators are not interrelated; i.e., we find that abnormal BHC stock returns do provide useful information with respect to changes in supervisory assessments when BHCs are far from their default points.

The paper proceeds as follows. In section 2, we provide a brief overview of the supervisory process for BHCs in the United States. In section 3, we estimate our proposed BOPEC off-site monitoring model (BOM) using both supervisory and securities market variables. We also examine the differential impact of the securities market variables based on the BHCs' relative distances from their default points. In section 4, we examine the model's out-of-sample performance using a statistical and a supervisory objective function. Section 5 concludes.

2. The U.S. Supervisory Process

The Federal Reserve is the supervisor of bank holding companies in the United States. Full-scope, on-site inspections of BHCs are a key element of this supervisory process. These inspections are generally conducted on an annual basis, particularly for the case of large and complex BHCs.⁴ Limited and targeted inspections that may or

⁴A complex BHC is defined as one with material credit-extending nonbank subsidiaries or debt outstanding to the general public. See DeFerrari and Palmer

may not be conducted on-site are also carried out. In this paper, we focus on full-scope, on-site inspections since they provide the most comprehensive supervisory assessments of BHCs.

At the conclusion of an inspection, supervisors assign the BHC a numerical rating, which prior to 2005 was called a composite BOPEC rating, that summarizes their opinion of the BHC's overall health and financial condition.⁵ The BOPEC acronym stands for the five key areas of supervisory concern: the condition of the BHC's Bank subsidiaries, Other nonbank subsidiaries, Parent company, Earnings, and Capital adequacy. A BOPEC rating of 1 is the best rating, with a rating of 5 being the worst rating short of closure. A rating of 1 or 2 indicates that the BHC is not considered to be of supervisory concern. Note that BOPEC ratings, as well as all other inspection materials, are confidential and are not made publicly available.

Between on-site inspections, when private supervisory information cannot be gathered as readily, supervisors monitor BHCs using quarterly regulatory reports filed by BHCs and their subsidiary banks. In addition, the supervisory CAMELS ratings assigned to banks within the holding company are used for monitoring the parent BHCs. As with BOPEC ratings, CAMELS ratings are confidential ratings that are assigned after a bank examination. The acronym refers to the six key areas of concern: the bank's Capital adequacy, Asset quality, Management, Earnings, Liquidity, and Sensitivity to risk. The composite CAMELS rating also ranges in integer value from 1 to 5 in decreasing order (i.e., banks that perform best are assigned a rating of 1). Since the condition of a BHC is closely related to the condition of its subsidiary banks, the off-site BHC surveillance program includes monitoring recently assigned CAMELS ratings.

As with on-site BHC inspections, on-site bank examinations occur at approximately a yearly frequency, which is long enough for the gathered supervisory information to decay and become less

(2001) for an overview of the supervisory process for large, complex banking organizations.

⁵Note that in January 2005, the BOPEC rating system was replaced by the RFI/C(D) rating system; see Board of Governors of the Federal Reserve System (2004). For an international survey of supervisory bank rating systems, see Sahajwala and van den Bergh (2000).

representative of the bank's condition.⁶ To address this issue, the Federal Reserve instituted an off-site monitoring system for banks, known as the System for Estimating Examiner Ratings (SEER), in 1993. The SEER system actually consists of two separate models that forecast bank failures over a two-year horizon as well as bank CAMELS ratings for the next quarter. The model that we are most interested in here is the latter, which is an ordered logit model with five categories corresponding to the five possible values of the CAMELS rating. The model is estimated every quarter in order to reflect the most recent relationship between the selected financial ratios and the two most recent quarters of CAMELS ratings. Significant changes in a bank's CAMELS rating as forecasted by the SEER model could be sufficient to warrant closer monitoring of the bank. The off-site BHC surveillance program also explicitly monitors the SEER model's forecasted CAMELS ratings.

2.1 The BOPEC Ratings Sample

The core database for our analysis is the set of supervisory BOPEC ratings assigned after on-site, full-scope inspections between the first quarter of 1990 and the second quarter of 1998. The sample endpoint is dictated by the availability of the bond data set.⁷ Our sample of BOPEC ratings is further refined to include only those inspections of top-tier BHCs with identifiable lead banks, four quarters of prior supervisory data, and at least one prior BOPEC rating. We focus on top-tier BHCs since they are typically the legal entities within the banking group that issues publicly traded equity. The lead-bank designation is often provided by banks in their regulatory filings. When such self-reporting is not available, we assign the lead-bank designation to the largest bank within the group. We need the BHCs in our sample to have identifiable lead banks in order to directly link their BOPEC ratings to their lead bank's CAMELS

⁶See Cole and Gunther (1998) as well as Hirtle and Lopez (1999) for further discussion of this issue.

⁷We are grateful to Rob Bliss for sharing his BHC bond database with us. A complete description of the database is presented in Bliss and Flannery (2001). The last quarter of bond data is the first quarter of 1998, which aligns with the second quarter of 1998 in our modeling framework.

ratings.⁸ Finally, we require each BHC to have at least four quarters of supervisory data and a lagged BOPEC rating in order to avoid issues regarding de novo BHCs and new BHCs arising from mergers. In addition, four quarters of supervisory data are required to calculate certain explanatory variables for the model described below.

Table 1 summarizes our sample of 3,010 BOPEC ratings assigned to 1,034 unique entities. Almost 65 percent of the BHCs in the sample are relatively small, with less than \$1 billion in total assets. Slightly more inspections occurred in the first half of the sample than in the second half, reflecting consolidation in the U.S. banking sector. For publicly traded BHCs, there are 1,291 BOPEC assignments corresponding to 363 unique entities. Note that public BHCs are generally larger than privately held BHCs, with a greater percentage having total assets ranging between \$1 billion and \$100 billion. Of the 41 BOPECs assigned to the largest BHCs, 39 are of public BHCs. With respect to BHCs with public debt outstanding, this subsample contains 309 BOPEC ratings corresponding to 63 unique BHCs. Again, these BHCs are typically larger than those in the full sample, with almost all BHCs having between \$1 billion and \$100 billion in assets. Finally, there are 283 BOPEC ratings corresponding to 58 unique BHCs that have both publicly traded equity and debt outstanding during the sample period.

Tables 2A and 2B present the distribution of BOPEC ratings assigned in each year for all BHCs, as well as for BHCs with both publicly traded equity and publicly traded bonds. The distributions in the two tables share many common features. The majority of the ratings fall in the upper two categories, indicating that a BHC's financial condition and risk profile are of little supervisory concern. For the full sample, while the distribution of ratings fluctuates over time, the percentage of ratings in the top two categories never falls below 63 percent. The maximum value is 96 percent in 1998. In contrast, for larger BHCs with both public equity and debt, the percentage of ratings in the upper two categories ranges from 47

⁸Note that this restriction does not imply that we limited the sample to single-bank BHCs. We simply focus on the CAMELS rating for a BHC's lead bank, whether self-identified or identified by asset size.

**Table 1. Asset Size Characteristics of the BHCs
in the BOPEC Sample**

	1990–98	1990–94	1995–98
Total # Ratings	3,010	1,735	1,275
Asset Size:			
Assets > \$100b	41	13	28
\$1b < Assets < \$100b	1,019	594	425
Assets < \$1b	1,950	1,128	822
Total # Ratings, Public BHCs	1,291	741	550
Asset Size:			
Assets > \$100b	39	13	26
\$1b < Assets < \$100b	807	487	320
Assets < \$1b	445	241	204
Total # Ratings, BHCs with Public Debt	309	174	135
Asset Size:			
Assets > \$100b	37	11	26
\$1b < Assets < \$100b	270	163	107
Assets < \$1b	2	0	2
Total # Ratings, Public BHCs with Public Debt	283	163	120
Asset Size:			
Assets > \$100b	36	11	25
\$1b < Assets < \$100b	247	152	95
Assets < \$1b	0	0	0
Note: The data sample spans the period from the first quarter of 1990 to the second quarter of 1998. The definition of a bank holding company (BHC) used in this table is the definition used in constructing our data set; i.e., a top-tier BHC with an identifiable lead bank, four quarters of available regulatory reporting data, and a lagged BOPEC rating. Public debt here refers to publicly traded bonds as listed in the Warga/Lehmann data set and used by Bliss and Flannery (2001).			

percent in 1991 to 100 percent from 1996 onward. Note that there are relatively few inspections culminating in a BOPEC rating of 4 or worse, with the exception of the early 1990s. This likely reflects the fact that both supervisors and bankers actively try to prevent this

Table 2A. All BOPEC Ratings in the Sample

	BOPEC Rating % of Total			
	1	2	3	4–5
1990	16%	52%	21%	10%
1991	16%	47%	25%	12%
1992	15%	52%	20%	14%
1993	24%	55%	14%	7%
1994	34%	53%	8%	5%
1995	36%	52%	8%	5%
1996	47%	47%	5%	1%
1997	47%	48%	4%	0%
1998	47%	49%	3%	0%
Total	31%	51%	12%	6%
Note: The data sample spans the period from the first quarter of 1990 to the second quarter of 1998.				

Table 2B. All BOPEC Ratings for BHCs with Public Equity and Bonds in the Sample

	BOPEC Rating % of Total			
	1	2	3	4–5
1990	21%	55%	17%	7%
1991	12%	35%	31%	23%
1992	18%	41%	26%	15%
1993	24%	68%	3%	6%
1994	45%	50%	5%	0%
1995	44%	56%	0%	0%
1996	60%	40%	0%	0%
1997	58%	42%	0%	0%
1998	68%	32%	0%	0%
Total	38%	48%	9%	5%
Note: The data sample spans the period from the first quarter of 1990 to the second quarter of 1998.				

outcome, as well as the general good health of the banking sector in the mid and late 1990s.

3. Multivariate Analysis Using the Ordered Logit Model

Our proposed BOPEC off-site monitoring (BOM) model is an ordered logit model and is similar in structure to the SEER model for CAMELS ratings; see Krainer and Lopez (2003, 2004) for further details.⁹ The model assumes that the supervisory rating assigned to BHC i in quarter t is a continuous variable, denoted BP_{it}^* , for which only discrete values, denoted BP_{it} , are observed. Recall that lower values of BP_{it}^* and BP_{it} correspond to better supervisory ratings. We model the continuous variable as

$$BP_{it}^* = \beta x_{it-2} + \varepsilon_{it}. \quad (1)$$

In this specification, x_{it-2} is a $(k \times 1)$ vector of supervisory variables unique to BHC i observed two quarters prior to the BOPEC assignment. We chose to lag the supervisory variables by two quarters because, in real time, these are often the most recent data available at the time of inspection.¹⁰

For firms with publicly traded securities, we augment equation (1) with securities market data, which is denoted as z_{it-1} ; i.e.,

$$BP_{it}^* = \beta x_{it-2} + \pi z_{it-1} + \varepsilon_{it}. \quad (2)$$

The vector z_{it-1} contains securities market variables—equity market variables, debt market variables, or both depending on the model—corresponding to BHC i at time $t - 1$, one quarter prior to the BOPEC assignment. The supervisory variables and the securities market variables enter into the model with different lags because securities market data are available on a more timely basis than are the supervisory variables that we use. The error term ε_{it} has a standard logistic distribution. Note that we estimate four BOM

⁹Recall that Gropp, Vesala, and Vulpes (2004, 2006) used logit and proportional hazard models, which are better suited for examination of the timing and determinants of banks' Fitch/IBCA ratings changes.

¹⁰See Gunther and Moore (2003) for further discussion.

models corresponding to the four alternative z_{it-1} vectors using BOPEC samples based on the availability of securities market variables. For example, we estimate the core model without securities market variables (i.e., $z_{it-1} = 0$) on all BOPEC ratings, but we estimate the BOM model using both sets of market variables only on the BOPEC sample for which both sets of variables are available.

3.1 Supervisory Variables

The choice of which supervisory variables to include in x_{it-2} is challenging. No simple behavioral models exist of how supervisors assign BOPEC ratings. As mentioned, there are more than 800 variables at the supervisors' disposal for this purpose. For this study, we selected nine explanatory variables that were judged to be reasonable proxies for the five components of the BOPEC rating. In addition, we included lagged BOPEC ratings in the specification to control for ratings persistence. As in Krainer and Lopez (2003, 2004), we chose a parsimonious specification in the hopes of generating reasonable out-of-sample forecasts.

Our supervisory variables are summarized in table 3. The first variable is the natural log of total BHC assets, which is our control variable for BHC size. The next four variables are used to capture the supervisory concerns regarding the BHC's bank subsidiaries, as summarized in the "B" component of the rating. The second variable is the CAMELS rating of the BHC's lead bank. The third variable is the ratio of the BHC's nonperforming loans, nonaccrual loans, and other real estate owned to its total assets. This variable proxies for the health and performance of the BHC's loan portfolio. The fourth variable is the ratio of the BHC's allowances for losses on loans and leases to its total loans, another proxy for the health and performance of the BHC's loan portfolio.

The fifth variable is an indicator of whether the BHC has a section 20 subsidiary, which is a subsidiary that can engage in securities activities that commercial banks were generally not permitted to engage in before the passage of the Gramm-Leach-Bliley Act of 1999. This variable is a proxy for the scale of the BHC's nonbank activities and thus speaks to the "O" component of the BOPEC rating. We also include as the sixth variable the ratio of a BHC's trading assets

Table 3. Summary Statistics for Financial Statement and Supervisory Variables

	Mean	Std. Dev.	25th Pctile.	Median	75th Pctile.
Assets (\$m)	\$6,336	\$23,700	\$250	\$493	\$2,068
CAMELS Rating	1.94	0.80	1	2	3
Nonperforming Loans/Assets	1.97%	1.87%	0.87%	1.47%	2.41%
Allowances for Loan Losses/ Assets	0.41%	0.69%	0.09%	0.21%	0.44%
Section 20 Subsidiary	0.04	0.19	0.00	0.00	0.00
Trading Assets/ Assets	1.10%	42.27%	0.0%	0.0%	0.0%
Double Leverage	55.24%	108.34%	7.29%	43.27%	98.21%
Return on Average Assets	0.82%	0.97%	0.66%	0.98%	1.22%
Equity Capital/Assets	8.18%	2.47%	6.71%	7.88%	9.26%

to its total assets as a proxy of its nonbanking activities, whether conducted in banking or nonbanking subsidiaries.¹¹

The seventh variable is the so-called “double leverage” ratio between the BHC and its lead bank, which is the ratio of the lead bank’s equity capital to that of the parent’s equity capital. This variable provides a measure of the soundness of the parent BHC, indicating the extent to which the parent’s equity capital can be used to buffer against damage to the lead bank’s equity capital. We use this variable as a proxy for the condition of the parent BHC as summarized in the “P” component of the BOPEC rating. The eighth variable is the BHC’s return on average assets (ROAA), defined as the ratio of the four-quarter average of the BHC’s net income to the four-quarter average of its assets. This variable is used to proxy for

¹¹Note that the trading assets variable as currently reported first became available in the first quarter of 1995. Before then, we proxy for BHC trading assets using the sum of the self-reported replacement cost of interest rate and foreign exchange derivative contracts.

the “E” component of the BOPEC rating. The ninth variable is the BHC’s ratio of equity capital to its total assets. This variable is used to proxy for the “C” component of the BOPEC rating.¹²

3.1.1 Equity Market Variables

The two equity market variables used in this study are based on BHC stock returns over the quarter prior to a BOPEC assignment. The motivation behind using equity market data lies in the conjecture that there is some agreement between BHC stock market investors and supervisors on what determines healthy financial condition. Stock market investors clearly are not trying to forecast BOPEC ratings. However, if the same financial developments that lead to supervisory rating changes also lead to changes in expected stock returns, it is possible that supervisors could use stock market signals as an additional off-site monitoring signal.

As per Campbell, Lo, and Mackinlay (1997), we decompose monthly BHC stock returns, denoted as R_{it} , before a BOPEC assignment into a systematic return, denoted SR_{it} , and an idiosyncratic, or abnormal, return, denoted AR_{it} . The decomposition is based on a two-factor market model; i.e.,

$$R_{it} = \alpha + \beta_1 R_{mt} + \beta_2 f_t + \nu_{it}, \quad (3)$$

where R_{mt} is the monthly return on the CRSP value-weighted index, f_t is the monthly change in the federal funds rate, and ν_{it} is a normally distributed error term. For each BOPEC assignment in our sample, the model’s parameters are estimated using monthly data over a period of at least two years leading up to one year prior to the assignment of the BOPEC rating. When available, we used up to five years of monthly data for the estimation window. In the twelve-month event window leading up to the assignment, the systematic portion of the return, SR_{it} , is calculated as $SR_{it} = \hat{\alpha} + \hat{\beta}_1 R_{mt} + \hat{\beta}_2 f_t$. The corresponding abnormal return, AR_{it} , is simply the difference between the realized return and the systematic return, $AR_{it} = R_{it} - SR_{it}$. The stock market variables

¹²A variety of capital measures have been used in previous studies, such as Estrella, Park, and Peristiani (2000) and Evanoff and Wall (2001). We chose a simple measure to facilitate comparison over the entire period.

are constructed to detect changes in BHC condition that occur between inspections and are relevant to an eventual BOPEC assignment but are not yet embedded in the available supervisory data. Thus, the stock market variables are used in event-study fashion, where we cumulate systematic and abnormal returns over the three-month window between quarter $t - 2$, when the latest supervisory data are available, and quarter $t - 1$, the quarter before the assignment.

In the empirical work to follow, the quarterly cumulative return at time $t - 1$, CR_{it-1} , is the actual return between month $t - 6$ and month $t - 3$. As an example, consider a BOPEC assignment in the last quarter of the year. The cumulative return is the sum of the three monthly returns from July to September. The systematic and idiosyncratic returns are formed in the same way. To permit comparison across BHCs and across time, we standardize the cumulative returns using the estimated standard errors for CAR_{it-1} . The standardized form of these variables is

$$\frac{CR_{it-1}}{\sqrt{\text{var}(CAR_{it-1})}} = \frac{CSR_{it-1}}{\sqrt{\text{var}(CAR_{it-1})}} + \frac{CAR_{it-1}}{\sqrt{\text{var}(CAR_{it-1})}}, \quad (4)$$

or equivalently, $SCR_{it-1} = SCSR_{it-1} + SCAR_{it-1}$.

BHC stock price changes that are large in magnitude with respect to general market activity may signal changes in condition that will eventually lead to a ratings change. The SCAR variable is designed specifically for identifying which stock price changes are large based on our normality assumption. However, relying exclusively on SCARs for market signals may cause us to miss important information available from the broader stock market. For example, an economy-wide shock that lowers returns for all stocks might not translate into abnormally negative returns for any particular BHC but could very well be an early indicator of changes in all supervisory ratings. To address this concern, we include the SCSR variable in our regressions. Hence, $z_{Eit} = [SCSR_{it}, SCAR_{it}]$.

It is important to note that there are circumstances where stock return data, especially if not normalized in some way, could be misleading from a supervisory perspective. For example, a BHC close to default has incentive to increase its risk profile, and the call option feature of the equity claim should cause the stock price to rise. In this scenario, supervisors could potentially mistake the stock

market signal of greater risk for a signal of improved condition. To address this concern, Gropp, Vesala, and Vulpes (2004) advocate using an equity-based distance-to-default measure that accounts for this increased risk. In contrast, we focus on the cumulative return measures described above, because these variables are more appropriate for forecasting exercises, where market signals are assessed not just by their size but by their relative precision (i.e., their standard errors).

3.1.2 Debt Market Variables

The debt market variable used in this study is adjusted bond yields from the Warga/Lehmann Brothers Corporate Bond Database, as used by Bliss and Flannery (2001) (henceforth, BF). Note that this database includes both subordinated and nonsubordinated BHC debt.¹³ There are two issues that need to be confronted before using the bond data. First, in cases where a BHC has multiple outstanding bonds, it is necessary to compress their market signals into a single observation. Following BF, we use weighted-average BHC bond yields, where the weights correspond to the size of the issue relative to the BHC's total amount of bonds outstanding in the quarter. Second, as with the equity market variables, we would like to have some measure of what constitutes an abnormal change in bond yield. Here, we follow the BF procedure of computing debt spreads from bond price indices based on eleven Moody's ratings categories and three term-to-maturity categories.¹⁴ The BF indices allow us to study debt yields relative to an index of similar bonds drawn from all industries. We also adjusted the BHC yield spreads to account for their last assigned BOPEC ratings.

To summarize, for BHC i with BOPEC rating j at time t , we define the yield on a bond (or a weighted average of several bonds) with terms k (i.e., maturity and Moody's rating) as y_{ijkt} . We then

¹³For a discussion of the market for BHC subordinated debt in the United States, see Board of Governors of the Federal Reserve System (1999), Hancock and Kwast (2001), Basel Committee on Banking Supervision (2003), and Goyal (2005).

¹⁴The "+" or "-" qualifiers attached to the basic rating definitions are suppressed. The maturity buckets are less than five years, five to ten years, and greater than ten years.

constructed the yield spread s_{ijkt} relative to the corresponding BF based on terms k ; i.e.,

$$s_{ijkt} = y_{ijkt} - \bar{y}_{kt}, \quad (5)$$

where $\bar{y}_{kt}$ is the yield on an index of like-termed bonds. We then further adjusted the yield spread to account for the BHC's last assigned BOPEC rating j ; i.e.,

$$d_{ijkt} = s_{ijkt} - \bar{s}_{jt}, \quad (6)$$

where $\bar{s}_{jt}$ is the median yield spread for all BHCs with BOPEC rating j at time t and publicly traded debt. In our empirical work, we found that these adjusted yield spreads appear to have more predictive power than the yield spreads based just on the BF indices; hence, $z_{Dit-1} = d_{ijkt-1}$.

3.2 Empirical Results

We estimate the BOM model over the four samples of BOPEC ratings. These four samples are the sample consisting of BOPEC ratings, the sample of ratings assigned to BHCs with publicly traded equity, the sample of ratings assigned to BHCs with publicly traded debt, and the sample of ratings assigned with both types of securities. The sample sizes are 3,010; 1,291; 309; and 283 observations, respectively. Note that the sample size decreases by more than a factor of ten from the largest to the smallest samples. The results are presented in tables 4A and 4B.

The basic results for the core model containing just supervisory variables are similar to those reported in Krainer and Lopez (2004). The lagged BOPEC rating and the current bank CAMELS rating are important drivers of the rating process; higher (or worse) values of these ratings lead to higher BOPEC ratings. Higher total assets are correlated with lower (i.e., better) ratings. High levels of problem loans and allowances tend to lead to worse BOPEC ratings. High levels of ROA and capital tend to lead to better BOPEC ratings.¹⁵

¹⁵We conducted two important robustness checks for this sample. First, regarding parameter stability across the sample period, we split the sample period in

Table 4A. BOM Model Estimation Results

	Full Sample		Equity Sample	
	Coefficients	<i>p</i> -value	Coefficients	<i>p</i> -value
Lagged BOPEC	1.292*	0.00	1.551*	0.00
CAMELS	1.223*	0.00	0.685*	0.00
Total Assets	−0.247*	0.00	−0.182*	0.00
Problem Loans	48.050*	0.00	50.208*	0.00
Allowances	56.676*	0.00	59.315*	0.01
Trading Assets	0.004	0.49	−1.502	0.42
Section 20 Sub.	1.819*	0.00	0.272	0.36
Double Leverage	0.054	0.25	−0.282	0.16
ROA	−1.015*	0.00	−0.567	0.06
Equity Capital	−22.103*	0.00	−45.280*	0.00
SCSR	—	—	−0.831*	0.00
SCAR	—	—	−0.523*	0.00
Adj. Yield Spread	—	—	—	—
Observations	3,010		1,291	
Pseudo <i>R</i> ²	0.47		0.48	

Note: The model estimated here across the four subsamples is summarized in equation (1). The results for the interacted supervisory variables and fixed effects are not reported, in order to conserve space. The models contained fixed effects for BHCs with public equity, BHCs with public equity for which SCAR and SCSR variables could not be properly calculated, and BHCs with public debt. The sample period range is 1990:Q1 to 1998:Q2. The model was estimated using robust standard errors and adjusting for clustered observations based on unique BHCs. A * denotes significance at the 5 percent level.

half into a “banking crisis” subsample (1990:Q1 to 1994:Q4) and a “banking recovery” subsample (1994:Q2 to 1998:Q2). The estimated coefficients in the core and extended models did not change much over the two subsamples. All coefficient estimates that were significantly different from zero in the full sample remained so and with the same sign in the subsamples. Second, to account for possible BHC merger effects during the period, we dropped all observations where the BHC had been involved in a merger, either as an acquirer or as an acquired institution, one year prior to its BOPEC assignment. This resulted in a loss of 771 observations, or about 26 percent of our sample. The coefficient estimates are very similar to those reported in tables 4A and 4B. Detailed results from these robustness checks are available from the authors upon request.

Table 4B. BOM Model Estimation Results

	Debt Sample		Equity and Debt Sample	
	Coefficients	<i>p</i> -value	Coefficients	<i>p</i> -value
Lagged BOPEC	2.144*	0.00	2.587*	0.00
CAMELS	0.427	0.08	0.465	0.11
Total Assets	0.119	0.54	0.092	0.68
Problem Loans	80.497*	0.00	95.568*	0.00
Allowances	2.313	0.95	14.104	0.63
Trading Assets	−5.865*	0.01	−4.803*	0.05
Section 20 Sub.	0.225	0.57	0.190	0.66
Double Leverage	−0.529	0.26	−0.224	0.63
ROA	−0.840	0.14	−0.509	0.31
Equity Capital	−60.633	0.00	−69.942*	0.00
SCSR	−	−	−0.552*	0.04
SCAR	−	−	−0.972*	0.00
Adj. Yield Spread	2.688*	0.00	2.382	0.07
Observations	309		283	
Pseudo <i>R</i> ²	0.48		0.58	
Note: The model estimated here across the four subsamples is summarized in equation (1). The results for the interacted supervisory variables and fixed effects are not reported, in order to conserve space. The models contained fixed effects for BHCs with public equity, BHCs with public equity for which SCAR and SCSR variables could not be properly calculated, and BHCs with public debt. The sample period range is 1990:Q1 to 1998:Q2. The model was estimated using robust standard errors and adjusting for clustered observations based on unique BHCs. A * denotes significance at the 5 percent level.				

For the sample associated with publicly traded BHCs, the empirical results for the supervisory variables are practically identical. The two equity market variables have negative signs as expected; i.e., increases in either systemic or idiosyncratic stock returns tend to be associated with better BHC performance and better (i.e., lower) BOPEC ratings.

For the sample associated with BHCs with publicly traded debt, the empirical results for the supervisory variables change somewhat. Total assets and allowances are no longer significant drivers of BOPEC ratings, while trading assets are significant, with higher values leading to better ratings. The *p*-values on current bank CAMELS

ratings and ROA increase beyond the 5 percent significance level. However, lagged BOPEC ratings, problem loans, and equity capital remain important explanatory variables. For the adjusted yield spread variable, the positive sign is also in line with expectations; i.e., higher yield spreads relative to the corresponding yield index are associated with worse BHC performance and worse (i.e., higher) BOPEC ratings.¹⁶ The empirical results of the sample for BHCs with both types of publicly traded securities are quite similar to those of the debt sample. Notably, the adjusted yield spread has a positive coefficient with a p -value of 7 percent, just outside of our defined significance level.

These in-sample results clearly suggest that securities markets provide information complementary to that generated by supervisors. Interestingly, the results for BHCs with both public equity and debt seem to suggest that equity market variables may be more useful than debt market variables. In the next subsection, we examine more directly the relative importance of these two sources of market information.

3.3 The Relative Importance of Equity and Debt Market Information

As noted in the introduction, a primary motivation for using both equity and debt market data in a supervisory monitoring model is that no single information source is likely to dominate the other in all states of the world. How, then, might these two types of securities market information differ? In which states of the world are the different sources of information most useful? We might expect that the residual claim feature of equity implies that equity market investors would be good at predicting BOPEC rating changes when BHC asset values are relatively far from the default point. For

¹⁶We estimated versions of the models above that allowed the coefficients on securities market variables to differ according to whether the market might perceive the institution to be “too big to fail.” For this analysis, we classified a BHC as possibly being too big to fail if it was one of the five largest BHCs by assets in a given year. We found that equity market prices have the same explanatory value for BOPEC ratings for this class of BHCs as for the rest of the sample. Debt prices for the largest BHCs were considerably less useful for signaling changes in condition, although this effect was not measured precisely. Results from this robustness check are available from the authors upon request.

such total asset values, changes in value correspond one-to-one with changes in equity value. Debt market investors, by contrast, might be more likely to predict BOPEC rating changes when asset values are relatively close to or below the default point, as that range for asset values is where bondholders are most at risk for taking losses.

To examine the relationship between securities market signals and how close a BHC might be to default, we explicitly calculate a default point for each public BHC at each point in time and examine how its interaction with the securities market variables affects the supervisory ratings. For this exercise, we use the Ronn-Verma (1986) model for estimating the value of a firm's assets. Given assumptions on the stochastic process describing changes in asset value, the firm's equity is modeled as a call option on those assets. This framework gives rise to a distance-to-default (DTD) measure, which is the difference between the estimated market value of the assets and the equity, scaled by estimated standard deviation of the rate of return on the assets. Note that the construction of this measure restricts our sample to public BHCs.

Working again with ordered logit models within the BOM framework, we create indicator variables for whether a BHC's DTD measure is within a certain percentile of the overall sample distribution of DTD measures. We constructed nine indicator variables corresponding to the first nine deciles of the DTD measure's unconditional distribution. We then interact each indicator variable with the securities market variables to generate nine sets of regression results. To focus the analysis, we use a more parsimonious model that has the lagged BOPEC rating as the only supervisory variable; i.e.,

$$BP_{it}^* = \beta BP_{it-2} + (\pi + \alpha_n I_{nit-1}) z_{it-1} + \varepsilon_{it}, \quad (7)$$

where z_{it-1} is the securities market variable in question and I_{nit-1} is equal to one if BHC i 's DTD measure is in the n th percentile at time $t - 1$, and zero otherwise. Our main object of interest is whether the coefficients on the securities market variables are different in magnitude depending on how close the BHC is to its default point.

The estimation results for the 1,291 BOPEC ratings for BHCs with publicly traded equity are presented in table 5A. For the SCAR

Table 5A. Estimation Results for the Simplified Rating Model Using DTD Interactions with SCARs

DTD Pctile.	π for SCAR Variable	p -value	α_n Interaction with DTD Measure	p -value	$\pi + \alpha_n$	p -value	Pseudo R^2
10th Pctile.	−0.421	0.00	0.137	0.48	−0.284	0.14	0.31
20th Pctile.	−0.406	0.00	0.022	0.89	−0.384	0.01	0.31
30th Pctile.	−0.423	0.00	0.054	0.71	−0.369	0.00	0.31
40th Pctile.	−0.431	0.00	0.058	0.67	−0.373	0.00	0.31
50th Pctile.	−0.418	0.00	0.030	0.84	−0.388	0.00	0.31
60th Pctile.	−0.313	0.00	−0.114	0.42	−0.427	0.00	0.31
70th Pctile.	−0.271	0.04	−0.155	0.32	−0.426	0.00	0.31
80th Pctile.	−0.286	0.09	−0.126	0.49	−0.412	0.00	0.31
90th Pctile.	−0.197	0.41	−0.214	0.40	−0.411	0.00	0.31
Note: The coefficient estimates presented here are for the regression in equation (9). The securities market variable of interest is the BHC SCAR as used previously. The interacted indicator variable is equal to one if the BHC’s distance-to-default (DTD) measure is in the n th percentile of its unconditional distribution. All models are estimated with a sample of 1,291 observations.							

variable described previously, we expect its aggregate coefficient to be negative, suggesting that an increased SCAR that reflects improved BHC condition should lead to a better BOPEC rating. We expect the disaggregated π coefficient to be negative across all nine regressions. We expect the α_n coefficient to have its largest impact for the lowest DTD percentiles; i.e., the stock market signals for BHCs relatively closer to their default point should be less clear for supervisory purposes. As shown in the table, the estimated π coefficients are invariably negative and statistically significant for all but the ninetieth DTD percentile. The α_n coefficients are estimated to be positive for closer-to-default deciles and become negative as the default decile is increased, but they are not significantly different from zero. Taken as a whole, the equity market impact measured as the $\pi + \alpha_n$ values is negative, statistically significant for all but the tenth percentile specification, and in a relatively narrow range. Thus, as expected, equity market signals are weakest for BHCs relatively close to the default point.

In table 5B, the z_{it-1} variable is the adjusted BHC yield spread. The regressions are based on 283 BOPEC ratings assigned to BHCs

Table 5B. Estimation Results for the Simplified Rating Model Using DTD Interactions with the Adjusted Yield Spreads

DTD Pctile.	π for Adj. Yield Spread	p -value	α_n Interaction with DTD Measure	p -value	$\pi + \alpha_n$	p -value	Pseudo R^2
10th Pctile.	0.626	0.56	8.906	0.00	9.532	0.00	0.39
20th Pctile.	0.245	0.86	1.445	0.56	1.690	0.42	0.39
30th Pctile.	1.180	0.55	-0.651	0.77	0.529	0.61	0.39
40th Pctile.	1.956	0.22	-1.726	0.36	0.230	0.82	0.39
50th Pctile.	2.894	0.01	-3.099	0.05	-0.205	0.85	0.40
60th Pctile.	1.646	0.16	-1.095	0.54	0.551	0.66	0.39
70th Pctile.	1.859	0.21	-1.315	0.52	0.544	0.66	0.39
80th Pctile.	0.116	0.06	0.733	0.76	0.849	0.46	0.39
90th Pctile.	-1.631	0.50	2.499	0.33	0.868	0.43	0.39

Note: The coefficient estimates presented here are for the regression in equation (9). The securities market variable of interest is the BHC adjusted yield spread as used previously. The interacted indicator variable is equal to one if the BHC's distance-to-default (DTD) measure is in the n th percentile of its unconditional distribution. All models are estimated with a sample of 283 observations.

with both public equity and debt. For this variable, we expect the π coefficients to be positive, such that an increase in adjusted yield spreads suggests a worsening of BHC condition and probably a worsening of the BOPEC rating. We expect the α_n coefficients to be decreasing and diminishing the overall impact of the market signal as the DTD measure increases; i.e., the debt market signal should be strongest for BHCs closer to their default points. In the table, we show that, in fact, the coefficient of the estimated $(\pi + \alpha_1)$ coefficient for the percentile closest to the default point is more than four times larger than the coefficients on the other percentiles. Furthermore, that first coefficient is the only one that is statistically significant, due to the statistical significance of the α_1 coefficient. The α_n coefficients for the higher DTD percentiles are not statistically significant.

In summary, our simplified ordered logit models suggest an asymmetric contribution of equity and debt market signals to explaining BOPEC ratings, and the asymmetry depends on how close the BHC is to its default point. The impact of the equity market signals

$(\pi + \alpha_n)$ is largest in magnitude for the BHCs further from default, while the impact of the adjusted yield spreads is much larger in magnitude for BHCs very close to default.¹⁷ These results are in line with theoretical expectations as well as prior empirical research. In particular, they provide further empirical insight into the finding by Berger, Davies, and Flannery (2000) that supervisory and debt-related assessments of BHCs are related. Furthermore, the results provide a refinement of their finding that supervisory and equity market assessments are not interrelated; that is, abnormal stock returns provide useful information with respect to changes in supervisory assessments only for BHCs further from default.

4. Out-of-Sample Forecast Performance

While the BOM model's in-sample fit and inference is interesting and important, a stricter test of its usefulness for supervisory purposes is its out-of-sample forecast accuracy. As in Krainer and Lopez (2003, 2004), we propose to estimate our model over rolling, four-quarter sample periods and evaluate their out-of-sample forecast performance. Clearly, the number of BOPEC ratings assigned with these rolling subsamples will be small relative to the full sample, but we accept the smaller estimation samples in order to simulate how securities market data might be used in practice by supervisors.¹⁸

However, a major difficulty in applying this forecasting approach to the four BOM models discussed in section 3 is that the number of observations in any given rolling subsample could be too small for the models with securities market data. That is, the number of BHCs with publicly traded securities within a given forecasting

¹⁷This result regarding adjusted BHC yield spreads seems to be consistent with Hanweck and Spellman (2002), who found that BHC subordinated debt yields were sensitive to solvency concerns if investors' forbearance expectations (i.e., time to closure or default) were less than one year.

¹⁸The need for four quarters of prior data also causes us to lose some of the early BOPEC assignments, mainly because we do not have bond market data prior to 1990. The sample used in the forecasting exercise consists of 2,878 BOPEC assignments. In this sample there were 638 upgrades, 331 downgrades, and 1,909 no-change inspections corresponding to 509, 290, and 866 unique BHCs, respectively.

window may be too small to permit model estimation. This is particularly true for the two samples of BHCs with publicly traded debt. To address this issue, we propose to pool the data and estimate a single model using a technique outlined in Griliches (1986). The underlying assumption to justify this pooling is that all ratings are generated by the same underlying process. In other words, for the equity sample, publicly traded BHCs are allowed to be different from private BHCs, but these differences are assumed to be observable. Thus, the relationship between a variable—say, the capital ratio—and the BOPEC rating is the same across publicly traded and private BHCs, holding the other covariates in the model fixed. If this assumption is satisfied, then we can treat the absence of a variable—say, stock market returns for BHCs that are not publicly traded—as though it were simply a case of missing data. We can then replace the missing values with the in-sample mean of the stock market return variable and estimate the model in equation (2) on the pooled sample, including fixed effects for the observations wherever we make this data adjustment. Formally, the estimation equation becomes

$$BP_{it}^* = \beta x_{it-2} + \pi z_{it-1} + I_{Eit-1} + I_{Dit-1} + \varepsilon_{it}, \quad (8)$$

where the indicator variables I_E and I_D take on the value of 1 if BHC i has publicly traded equity and debt, respectively.

This estimation approach has two advantages. First, the coefficient estimates on the securities market variables should not be affected, since we are simply replacing the missing values with the variables' mean. Second, the pooling should allow us to get more efficient estimates of the coefficients on the supervisory variables because we can use all available observations.

To confirm that this procedure does not impact the empirical results, we conducted two robustness tests. First, we checked whether the distribution of BOPEC ratings for publicly traded BHCs differed from that of private BHCs. On average, we found no statistical differences, suggesting that the subsamples based on public equity and/or debt are not markedly different from the whole sample with respect to assigned BOPEC ratings. Second, we compared the in-sample estimation results in tables 4A and 4B with those corresponding to the appropriate specification of equation (8)

using the Griliches adjustment. For example, for the case of the equity BOM model, we compare the in-sample estimation results based on the sample of 1,291 inspections of BHCs with publicly traded equity to the results where we use the full sample of inspections and plug in the mean abnormal return values for BHCs with no stock return data. Table 6 shows that the estimates for the securities market variables are very similar. The largest difference is observed for the sample based on BHCs with both public equity and debt, as expected, since it has the smallest sample size. However, even here, the coefficients' signs and significance are similar across the estimation methods. Given that the Griliches adjustment does not alter our results markedly and permits us to conduct our forecasting exercise for all four ratings samples, our forecasting results presented below are based on estimating equation (8) across our rolling data subsamples using the Griliches adjustment.¹⁹

4.1 *Forecasting Results*

Our measure of whether a model forecasts well is to ask how often its directional forecasts are correct. For example, if the model generates a signal suggesting a BOPEC upgrade, what percentage of the time does an upgrade actually take place? For our purposes, we generate a forecasted BOPEC rating $\widehat{BP}_{it}$ as follows:

$$\widehat{BP}_{it} = \sum_{j=1}^5 j * \widehat{\Pr}(BP_{it} = j), \quad (9)$$

where $\widehat{\Pr}(BP_{it} = j)$ is the ordered logit model's forecasted probability based on information at time $t - 1$ that BHC i will be assigned a rating of j at time t . We then compare these forecasted ratings to BP_{it-1} , the ratings as of time $t - 1$. An upgrade (or downgrade) signal is generated when the forecasted rating is a full rating better (or worse) than the existing rating. That is, an upgrade signal is given when $BP_{it-1} - \widehat{BP}_{it} \geq 1$, and a downgrade signal is

¹⁹In addition, we performed a third robustness check by conducting the out-of-sample exercise for the equity BOM model based on both adjusted and unadjusted data. Comparisons of these two sets of forecasts are both qualitatively and quantitatively similar. The results from these robustness checks are available from the authors upon request.

Table 6. Coefficient Estimates on the Securities Market Variables for the Subsample and Griliches Adjustment Estimation Techniques

	Equity Only		Debt Only		Equity and Debt	
	Subsample Estimation	Griliches Adjustment	Subsample Estimation	Griliches Adjustment	Subsample Estimation	Griliches Adjustment
SCSR	-0.831*	-0.828*	-	-	-0.972*	-0.804*
SCAR	-0.523*	-0.529*	-	-	-0.552*	-0.522*
Adj. Yield Spread	-	-	+2.686*	+2.387*	+2.382	+2.234*
Sample Size	1,291	3,010	309	3,010	283	3,010

Note: A * represents significance at the 5 percent level.

given when $BP_{it-1} - \widehat{BP}_{it} \leq -1$. A “no change” signal occurs when $|BP_{it-1} - \widehat{BP}_{it}| < 1$. Note that we only generate one-quarter-ahead forecasts, even though we evaluate them at various horizons.

In tables 7A–D, we report the percentages of signals that are correct. As shown in table 7A for the core model, an upgrade signal received four quarters prior to a BOPEC assignment results in an actual upgrade 55 percent of the time; 36 percent of the time, the actual outcome of the inspection is no change in rating; and 9 percent of the time, the actual outcome is a downgrade. By one quarter prior to the inspection, the upgrade signal is accurate 90 percent of the time. The model appears to be just as effective at picking up downgrades. Four quarters prior to the inspection, a downgrade signal is accurate 68 percent of the time, improving to 91 percent accuracy one quarter prior to the inspection.

To evaluate the BOPEC forecasts and their directional signals, we compare the conditional probabilities in the tables to a benchmark. For the core BOM, a natural benchmark for comparison purposes is the unconditional distribution of BOPEC ratings changes. In our sample, the unconditional probabilities of upgrades, downgrades, and no changes are 22 percent, 12 percent, and 66 percent, respectively. That is, given no information about a BHC, the probability of an upgrade is 22 percent. Given a signal from the core BOM, the upgrade probability rises to 55 percent at the four-quarter horizon and to 90 percent at the one-quarter horizon; see table 7A. Thus, conditioning on supervisory information in the core BOM model is clearly useful relative to the unconditional probabilities. This notion is formalized by the Pearson goodness-of-fit results in the right-most column, which tests whether the model’s conditional change probabilities are statistically different from the unconditional probabilities. We easily reject the null hypothesis that the conditioning variables do not provide useful information for forecasting BOPEC ratings.

As shown in tables 7B through 7D, forecast performance does not change much when securities market data are added into the BOM model; in fact, performance is a bit worse in some cases. With respect to the equity BOM model, the results in table 7B show that upgrade-signal accuracy increases from 64 percent four quarters prior to inspection to 91 percent accuracy within one quarter of the inspection, compared to 55 percent and 90 percent, respectively,

Table 7A. Out-of-Sample Forecast Performance of the Core BOM Model

	# Signals	Actual Inspection Outcome			Pearson Statistic
		Upgrade %	No Change %	Downgrade %	
Signal at −4 Quarters					
Upgrade	22	55%	36%	9%	145.0*
No change	2,825	22%	67%	11%	
Downgrade	31	3%	29%	68%	
Signal at −3 Quarters					
Upgrade	28	68%	21%	11%	254.3*
No change	2,820	22%	67%	11%	
Downgrade	30	0%	17%	83%	
Signal at −2 Quarters					
Upgrade	45	80%	13%	7%	460.1*
No change	2,793	22%	68%	10%	
Downgrade	40	0%	10%	90%	
Signal at −1 Quarter					
Upgrade	60	90%	8%	2%	624.6*
No Change	2,773	21%	69%	10%	
Downgrade	45	0%	9%	91%	
Note: This table presents the forecast performance results based on the 2,878 BOPEC change signals from the core BOM model at different horizons. A forecast signal is the difference between the forecasted BOPEC rating and the previously assigned BOPEC rating. Signals greater than one (less than one) are forecasts of upgrades (downgrades), respectively. The figures in bold indicate the outcome expected, conditional on the signal. Percentages in rows may not sum to 100 percent due to rounding. The Pearson goodness-of-fit statistic tests the null hypothesis that the distribution of BOPEC change outcomes conditional on the core model forecasts is not different from the unconditional probabilities of BOPEC upgrades (22 percent), no changes (66 percent), and downgrades (12 percent). The statistic is distributed $\chi^2(10)$. A * denotes significance at the 5 percent level.					

Table 7C. Out-of-Sample Forecast Performance of the Debt BOM Model

[illegible]

Table 7D. Out-of-Sample Forecast Performance of the Extended BOM Model

		Actual Inspection Outcome				
	# Obs.	Upgrade %	No Change	Downdgrade %	Pearson Stat. I	Pearson Stat. II
Signal at −4 Quarters						
Upgrade	28	57%	36%	7%	126.7*	6.4*
No Change	2,813	22%	67%	11%		
Downdgrade	37	0%	43%	57%		
Signal at −3 Quarters						
Upgrade	31	68%	267%	6%	270.5*	3.3
No Change	2,808	22%	67%	11%		
Downdgrade	39	0%	23%	77%		
Signal at −2 Quarters						
Upgrade	48	83%	13%	4%	633.1*	3.1
No Change	2,767	22%	68%	10%		
Downdgrade	63	0%	13%	87%		
Signal at −1 Quarter						
Upgrade	67	90%	7%	3%	715.5*	11.9*
No Change	2,746	21%	69%	10%		
Downdgrade	65	2%	17%	83%		

Note: This table presents the forecast performance results based on the 2,878 BOPEC change signals from the extended BOM model at different horizons. A forecast signal is the difference between the forecasted BOPEC rating and the previously assigned BOPEC rating. Signals greater than one (less than one) are forecasts of upgrades (downdgrades), respectively. The figures in bold indicate the outcome expected, conditional on the signal. Percentages in rows may not sum to 100 percent due to rounding. The Pearson goodness-of-fit statistic I tests the null hypothesis that the distribution of BOPEC change outcomes conditional on the extended BOM model forecasts is not different from the in-sample probabilities of BOPEC upgrades (22 percent), no changes (66 percent), and downdgrades (12 percent). The statistic is distributed $\chi^2(10)$. The Pearson goodness-of-fit statistic II tests the null hypothesis that the distribution of BOPEC change outcomes conditional on the extended BOM model forecasts is not different from the distribution of outcomes forecasted by the core BOM model. The statistic is distributed $\chi^2(2)$. A * denotes significance at the 5 percent level.

for the core model. For downgrades, however, forecast performance improves from 55 percent to 83 percent as the inspection approaches, which is not quite as large an improvement as for the core model. Overall, the Pearson test results in the last column suggest that the addition of the equity market variables does not improve forecast performance relative to the core model. In fact, at four quarters and one quarter prior, these variables reduce forecast performance.

In table 7C, the debt BOM model's forecasting performance is actually a bit worse than that of the core model with only supervisory variables. For example, an upgrade signal four quarters prior to the inspection is correct 52 percent of the time, while a downgrade signal at four quarters prior results in an actual downgrade 60 percent of the time. When compared to the core model's forecast signals using the Pearson test, we cannot reject the null hypothesis that the debt market variables do not improve forecast performance. Finally, as shown in table 7D, the forecast performance of the extended model with both equity and debt market variables is again marginally worse than that of the core model. In summary, the introduction of securities market variables to the core BOM model fails to improve overall BOPEC forecast performance for rating changes at conventional levels of significance.

4.2 Information in the Forecasts

Although BOPEC change forecasts do not appear to be appreciably different across the core and the extended BOM models presented in tables 7A–D, the set of BOPEC rating changes correctly signaled by these models are not identical. This outcome suggests the need to think carefully about the cost of forecast errors to supervisors. The forecasting literature has shown that combining forecasts from different models can improve certain aspects of forecast performance; see Granger and Newbold (1986) as well as Diebold and Lopez (1996) for further discussion. Hence, another way to gauge the contribution of securities market information is to examine the additional forecast signals for BHCs with public securities as generated by the extended models relative to the core model's signals. Seen in this light, the marginal benefit of monitoring securities market information is notable.

Table 8A. Improvements in BOPEC Downgrade Signals

	Equity BOM Model	Debt BOM Model	Extended BOM Model
4 Quarters Prior	24%	14%	22%
3 Quarters Prior	28%	10%	31%
2 Quarters Prior	28%	9%	29%
1 Quarter Prior	24%	5%	25%
Note: This table presents the percentage improvement in correct BOPEC downgrade signals when combining the downgrade signals from the core BOM model with those from the models incorporating securities market data. Downgrade signal is defined as forecasted rating – current rating > 1. The table reports the number of downgrades correctly signaled by the alternative models and not identified by the core model, expressed as a percentage of downgrades correctly identified by the core model. Sample contains 2,878 inspections, with 331 downgrades.			

Comparing BOPEC change signals generated by the core and the three extended BOM models, we ask what is the percentage increase in the number of correct downgrade signals when securities market data are incorporated into the core BOM model?²⁰ We focus on downgrades because these are the events of most interest to supervisors. As reported in table 8A, the extended model with adjusted bond yields and stock return data produces an additional 22 percent more correct signals at the four-quarter horizon over and above those produced by the core model. At the one-quarter horizon, the improvement is 25 percent more correct signals. Another interesting point is the similarity between the marginal contributions of the equity BOM model and the extended BOM model with both equity and debt market variables. Evidently, in this particular framework, most of the additional downgrade signals a supervisor can extract from securities market data come from the equity market. This result contrasts with the in-sample results and may be due to the relatively small number of BHCs with publicly traded debt in any given subsample period.

Of course, these three BOM models also produce incorrect signals over and above those produced by the core model. Since table 8A

²⁰Note that all of these additional correct forecasts are for ratings changes at BHCs with publicly traded securities.

Table 8B. Trade-Off between Correct and Incorrect Downgrade Signals

	Equity BOM Model	Debt BOM Model	Extended BOM Model
4 Quarters Prior	1/2	7/11	11/25
3 Quarters Prior	17/12	6/7	19/17
2 Quarters Prior	23/17	7/4	6/5
1 Quarter Prior	3/2	5/1	5/3
Note: This table presents the trade-off between additional correct and additional incorrect BOPEC downgrade signals provided by the alternative BOM models relative to the core BOM model. A cell entry of x/y suggests that the alternative BOM model identifies x additional correct downgrades signals beyond those of the core model, at the rate of y additional incorrect downgrade signals. Sample contains 2,878 inspections, with 331 downgrades.			

shows that these models identify additional BOPEC downgrades, their mistakes may be responsible for our earlier result that their overall forecast performance was almost the same as the core model. We examine this trade-off more closely in table 8B, where we express the models' ratios of correct downgrade signals to incorrect signals, which are also known as false positives or type 2 errors. For the extended BOM model at the four-quarter horizon, the model produces eleven additional correct signals at the cost of twenty-five incorrect downgrade signals. By the one-quarter horizon, however, the performance improves dramatically to five additional correct signals at the cost of only three additional incorrect signals. The models extended by equity variables and debt variables alone behave similarly.²¹ Interestingly, this signal trade-off for the debt BOM model is quite good; by one quarter out, it produces five correct signals for every incorrect signal. However, as indicated in table 8A, the

²¹Gropp, Vesala, and Vulpes (2004, 2006) conduct a similar analysis, but they focus on in-sample model fit. In their analysis, incorporating securities market variables into their model reduces false positive errors (i.e., type 2 errors). Our results align well with this result in that the combination of BOPEC downgrade signals from the core BOM model with the three expanded models adds more correct than incorrect signals, especially nearer to the BOPEC assignment; i.e., the combined set of signals has a reduced type 2 error rate than the set based just on the core model.

drawback to this model is that it produces relatively fewer signals beyond those from the core model.

In summary, our analysis of combined BOPEC downgrade signals indicates that models using securities market data can generate a reasonably large number of additional correct downgrade signals relative to the core BOM model. Given the emphasis placed by supervisors on BOPEC downgrades, we believe that these forecasting results argue for closer monitoring of securities market information for off-site supervisory monitoring.

5. Conclusion

Our empirical results indicate that both equity and debt market information are useful in improving the in-sample fit of our proposed BOM model for BOPEC ratings. Both types of securities market information also appear to be useful in explaining BOPEC upgrades and downgrades. Moreover, for our data set, we detect nonlinearities in the impact of securities market variables on BOPEC ratings. For BHCs closer to their estimated default points, the effect of our adjusted BHC bond yield spreads on BOPEC ratings is larger in magnitude than it is for BHCs further from their default points, and vice versa for equity market data. These results provide further empirical support for the finding by Berger, Davies, and Flannery (2000) that supervisory and debt-related assessments of BHCs are correlated and a refinement of their finding that supervisory and equity market assessments are not related; i.e., abnormal BHC stock returns provide useful information regarding supervisory assessments for BHCs that are far from their default points.

When we turn to out-of-sample forecasting, however, evidence for the usefulness of market information is disappointingly weak. For our analysis, we estimate our four BOM model specifications on a rolling subsample of data and then forecast BOPEC ratings out of sample. We find the forecast performance of the three BOM models that include securities market data is not much different than the performance of the core model based on supervisory data alone.

While the overall forecasting performance of the four BOM models is similar, the sets of forecasted rating changes are not identical. That is, the core model correctly identifies one set of BOPEC rating

changes, particularly downgrades, while the extended models correctly identify other sets of downgrades. The extended BOM models correctly identify additional BOPEC downgrades for publicly traded BHCs over and above the correct forecasts from the core model. These additional correct forecasts can be achieved at a relatively modest cost of additional incorrect signals. Hence, supervisory use of securities market information within the context of an off-site monitoring model, such as our proposed BOM model, appears to be reasonable.

References

- Basel Committee on Banking Supervision. 2003. "Markets for Bank Subordinated Debt and Equity in Basel Committee Member Countries." Working Paper No. 12, Bank for International Settlements.
- Berger, A. N., S. M. Davies, and M. J. Flannery. 2000. "Comparing Market and Supervisory Assessments of Bank Performance: Who Knows What When?" *Journal of Money, Credit, and Banking* 32 (3): 641–67.
- Bliss, R., and M. Flannery. 2001. "Market Discipline in the Governance of U.S. Bank Holding Companies: Monitoring versus Influencing." In *Prudential Supervision: What Works and What Doesn't*, ed. F. S. Mishkin, 107–43. Chicago: University of Chicago Press.
- Board of Governors of the Federal Reserve System. 1999. "Using Subordinated Debt as an Instrument of Market Discipline." Report of the Study Group on Subordinated Notes and Debentures. Federal Reserve System Staff Study No. 172.
- . 2004. "Bank Holding Company Rating System." Supervision and Regulation Letter No. 04-18, Division of Banking Supervision and Regulation.
- Campbell, J. Y., A. W. Lo, and A. C. MacKinlay. 1997. *The Econometrics of Financial Markets*. Princeton, NJ: Princeton University Press.
- Cole, R. A., and J. W. Gunther. 1998. "Predicting Bank Failures: A Comparison of On- and Off-Site Monitoring Systems." *Journal of Financial Services Research* 13 (2): 103–17.

- Covitz, D. M., D. Hancock, and M. L. Kwast. 2004a. "A Reconsideration of the Risk Sensitivity of U.S. Banking Organization Subordinated Debt Spreads: A Sample Selection Approach." *Economic Policy Review* (Federal Reserve Bank of New York) 10 (2): 73–92.
- . 2004b. "Market Discipline in Banking Reconsidered: The Roles Funding Manager Decisions and Deposit Insurance Reform." Finance and Economics Discussion Series No. 2004-53, Board of Governors of the Federal Reserve System.
- DeFerrari, L. M., and D. E. Palmer. 2001. "Supervision of Large Complex Banking Organizations." *Federal Reserve Bulletin* (February): 47–57.
- DeYoung, R., M. J. Flannery, W. W. Lang, and S. M. Sorescu. 2001. "The Information Content of Bank Exam Ratings and Subordinated Debt Prices." *Journal of Money, Credit, and Banking* 33 (4): 900–25.
- Diebold, F. X., and J. A. Lopez. 1996. "Forecast Evaluation and Combination." In *Handbook of Statistics 14: Statistical Methods in Finance*, ed. G. S. Maddala and C. R. Rao, 241–68. Amsterdam: North-Holland.
- Estrella, A., S. Park, and S. Peristiani. 2000. "Capital Ratios as Predictors of Bank Failure." *Economic Policy Review* (Federal Reserve Bank of New York) 6 (2): 33–52.
- Evanoff, D., and L. Wall, L. 2001. "Sub-debt Yield Spreads as Bank Risk Measures." *Journal of Financial Services Research* 20 (2): 121–46.
- . 2002. "Measures of the Riskiness of Banking Organizations: Subordinated Debt Yields, Risk-Based Capital and Examination Ratings." *Journal of Banking and Finance* 26 (5): 989–1009.
- Flannery, M. J. 1998. "Using Market Information in Prudential Bank Supervision: A Review of the U.S. Empirical Evidence." *Journal of Money, Credit, and Banking* 30 (3): 273–305.
- Flannery, M. J., and S. U. Sorescu. 1996. "Evidence of Bank Market Discipline in Subordinated Debenture Yields: 1983–1991." *Journal of Finance* 51 (4): 1347–77.
- Goyal, V. K. 2005. "Market Discipline of Bank Risk: Evidence from Subordinated Debt Contracts." *Journal of Financial Intermediation* 14 (3): 318–50.
- Granger, C. W. J., and Newbold, P. 1986. *Forecasting Economic Time Series*. 2nd ed. San Diego: Academic Press.

- Griliches, Z. 1986. "Economic Data Issues." In *Handbook of Econometrics*, Vol. III, ed. Z. Griliches and M. D. Intriligator. Amsterdam: Elsevier Science Publishers BV.
- Gropp, R., and A. J. Richards. 2001. "Rating Agency Actions and the Pricing of Debt and Equity of European Banks: What Can We Infer About Private Sector Monitoring of Bank Soundness?" *Economic Notes* 30 (3): 373–98.
- Gropp, R., J. Vesala, and G. Vulpes. 2004. "Market Indicators, Bank Fragility, and Indirect Market Discipline." *Economic Review* (Federal Reserve Bank of New York) 10 (2): 53–62.
- . 2006. "Equity and Bond Market Signals as Leading Indicators of Bank Fragility." *Journal of Money, Credit, and Banking*, 38 (2): 399–428.
- Gunther, J. W., and R. R. Moore. 2003. "Early Warning Models in Real Time." *Journal of Banking and Finance* 27 (10): 1979–2002.
- Hancock, D., and M. Kwast. 2001. "Using Subordinated Debt to Monitor Bank Holding Companies: Is It Feasible?" *Journal of Financial Services Research* 20 (2): 147–87.
- Hanweck, G. A., and L. J. Spellman. 2002. "Forbearance Expectations and the Subordinated Debt Signal of Bank Insolvency." Manuscript, School of Management, George Mason University.
- Hirtle, B. J., and J. A. Lopez. 1999. "Supervisory Information and the Frequency of Bank Examinations." *Economic Policy Review* (Federal Reserve Bank of New York) 5 (1): 1–20.
- Krainer, J., and J. A. Lopez. 2003. "How Might Financial Market Information Be Used for Supervisory Purposes?" *Economic Review* (Federal Reserve Bank of San Francisco), 29–45.
- . 2004. "Incorporating Equity Market Information into Supervisory Monitoring Models." *Journal of Money, Credit, and Banking* 36 (6): 1043–67.
- Meyer, P. A., and H. W. Pifer. 1970. "Prediction of Bank Failures." *Journal of Finance* 25 (4): 853–68.
- Pettway, R. H., and J. F. Sinkey. 1980. "Establishing On-Site Bank Examination Priorities: An Early-Warning System Using Accounting and Market Information." *Journal of Finance* 35 (1): 137–50.
- Ronn, E. I., and A. K. Verma. 1986. "Pricing Risk-Adjusted Deposit Insurance: An Option-Based Model." *Journal of Finance* 41 (4): 871–95.

- Sahajwala, R., and van den Bergh, P. 2000. "Supervisory Risk Assessment and Early Warning Systems." Basel Committee on Banking Supervision Working Paper No. 4, Bank for International Settlements.
- Sinkey, J. F. 1975. "A Multivariate Statistical Analysis of the Characteristics of Problem Banks." *Journal of Finance* 30 (1): 21–36.

Capital Regulation and Banks' Financial Decisions*

Haibin Zhu

Monetary and Economic Department, Bank for International Settlements

This paper develops a stochastic dynamic model to examine the impact of capital regulation on banks' financial decisions. In equilibrium, lending decisions, capital buffer, and the probability of bank failure are endogenously determined. Compared to a flat-rate capital rule, a risk-sensitive capital standard causes the capital requirement to be much higher for small (and riskier) banks and much lower for large (and less risky) banks. Nevertheless, changes in actual capital holdings are less pronounced due to the offsetting effect of capital buffers. Moreover, the nonbinding capital constraint in equilibrium implies that banks adopt an active portfolio strategy, and hence the countercyclical movement of risk-based capital requirements does not necessarily lead to a reinforcement of the credit cycle. In fact, the results from the calibrated model show that the impact on cyclical lending behavior differs substantially across banks. Lastly, the analysis suggests that the adoption of a more risk-sensitive capital regime can be welfare improving from a regulator's perspective, in that it causes less distortion in loan decisions and achieves a better balance between safety and efficiency.

JEL Codes: G21, G28.

*I would like to thank Gianni de Nicolo, Diemo Dietrich, Arturo Estrella, Vincenzo Quadrini, Hyun Shin (the editor), Kostas Tsatsaronis, Skander Van den Heuvel, and participants at a BIS workshop, the EFMA 2007 Annual Conference, the Riksbank conference on "The Evolution of Financial Markets and Financial Institutions: New Threats to Financial Stability," and the "Banking and the Macroeconomy" workshop at ETH Zurich for insightful comments. The views expressed in this paper are my own and do not necessarily represent those of the Bank for International Settlements. Author contact: Monetary and Economic Department, Bank for International Settlements, Basel, Switzerland. Tel: 41 61 280 9164. Fax: 41 61 280 9100. E-mail: haibin.zhu@bis.org.

1. Introduction

Banking is one of the most regulated industries in the world. Among various regulatory measures, the regulation of bank capital is crucial due to the important role it plays in banks' soundness and risk-taking behavior, and its influence on the competitiveness of banks.¹ In practice, a key aspect of capital regulation is the calculation of minimum regulatory capital, which is typically based on the credit risk of bank assets.

A unified regulatory capital framework, which applied to all internationally active banks, was first introduced in 1988 by the Basel Committee on Banking Supervision (BCBS). It was commonly known as Basel I (BCBS 1988). However, the risk matrix was very simple in the sense that the risk weights were practically the same for bank loans of different risk.² The low sensitivity arguably could have led to severe market distortions as banks swapped low-risk assets against high-risk ones with more favorable risk weighting (regulatory arbitrage). Reflecting this, in June 2004, the Basel Committee on Banking Supervision released a revision of the existing accord on capital regulation, known as Basel II (BCBS 2006). One of the most remarkable changes under the new framework was the introduction of a more risk-sensitive capital standard.

The transition toward a more risk-sensitive capital regime has generated the discussion on its implications for the banking industry. Two major concerns arise. First, it tends to cause substantial changes in the level of bank capital. The study of Ervin and Wilde (2001), which is based on a hypothetical bank portfolio in the United States in 1990, suggests that the minimum capital requirement could drop from 8 percent to 6.8 percent with the introduction of the risk-based capital rule. Similarly, a study conducted by U.S. federal regulators (OCC, FRB, FDIC, and OTS 2004) shows that the new capital rule will cause the minimum required risk-based capital (MRC), the sum of expected and unexpected losses,

¹See Berger, Herring, and Szegö (1995) and Santos (2001) for surveys on the motivations behind capital regulation.

²Generally, G-10 government debt is weighted 0 percent, G-10 bank debt is weighted 20 percent, and other debts, including corporate debt and the debt of non-G-10 governments, are weighted 100 percent. As an exception, mortgage on residential property receives a risk weight of 50 percent.

to drop on average by 12.5 percent. Across entities, the changes in MRC vary substantially within a range of -50 percent to 70 percent, with a median reduction of 24 percent. These results raise concerns that the introduction of the new capital standard may cause abrupt and undesirable changes to the banking industry. Second, the greater risk sensitivity in the new capital regime may cause additional volatility in economic activity, sometimes cited as the procyclicality problem.³ In particular, capital requirements can increase as the economy falls into recession and fall as the economy enters an expansion. The increase in capital requirements during the downswing may result in a credit crunch and thus worsen already adverse economic conditions. The empirical relevance of this concern was later confirmed in several countries, including the United States (Catarineu-Rabell, Jackson, and Tsomocos 2005; Gordy and Howells 2006; Jordan, Pek, and Rosengren 2003; Kashyap and Stein 2004), Spain (Corcostegui et al. 2002), and Mexico (Segoviano and Lowe 2002).

However, there are several major caveats in these studies, which suggest that the above conclusions are debatable. First, these studies focus exclusively on the change in regulatory capital. Nevertheless, in order to understand the impact of regulatory measures, it is more important to examine the change in banks' actual capital holdings.⁴ A well-known fact is that most banks tend to hold a significant amount of capital above the regulatory requirement in practice, either for efficiency reasons or because the capital cushion is established as a precaution against contingencies (adverse events or regulatory penalties; see Barrios and Blanco 2003; Elizalde and Repullo 2006). The existence of capital buffers suggests that banks' actual capital holdings may not change as remarkably as regulatory capital. Equally important, if banks hold enough capital buffer during economic downturns, the procyclicality

³See Borio, Furfine, and Lowe (2001) and Jokivuolle and Kauko (2001) for a related discussion and Allen and Saunders (2004) for a literature survey.

⁴This paper uses three distinct definitions of bank capital. *Economic capital* refers to banks' capital holdings in the absence of capital regulation, *regulatory capital* refers to the minimum capital requirements based on specific capital rules, and *actual capital* (or total capital) refers to banks' equilibrium capital holdings in an economy with capital regulation. Accordingly, *capital buffer* is defined as the difference between actual and regulatory capital.

concern tends to be mitigated. Given its importance, the cyclical behavior of bank capital buffer has been investigated by many researchers, but so far the evidence has been mixed. Estrella (2004b), Heid (2005), Koopman, Lucas, and Klaassen (2005), and Peura and Jokivuolle (2004) suggest that the existence of capital buffers can potentially mitigate the volatility in total capital. By contrast, the empirical evidence in Germany (Stolz and Wedow 2005) and Spain (Ayuso, Pérez, and Saurina 2004) shows that capital buffers are also anticyclical; therefore, it is not clear whether actual capital holdings are more volatile than regulatory capital or vice versa.⁵

Second, these studies adopt a passive portfolio assumption; e.g., banks hold the same portfolio under the two capital regimes or hold portfolios as specified by the minimum capital requirements. This is questionable, given that an important objective of the revision of capital rules is to avoid excessive risk taking that may arise under the old (less risk-sensitive) regime. Hence, it is more interesting to investigate how banks adjust their risk-taking behavior in their loan-extension process. In other words, an active portfolio strategy is more appropriate for the purpose of this study.

Moreover, some of these impact studies are based on the industry information at a particular point at time. Therefore, the results fail to recognize the dynamic aspect of regulatory implications. The dynamic perspective is crucial in the sense that the underlying credit risk of bank assets can change substantially over the business cycle. As a result, the impacts on banks' capital holdings (both regulatory and actual capital) and portfolio composition tend to be different at different phases of the business cycle.

This paper seeks to fill these gaps and provide new insights on these important issues. I propose a stochastic dynamic model in which a bank's objective is to maximize shareholder value by making loans. A larger loan book yields lower return but is associated with lower risk due to benefits from portfolio diversification. Each

⁵Cross-country studies, such as Bikker and Metzmakers (2004) and Kim and Lee (2006), find mixed results on the impact of capital regulation on the dynamics of actual capital holdings. A recent study by Repullo and Suarez (2007) suggests that capital buffer tends to be countercyclical under Basel I but is procyclical under Basel II.

bank has to choose its optimal capital structure, facing the trade-off between one-period investment gains and future franchise value. In equilibrium, the bank's leverage ratio, lending strategy, and failure probability are jointly determined. I consider three economic environments: (i) the baseline economy in which capital regulation is absent, (ii) a flat-rate capital regime in which a uniform capital-asset ratio requirement is imposed, and (iii) a risk-based capital regime in which the minimum capital requirement is determined by the credit risk of bank assets. The model is calibrated to match aspects of the U.S. economy in 1980–2002, and the equilibrium outcomes under the three economic environments are compared with each other.

The numerical results show that banks' financial decisions depend crucially on the state of the business cycle, banks' initial equity holdings, and the form of capital standards. Comparing equilibrium outcomes under different capital regimes sheds new light on the issues of interest. On the one hand, the results support the view that the adoption of a risk-sensitive capital rule causes significant changes in minimum capital requirements, and the impact is substantially different across banks, depending on the risk profile of their loan portfolios. However, they also suggest that changes in actual capital holdings can be much smaller than changes in minimum capital requirements due to the offsetting effect of capital buffers. On the other hand, the results find little support for the idea that capital regulation amplifies procyclicality in bank lending. The volatility of bank credit (and economic output) over the business cycle is smaller under both capital regimes than the volatility under the baseline economy (without capital regulation). Even when comparing the two capital regimes, higher volatility of capital requirements under the risk-based capital rule does not necessarily lead to amplified swings in banks' lending activity, because the non-binding capital constraint implies that the loan portfolios chosen by banks in equilibrium are often different from those implied by regulatory capital constraints. In fact, for smaller banks, the swing in bank lending is less pronounced under the risk-based capital rule than under the flat-rate capital regime.

The model also provides a framework to compare the optimality of the two capital regimes from a regulator's perspective, assuming that the regulator has to choose between maximizing franchise value and minimizing the negative externality caused by bank failures.

Illustrative examples suggest that the risk-based capital rule can outperform the flat-rate capital requirement, in that expected franchise value is higher for the same level of bank safety. The results hold both at the micro level (comparison based on individual banks) and at the macro level (comparison based on the banking system as a whole). The intuition behind this observation is that under the flat-rate capital requirement, the failure to reward sound investments causes a big distortion in limiting the scale of lending for those banks that have access to high-return, low-risk investment opportunities. From a dynamic perspective, such distortion implies that banks have to make a large sacrifice of future franchise value in order to reduce the default risk in the current period. By contrast, the future franchise value at risk is much smaller under the risk-based capital regulation because banks can invest in high-return, low-risk projects without incurring the high cost of holding capital.

The remainder of the paper is organized as follows. Section 2 outlines the model setup and section 3 discusses the qualitative results. In section 4, the model is then calibrated to the U.S. economy, allowing us to examine the equilibrium financial decisions under different capital regimes. Section 5 compares the two capital regimes from a regulator's perspective. The final section concludes.

2. The Model

The benchmark model follows a modified version of the model developed by Cooley and Quadrini (2001) for a generic firm. I treat a bank as a special type of firm, whose objective is to maximize the expected discounted value of dividends, $E_0\{\sum_{t=0}^{\infty} \beta^t d_t\}$, where d_t is the dividend payment at time t , β is the discount factor of bank shareholders, and E is the expectations operator. Accordingly, the assumptions on the asset and liability of the *firm's* balance sheet will be modified to reflect distinctive features of the banking sector. The modified framework bears considerable similarity to the one developed by Van den Heuvel (2002).

2.1 Loans

In each period, the bank tries to maximize shareholder value by making loans. To maintain analytical tractability, this model abstracts

from the determination of lending rates by working directly with a return function of $Y_t = (z_t + \epsilon_t + \mu_t)F(L_t)$, where Y_t is the revenue from making loans (adjusted by cost of defaults), L_t is the size of the loan portfolio, and $F(\cdot)$ satisfies $F(0) = 0$, $F' > 0$, and $F'' < 0$.⁶ The concavity of the $F(\cdot)$ function implies decreasing returns to scale for bank investments, which is standard in the literature and has been supported by empirical findings in Berger et al. (2005), Carter and McNulty (2005), and Cole, Goldbert, and White (2004), among others.

The other three variables— z_t , ϵ_t , and μ_t —jointly determine the profitability of bank lending. In particular, z_t represents the state of the economy that determines the level of investment yields and follows a first-order Markov process (assumption 1), ϵ_t is a bank-specific shock that determines the risk of bank assets (assumption 2), and μ_t measures the benefit from monitoring as it is explained below (assumption 3).

ASSUMPTION 1. The state of the economy, z , takes values in a finite set $Z = \{z_1, z_2, \dots, z_N\}$ and follows a first-order Markov process with transition probabilities $P(z_{t+1}|z_t)$.

ASSUMPTION 2. The bank-specific shock ϵ is independently and identically distributed and has a normal density function $N(0, \sigma_\epsilon^2(L))$. The volatility is inversely related to the size of bank assets, i.e., $\frac{d\sigma_\epsilon}{dL} < 0$.

At the beginning of each period, each bank observes z_t but not ϵ_t . In other words, each bank knows the level of expected investment yields in the current period and is aware that the underlying risk depends on its lending strategy. The above two assumptions suggest that when a bank chooses the size of its loan portfolio, it faces the trade-off between return and risk. A larger loan portfolio produces a lower average return but is also less risky, the latter of which can

⁶This partial-equilibrium assumption enables me to isolate the potential cyclicity coming from the supply side of the bank loan market. As a trade-off, the financing strategies of borrowers and the impact of banks' lending decisions on the real economy, which would be important in a general equilibrium setting, are ignored in this paper.

be justified on the ground of greater scope for diversification as proposed in finance theory. For instance, a report by the Group of Ten (2001) suggests that consolidation of banks within the United States has led to reductions in risk due to geographic diversification.

Finally, there exists another bank-specific term, μ , which represents the contribution of bank-specific efforts in improving investment returns. Such efforts can vary across banks due to differences in managerial skills, market power, and monitoring incentives. By paying a monitoring cost of $\theta(\mu)$ per unit of loan, the bank is able to improve the average investment return by μ .⁷ This unit cost increases in μ at a growing rate, implying that the room for return enhancement tends to become smaller, i.e., a diminishing benefit from monitoring.

ASSUMPTION 3. *The function $\theta(\mu)$ satisfies $\theta(0) = 0$, $\theta'(\cdot) > 0$, and $\theta''(\cdot) > 0$.*

The model abstracts from the determination of lending rates and assumes instead a continuous distribution of net revenues.⁸ The continuity assumption can be justified as follows. Suppose that a bank can make many individual loans, with a fixed lending rate for each loan. Since each loan can go bad with positive probability, the effective average return of the whole portfolio tends to be distributed over a large set of real numbers, with the maximum being the highest possible return if no default occurs and the minimum the worst scenario if all loans default.

2.2 Deposits and Equity

A bank begins each period with equity capital equal to e_t . The bank finances its investment by equity capital and bank deposits, the latter being a debt instrument with a fixed return of r_t^d .

Bank deposits are not risk free, in the sense that the bank may default if its loan investment does not generate enough profits.

⁷In the literature, the monitoring effort could have two effects. One is to induce entrepreneurs to choose high-productivity projects; the other is to reduce the business risk. Since the revenue function refers to default-adjusted returns, both effects tend to increase the expected revenue.

⁸The same assumption is adopted by Acharya (2003).

Therefore, a risk premium is charged to compensate for the default risk. This paper abstracts from the supply of bank deposits from the household sector by assuming that the elasticity of the deposit supply is infinite and by making the following assumption.

ASSUMPTION 4. Depositors are risk neutral, and their reservation return is the risk-free rate, r_t^f .

The assumption of perfect access to nonreservable deposits is also adopted in Van den Heuvel (2002).⁹ The advantage of this assumption is that it allows me to focus on the bank capital channel through which the change in capital regimes affects banks' lending behavior. On the other hand, it implies that the bank lending channel does not exist in this model, since a necessary condition for the existence of an effective bank lending channel—that shocks to reservable deposits should affect the bank's loan supply (Bernanke and Blinder 1988; Kashyap and Stein 1994)—is absent from this model.

However, this paper departs from the assumption in Van den Heuvel (2002) that bank deposits are fully insured (with zero premium) and, by extension, the deposit rate equals the risk-free rate.¹⁰ Instead, the bank will need to pay a risk premium in line with its likelihood of default. This is important in that a blanket deposit-insurance scheme tends to cause the moral hazard problem, which introduces undesirable bias in banks' financial decisions. Assuming that depositors are risk neutral and the expected payoff on bank deposits is $E(r)$, the supply function is

$$D_t^s = \begin{cases} 0 & \text{if } E(r) < r_t^f \\ [0, \infty] & \text{if } E(r) = r_t^f \\ \infty & \text{if } E(r) > r_t^f. \end{cases} \quad (1)$$

At the end of each period, banks receive revenues on their asset portfolios and pay back deposits. A solvent bank chooses its dividend

⁹By assuming a perfect access to deposits, this paper also abstracts from the vulnerability issue related to the bank-run phenomenon, as examined in Diamond and Dybvig (1983).

¹⁰The assumption that a risk premium is charged for bank deposits is approximately equivalent to the alternative assumption that bank deposits are fully insured by a risk-based deposit-insurance scheme.

payment (d_t) or, equivalently, its initial equity holding (e_{t+1}) for the next period. Importantly, the dividend is allowed to be negative in this model; i.e., the bank is able to issue new equity if necessary. Nevertheless, issuing new equity is costly. This is consistent with the statement made by practitioners and can be justified with respect to transaction costs and differential taxation levels for equity. Specifically, the cost of raising new equity is assumed to be linear following Cooley and Quadrini (2001).¹¹

ASSUMPTION 5. *The bank pays an issuance cost of $\Lambda(-d_t) = -\lambda d_t$ to raise new equity in the amount of $-d_t$ ($d_t < 0$).*

Finally, this model makes the following assumption on the exit of a bank.

ASSUMPTION 6. *A bank defaults if its net asset value drops below zero. A bank cannot reenter the market once it exits; i.e., its franchise value is zero thereafter.*

Here the bank's net asset value is defined as the end-of-period total revenue minus the bank's debt liability, i.e., the principal and interest payments on bank deposits. If the bank's net asset value drops below zero, its equity capital is negative and it claims bankruptcy. In other words, the renegotiation of bank debt is not allowed in this paper and a default automatically leads to permanent exit.

2.3 Capital Regulation

Regulation is introduced in the form of a minimum capital requirement imposed by the regulatory authority. Each bank has to meet this requirement at all times. The minimum level of bank capital typically depends on the size of the bank's loan book and, under a risk-based capital regime, is related to the risk profile of bank assets.

¹¹The cost of raising new equity is assumed to be constant over time. In practice, it is often more expensive for banks to raise fresh capital in economic downturns. Numerical studies show that introducing time-varying cost of equity issuance does not change the results qualitatively.

ASSUMPTION 7. *The bank has to meet the minimum capital requirement, e_{min} , which depends on the characteristics of bank assets.*

Notice that the capital regulation, combined with the equity issuance cost assumption, causes deviations from the Modigliani-Miller framework. Without these assumptions, whether bank loans are financed with deposits or equity is irrelevant. As explained later, the probability of default and the subsequent losses increase banks' willingness to hold more capital than required.

Owing to the simplified maturity structure of bank assets and liabilities in this paper, the minimum capital requirement is effectively operative only at the moment that the bank decides on the size of its loan portfolio.¹² Throughout the paper, I assume that the minimum capital requirement cannot be violated or, equivalently, that the penalty cost is infinite and the bank's franchise value goes to zero if it fails to meet the minimum capital requirement.¹³

This paper examines two alternative forms of capital requirements. The first form imposes a flat capital/asset ratio requirement (κ) of 8 percent on all banks. That is, the level of regulatory capital depends on the size of bank assets but not their riskiness. By contrast, under the alternative form, the minimum regulatory capital is calculated on the basis of the risk profile of bank assets. In particular, it is determined in a way that the probability of default for the bank does not exceed a threshold level (0.1 percent). That is, the minimum level of regulatory capital (e_{min}) is determined by¹⁴

$$(z + \epsilon_{0.001} + \mu)F(L) - \theta(\mu)L = (L - e_{min})r^d,$$

¹²Given that a bank's portfolio is fixed during each period, meeting the minimum capital requirement at the beginning of a period (when the loan is extended) implies that the requirement is not violated throughout the period. Moreover, it is not meaningful to discuss capital charges after the realization of bank-specific shocks, particularly under the risk-based capital regime, because the uncertainty in asset returns is eliminated.

¹³In a model with richer maturity structure of bank assets and liabilities, the capital requirement might be violated subject to a non-infinitely harsh penalty.

¹⁴Here the calculation is based on the credit VaR (value-at-risk) for the bank portfolio as a whole rather than on the basis of individual loans. Gordy (2003) and Kupiec (2004) prove that when the default risk is driven by a single common factor and the bank portfolio is fully diversified, the two definitions yield the same level of regulatory capital.

where $\epsilon_{0.001}$ is the 0.1 percent critical value of the normally distributed variable ϵ with mean zero and standard deviation of σ_ϵ .¹⁵ The above equation implies that if a bank holds the required level of capital, there is 0.1 percent probability that the bank's revenue will be insufficient to finance the bank's deposit liability; i.e., the bank will become insolvent. Because the default risk is very small, the deposit rate should be very close to the risk-free rate. Therefore, the above equation can be approximated by

$$e_{min} \approx -\frac{(z + \mu)F(L) - (r^f + \theta(\mu))L}{r^f} - \frac{\epsilon_{0.001}F(L)}{r^f}. \quad (2)$$

In equation (2), the minimum capital requirement consists of two components. The first component refers to the expected loss and the second to the unexpected loss of the bank's portfolio. The expected loss is jointly determined by the state of the economy and the bank's choice of monitoring effort.

2.4 *Timing*

At the beginning of each period, there exists a continuum of banks each with equity equal to $e_{i,t}$, which is carried over from last period. The state of the economy, z_t , is known since the end of last period, while the bank-specific shock $\epsilon_{i,t}$ is realized at the end of the current period. That is, at the moment of making lending decisions, each bank knows z_t but does not know $\epsilon_{i,t}$. Since banks' decisions are independent of each other, in the remaining part the analysis will focus on the optimization problem of an individual bank, and the subscript i , which represents bank-specific variables, will be dropped.

The bank then decides on the volume of loans it will supply (L_t) and the amount of monitoring cost it will pay (θ_t). To finance the loan portfolio, the bank's liability comprises equity holding (e_t) and deposits ($L_t - e_t$), the latter paying an equilibrium deposit rate of

¹⁵Implicitly, I assume that banks have an incentive to disclose accurate information on their capital adequacy, monitoring effort, and loan risk. Estrella (2004a) suggests that it can be achieved via the combination of voluntary disclosure, direct supervision, and financial market discipline, which is beyond the scope of this paper.

r_t^d . The bank's financial structure has to meet the regulatory capital requirement, if one exists.

At the end of each period, the bank-specific shock ϵ_t is realized and depositors are paid back. The bank either continues its operation (if its net asset value is positive) or files for bankruptcy (if its net asset value drops below zero). If a bankruptcy occurs, all depositors receive an equal amount of payment by splitting residual assets on the bank's balance sheet.

Lastly, at the end of each period, the state of the economy in the next period, z_{t+1} , is revealed. Depending on the amount of net profits and the future state of the economy, a solvent bank chooses whether to distribute dividends to its shareholders or to issue new equity (d_t). This choice is reflected on the bank's initial equity holding for the next period, e_{t+1} .

3. The Solution to the Model

The bank's problem can be solved using dynamic programming techniques and the backward-induction method. At each period, a bank makes three important decisions: choosing the size of its loan portfolio (L_t) and the monitoring effort (θ_t) at the beginning of the period and deciding on the dividend policy (d_t) at the end of the period.

I first examine the solvency condition for each bank at the end of each period. For given initial equity (e_t), bank lending (L_t), deposit rate (r_t^d), and monitoring effort (θ_t), the end-of-period net asset value of the bank is

$$\pi(L_t, \mu_t, \epsilon_t, r_t^d) = (z_t + \epsilon_t + \mu_t)F(L_t) - \theta(\mu_t)L_t - r^d(L_t - e_t). \quad (3)$$

There are two possible outcomes. If the bank's net asset value is positive—i.e., $\pi_t > 0$ —depositors receive full payment. Otherwise, the bank files for bankruptcy and depositors have to accept a smaller payment prorated by the amount of residual assets on the bank's balance sheet, which equals $(z_t + \epsilon_t + \mu_t)F(L_t) - \theta(\mu_t)L_t$.

Accordingly, the critical value of the bank-specific shock, $\bar{\epsilon}$, is defined by

$$\pi(L_t, \mu_t, \bar{\epsilon}_t, r_t^d) = 0. \quad (4)$$

For $\epsilon \geq \bar{\epsilon}$, depositors receive full payment; otherwise, the bank defaults. Since depositors are risk neutral, in equilibrium their

expected payoff should equal the return on risk-free investments (see equation (1)). That is,

$$\begin{aligned} \int_{\bar{\epsilon}_t}^{\infty} r_t^d(L_t - e_t)\phi(\epsilon)d\epsilon + \int_{-\infty}^{\bar{\epsilon}_t} [(z_t + \epsilon_t + \mu_t)F(L_t) - \theta(\mu_t)L_t]\phi(\epsilon_t)d\epsilon \\ = r_t^f(L_t - e_t), \end{aligned} \quad (5)$$

where $\phi(\epsilon)$ is the probability density function of the shock variable ϵ . Combining equations (4) and (5), $\bar{\epsilon}_t$ and r_t^d can be jointly determined conditional on the choice of L_t and μ_t .

$$\bar{\epsilon}_t = \bar{\epsilon}(L_t, e_t, \mu_t) \quad (6)$$

$$r_t^d = r^d(L_t, e_t, \mu_t) \quad (7)$$

LEMMA 1. $\bar{\epsilon}$ increases with bank asset (L) and decreases with bank equity (e).

Proof. Based on equations (4) and (5), it can be easily derived that

$$\begin{aligned} \frac{\partial \bar{\epsilon}}{\partial L} &= \frac{1}{1 - \Phi(\bar{\epsilon})} \cdot \left\{ \frac{\theta(\mu) + r_f}{F(L)} \cdot \left(1 - \frac{F'(L)L}{F(L)} \right) + \frac{F'(L)e}{F(L)^2} \right\} > 0 \\ \frac{\partial \bar{\epsilon}}{\partial e} &= -\frac{r^f}{F(L) \cdot (1 - \Phi(\bar{\epsilon}))} < 0. \end{aligned}$$

Lemma 1 is quite intuitive. Notice that $\bar{\epsilon}$ is the threshold value above which a bank can generate positive net profits; hence, a higher $\bar{\epsilon}$ indicates lower profitability and a higher probability of bank failure. Everything else constant, an increase in bank assets tends to lower bank profits and increase bank fragility for two reasons. For one, it lowers the marginal return of bank assets because of the concavity of the revenue function. Moreover, it increases the leverage ratio, pushes upward the default boundary ($\bar{\epsilon}$), and, consequently, raises the probability of default ($\Phi(\bar{\epsilon})$). Similarly, lower bank equity increases the probability of bank failure because of higher leverage.

Once $\bar{\epsilon}_t$ is determined, the net asset value of the bank can be rewritten as

$$\pi(L_t, \mu_t) = \begin{cases} (\epsilon_t - \bar{\epsilon}_t)F(L_t) & \text{if } \epsilon \geq \bar{\epsilon}_t \\ 0 & \text{if } \epsilon_t < \bar{\epsilon}_t. \end{cases} \quad (8)$$

Define $V(e, z)$ as the value function for a bank with an initial equity of e and the initial state of the economy z . The bank's optimization problem can be written as follows (equations (9)–(14)):

$$V(e_t, z_t) = \max_{(L_t, \mu_t)} \sum_{z_{t+1}} \int_{\bar{\epsilon}_t}^{\infty} W(\pi(L_t, \mu_t, \epsilon_t, e_t), z_{t+1}) \cdot P(z_{t+1}|z_t) \phi(\epsilon) d\epsilon \quad (9)$$

subject to equations (6), (7), (8), and

$$L_t \geq e_t \quad (10)$$

$$e_t \geq e_{\min}(L) \quad (11)$$

$$W(\pi(L_t, \mu_t), z_{t+1}) = \max_{e_{t+1}} \{d(\pi_t, e_{t+1}) + \beta V(e_{t+1}, z_{t+1})\} \quad (12)$$

subject to

$$d(\pi(L_t, \mu_t), e_{t+1}) = \begin{cases} \pi(L_t, \mu_t) - e_{t+1} & \text{if } \pi_t \geq e_{t+1} > 0 \\ [\pi(L_t, \mu_t) - e_{t+1}](1 + \lambda) & \text{if } 0 \leq \pi_t < e_{t+1} \end{cases} \quad (13)$$

$$e_{t+1} = 0 \quad \text{if } \pi_t < 0. \quad (14)$$

By solving the above dynamic problem, the bank maximizes the expected value for its shareholders. The uncertainty in future franchise value comes from both the uncertainty about the future state of the economy and the riskiness of bank assets. Franchise value drops to zero when $\epsilon < \bar{\epsilon}(L_t, \mu_t)$, because in that case the bank loses all its equity capital and is forced to exit the market. Equation (11) is the regulatory capital constraint (if applicable). The bank's equity holding must be higher than the minimum requirement, which equals $\kappa \cdot L$ under the flat-rate capital requirement and is endogenously determined under the risk-based capital requirement (equation (2)). The embedded dynamic programming problem

(equations (12)–(14)) is the dividend policy to be chosen by the bank at the end of each period, which is equivalent to the choice of e_{t+1} . If the bank chooses to issue new equity, d_t is negative and the bank pays an additional cost of λ for each unit of fresh capital.

The solution to this problem has the following properties.

PROPOSITION 1. *There exists a unique function $V^*(e, z)$ that solves the dynamic programming problem (9). The optimal value function is strictly concave in e .*

Proof. See the appendix.

PROPOSITION 2. *(optimal policy functions)*

- (i) *The optimal monitoring effort, μ^* , is a function of L and is independent of z and ϵ . It satisfies $F(L) = \theta'(\mu^*)L$.*
- (ii) *The optimal dividend policy follows a cutoff rule; i.e., there exists a pair (π_1, π_2) with $\pi_1 \leq \pi_2$, for which the bank issues new equity ($d < 0$) if $\pi < \pi_1$, retains all profits ($d = 0$) if $\pi \in [\pi_1, \pi_2]$, and distributes dividends ($d = \pi - \pi_2$) if $\pi > \pi_2$.*

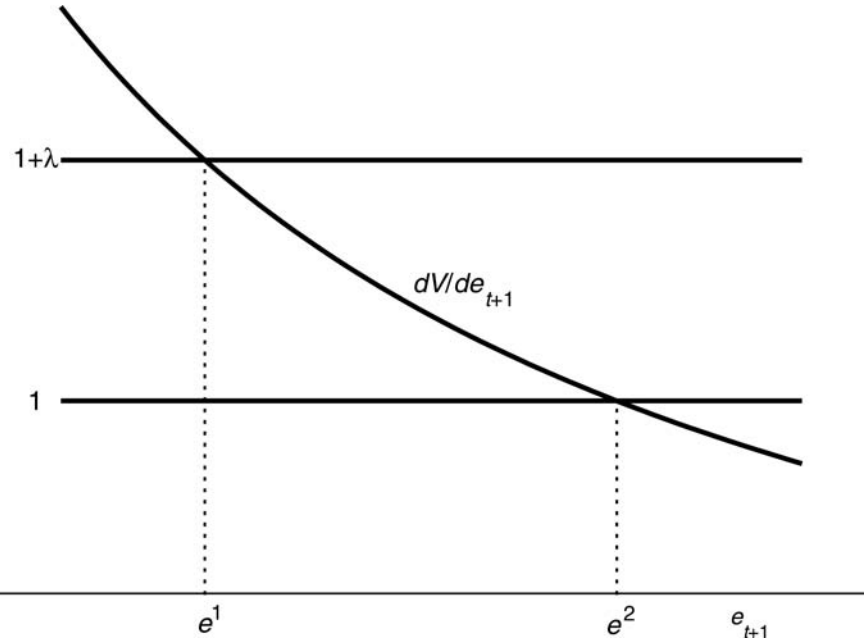
Proof. See the appendix.

Proposition 2(i) suggests that the level of monitoring effort is chosen to maximize the expected net profit in the current period and is independent of the uncertainty in asset returns and the bank's future franchise value. This is because, by assumption, the monitoring effect affects the bank's revenue as an additive component, $\mu_t F(L_t) - \theta(\mu_t)L_t$. As a result, the optimal monitoring effort that maximizes this component also generates a current-period profit outcome that strictly dominates any other solution and, by extension, maximizes the bank's franchise value. This result is very useful, as it reduces the number of free parameters by one in the value function iteration procedure and simplifies the algorithm significantly.

The strict concavity of V^* implies that the dividend policy assumes a simple form as specified in proposition 2(ii). Suppose that the feasible values of e are restricted to a set $[e_{min}, e_{max}]$. In the

embedded dynamic programming problem (13), the marginal cost of bank equity is 1 if the bank does not issue new equity or $1 + \lambda$ if the bank issues new equity. At the same time, the marginal benefit of bank equity, irrespective of whether new equity is issued or not, is downward sloping and equals $V'(e_{t+1}|z_t)$. This trade-off is illustrated in figure 1. Suppose that e^1 (or e^2) is the critical value of e_{t+1} at which the marginal benefit equals the marginal cost if new bank equity is issued (or if no new equity is issued); the bank's dividend policy must follow the cutoff rule as stated above with $\pi_1 = e^1$ and $\pi_2 = e^2$. The reason is as follows. If the bank's net asset value is smaller than e^1 , the bank has the incentive to raise new equity to the target level e^1 , because the increase in the bank's franchise value

Figure 1. The Determination of the Bank's Dividend Policy



Note: The two horizontal lines represent the marginal cost of bank equity when new equity is issued (upper line) or is not issued (lower line). The downward-sloping curve represents the marginal benefit of bank equity.

exceeds the equity issuance cost. By contrast, when $\pi > e^2$, the bank chooses to distribute dividends in the amount of $\pi - e^2$, at which point the marginal decrease in bank value equals the marginal benefit of dividend payment. When the net asset value is in the middle range, the bank prefers to reinvest all profits because it increases the shareholders' value by more than paying the profits out as dividends; nevertheless, no extra equity will be raised, because the cost is relatively too high.

Notice that the two critical values, π_1 and π_2 , change with the state of the business cycle but are independent of bank characteristics. In other words, the "optimal" range of equity holding is the same for all banks. Of course, the choice of equity holding within this range depends largely on the realization of bank-specific shocks, suggesting that the equilibrium size distribution of bank equity exhibits substantial cross-sectional differences (this issue will be reexamined later).

The optimal cutoff rule implies that when a bank chooses its lending strategy (L), it faces the following trade-offs.¹⁶ First, without considering the possibility of bank failure, banks tend to choose an optimal loan volume that maximizes their current-period expected profit. Importantly, this optimal loan volume is the same for all banks because they have the same investment opportunities.¹⁷ In this sense, initial equity size does not play a role, as all banks tend to choose the same portfolio size, which implies that smaller banks would choose higher leverage ratios. Second, the possibility that a bank may fail would affect its lending decision in an important way. As lemma 1 suggests, the probability of bank failure increases in portfolio size and decreases in banks' initial equity holdings. Therefore, in making portfolio decisions, small banks would pay more attention to expected losses in franchise value (due to bank failure) and choose a lower leverage ratio than the one driven by the efficiency concern alone (as discussed above). Lastly, the existence of equity issuance cost adds another layer of complexity to banks'

¹⁶The first-order condition is complicated, so I do not state it explicitly here.

¹⁷Based on equation (3) and proposition 2(i), expected profit in the current period equals $E(\pi|z) = zL^\alpha + \frac{L^{2\alpha-1}}{4c_0} - r^f(L - e)$. It is straightforward to show that there is a uniform L that maximizes the expected profit for all banks.

lending decisions, because the size of bank asset has an ambiguous effect on the magnitude of this cost.¹⁸ As a result, a closed-form solution is ruled out in the model. The remaining part of the paper employs numerical examples to illustrate the properties of the solution and the equilibrium relationship among bank size, leverage, profits, and the probability of failure.

4. Calibration

4.1 Parameterization

The model parameters are calibrated based on the U.S. data in 1981–2002. Table 1 summarizes the input parameters used in the numerical analysis.

- The risk-free rate is fixed at 3 percent, close to the average real risk-free rate (short-term Treasury rate adjusted by inflation) in the United States in the period 1981–2002 (which equals 3.07 percent).

Table 1. Model Parameters for the Calibration

Risk-Free Rate	r^f	1.030
Returns-to-Scale Parameter	α	0.985
Economic State	z_H	1.070
	z_L	1.052
Transition Matrix	H $\rightarrow$ H	0.800
	L $\rightarrow$ L	0.800
Standard Deviation of the Shock	$\sigma_\epsilon(L)$	$\frac{1.05}{L+15}$
Monitoring-Cost Parameter	c_0	6.500
Intertemporal Discount Rate	β	0.950
New-Equity Premium	λ	0.300
Minimum Capital Requirement	κ	0.080

¹⁸Based on equation (8), the two critical net asset values can be transformed into two critical values for bank-specific shocks. Specifically, a bank chooses to issue new equity when the bank-specific shock is greater than $\bar{\epsilon}$ but lower than $\epsilon_1 = \bar{\epsilon} + \frac{e^1}{F(L)}$. The change in loan volume (L) has an ambiguous effect on the probability of new equity issuance, $\Phi(\epsilon_1) - \Phi(\bar{\epsilon})$, and the expected cost of equity issuance.

- The concave return function $F(\cdot)$ is specified as $F(L) = L^\alpha$ with $\alpha = 0.985$. This returns-to-scale parameter is close to the one adopted by Cooley and Quadrini (2001). The parameter must be strictly less than one, because otherwise the optimal size of the bank goes to infinity. In addition, it is chosen close to one to avoid the large discrepancy in investment returns between large and small banks.
- The state of the economy is represented by z . For simplicity, I assume that z only takes two possible states: high (H) or low (L). The probability of staying at the same state in the next period is set to be 0.800, which implies that the same state of economy lasts for four years on average. The bank-specific shock ϵ is assumed to be normally distributed with mean zero and standard deviation σ_ϵ , which will be discussed later.

The choice of two possible states of the economy, z_H and z_L , attempts to match the loan premium observed in the data. I rank the annual loan premium (defined as loan rates minus short-term Treasury rates) during the period 1981–2002 and divide the sample into two subgroups. The average loan premium in the good eleven years is 3.88 percent and in the bad eleven years is 2.12 percent. Hence I set the two shocks as $z_H = 1.070$ and $z_L = 1.052$ (i.e., excess returns of 4 percent and 2.2 percent), respectively.

- The determination of σ , the return volatility, is not straightforward. Following conventional finance theory, I first assume a particular functional form for the relationship between the volatility and the size of bank assets: $\sigma(L) = \frac{a_1}{L+a_2}$. The two constant terms a_1 and a_2 are then chosen so that the Sharpe ratios, defined as the bank's excess return on assets (over the risk-free rate) divided by its standard deviation, are within a reasonable range in equilibrium. Rosen (2005) documents that the median Sharpe ratio for U.S. banks during the period 1977–2003 is about 0.90. For this study, $a_1 = 1.05$ and $a_2 = 15$ so that, in equilibrium, the Sharpe ratios vary between 0.40 and 1.40, with a median of 0.92.
- The cost of monitoring effort is specified as a quadratic form $\theta(\mu) = c_0\mu^2$, where $c_0 = 6.5$. I choose this particular value to offset the decreasing returns-to-scale effect due to the concavity of the return function, so that in equilibrium the

average loan premium matches the target level. Nevertheless, the monitoring-cost parameter is not crucial in this study, and varying this parameter has only a minor impact on the results.

- The intertemporal discount factor equals 0.95. This is very close to the one adopted by Cooley and Quadrini (2001) and is also consistent with the business-cycle literature.
- Following Cooley and Quadrini (2001), the cost of issuing new equity, or the new share's premium, is set to be 30 percent. The magnitude of the equity issuance cost is also in line with the size of the tax burden (40 percent) discussed in Van den Heuvel (2002), an alternative way to introduce the differentiation between equity and debt financing.
- Under the flat-rate capital requirement, the minimum capital-asset ratio equals 8 percent.

The computational procedure of the dynamic programming problem is described in the appendix. The results presented are based on the discretization using a grid with 240 points for bank equity (e), net asset value (π), and loan volume (L).

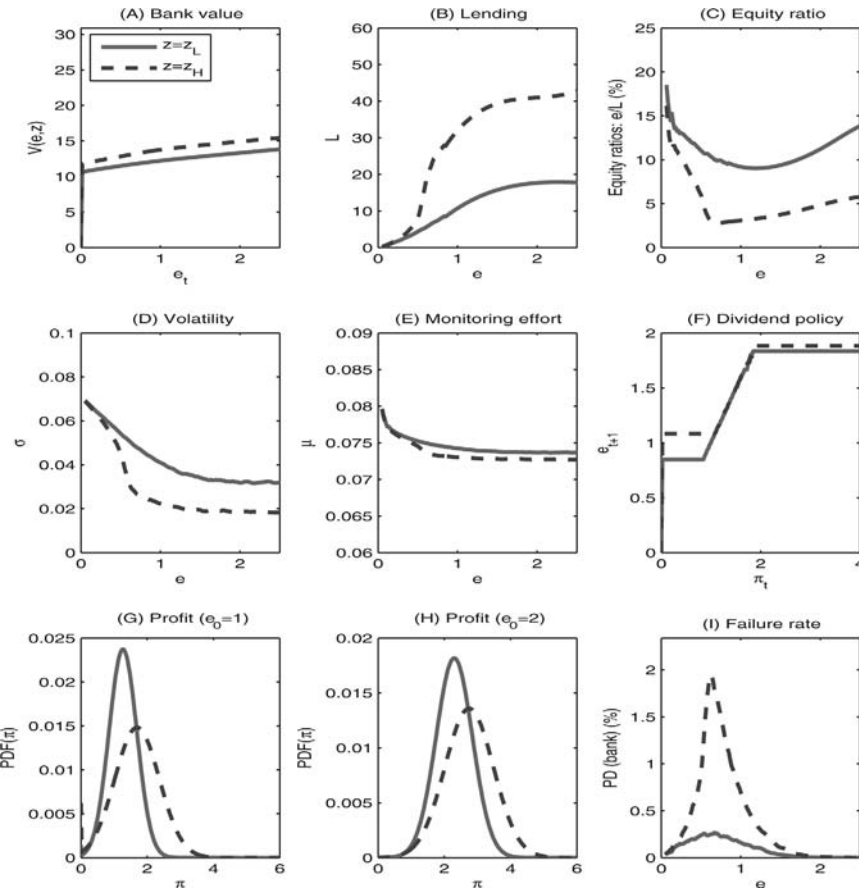
4.2 *A Baseline Model without Capital Regulation*

The baseline model examines the banks' financial decisions in the absence of regulation, i.e., when the capital requirement is zero. This provides a benchmark for judging the impact of different forms of capital regulation. The characteristics of banks' equilibrium behavior are shown in figure 2. Noticeably, the equilibrium financial decisions of banks largely depend on their initial equity holdings (hereafter referred to as "size") and the state of the business cycle.

First of all, panel A of figure 2 indicates that the bank's franchise value is higher in the good state and increases with bank size.

Second, panels B and C plot the bank's lending decisions. Generally, bank lending increases during the expansion period and falls during the downturn period. This procyclical lending behavior is a result of maximizing profits when the average return is high. The optimal leverage ratio (the inverse of equity/asset ratio) can be as high as in the range of 15–35 in the good state and as low as 6–10 in the bad state.

Figure 2. Equilibrium Outcomes under the Baseline Economy



Note: The figure plots the equilibrium outcomes under the baseline economy (without capital regulation), including banks' franchise values (V), the volume of loan portfolio (L), equity/asset ratios, asset-return volatility (σ), monitoring efforts (μ), the dividend policy, the probability distribution of end-of-period net asset values for a small bank (initial equity = 1) and a large bank ($e_0 = 2$), and the probability of bank failure. Notice that the dividend policy is equivalent to choosing the next period's equity (e_{t+1}), and all banks choose the same optimal dividend policy in equilibrium.

Moreover, the optimal leverage ratio differs across banks. As shown in the figure, large and very small banks choose low leverage in equilibrium, while mid-sized banks choose high leverage.¹⁹ This cross-sectional difference is driven by the trade-off between one-period gains (an efficiency concern) and future franchise value (a safety concern), as discussed in section 3. On the one hand, there is an optimal scale of bank lending that maximizes the expected profit of each bank in the current period due to the same lending opportunity (see footnote 17). Other things equal, each bank has the incentive to choose this optimal lending scale, implying that the optimal leverage ratio would be higher for smaller banks. On the other hand, high leverage increases the probability of bank failure and reduces the bank's value in the future. The equilibrium lending decision depends on the relative importance of these two effects. The numerical example suggests that, for very small banks, focusing solely on the efficiency leads to very high leverage and high default risk. The future franchise value cost is so high that these banks will choose a prudent lending strategy (i.e., the safety concern dominates). As bank equity increases, the stronger capital position implies that banks can maximize current-period gains without increasing the probability of default substantially. Hence the optimal leverage ratio increases to reach the balance between the marginal benefit from investment gains and the marginal cost of bank failure. However, at a certain point (when the capital position is strong enough) the efficiency aspect dominates the safety concern and the optimal leverage ratio decreases with bank size. The combined effects explain the "U-shaped" optimal equity/asset ratios as observed in the graph.

Third, panels D and E show the risk profile of bank investments and banks' monitoring efforts. Smaller banks are associated with smaller portfolio investments, which have high return but high risk as well. In order to stay in the business, smaller banks are willing to

¹⁹The optimal dividend policy, as shown in panel F, suggests that bank size is approximately between 1 and 2 for all banks. Although bank size here does not map directly into the observed data, it is reasonable in this paper to interpret banks of size 2 as the largest ones in the industry, banks of size 1 as the small and established banks, and banks of any size in between as mid-sized banks. Banks with a size smaller than 1 represent the smallest ones that are still at their infancy stage.

devote a higher monitoring effort (per unit of asset) to reducing the probability of failure. In addition, when the economic conditions are favorable, high productivity induces banks to lower their monitoring efforts because the marginal benefit of inducing higher returns through this channel is smaller.

Fourth, panel F shows the optimal dividend policy or, equivalently, the choice of next period's initial equity, which follows a cutoff rule as described in proposition 2(ii). As mentioned above, the threshold values in the cutoff rule only depend on the state of the business cycle but are independent of bank characteristics. That is, irrespective of banks' initial equity in the current period, their initial equity in the next period will be distributed within the same "optimal" range. However, the size distribution within this range is different across banks, because it depends on the realization of end-of-period net asset values. Panels G and H illustrate the probability distributions of end-of-period net asset values for two hypothetical banks—a small bank ($e_0 = 1$) and a large bank ($e_0 = 2$). Large banks are more likely to earn higher profits and remain the business leaders in the next period. This is not surprising, because strong capitalization implies that these banks can choose to maximize investment profits without worrying too much about the possibility of failure. By contrast, the growth capacity of small banks is limited due to the concern that a high-leveraged capital structure, which is desirable from the perspective of profit maximization, tends to increase the probability of bank failure substantially.

Lastly, panel I shows the probability of bank failure during the two stages of the business cycle. Overall, the predicted bank failure rates are within a reasonable range, with a peak level of 1–2 percent and a low level of 0.01–0.04 percent. This is in line with the historical observations before a mandatory capital requirement was applied in 1988.²⁰ Importantly, the failure probability differs substantially across banks. In line with the difference in bank leverage, an inverse U-shaped curve is observed across banks of different sizes. The largest and smallest banks have lower leverage, and their default probabilities are very low. By contrast, mid-range banks

²⁰Because the probability of default is very small for all banks, the difference between default-adjusted deposit rates and the risk-free rate is negligible (recall assumption 4).

engage in aggressive lending and, consequently, incur a high risk of default.

Interestingly, the probability of bank failure is much higher in the upturn of the business cycle. This supports the view that business risks build up during the expansion period.²¹ As former Federal Reserve Chairman Alan Greenspan noted, "The worst loans are made at the top of the business cycle" (Greenspan 2001). Borio, Furfine, and Lowe (2001) and Borio and Lowe (2002) document the same pattern that systemic risk tends to build up in the booming period. Notice, however, that in this paper, the higher systemic risk during upswings is not caused by underestimation of the business risk. Instead, the forward-looking banks are fully aware of the risk involved: they are willing to increase their risk exposure in the pursuit of higher returns. In other words, the one-period expected gain from risk taking exceeds the franchise value at risk.

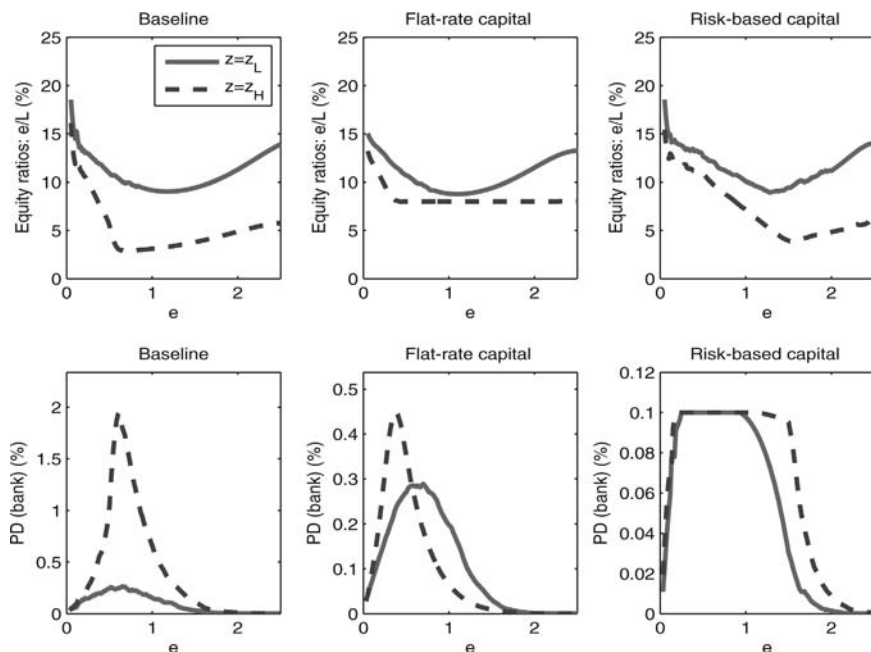
4.3 Equilibrium Outcomes under Capital Regulation

The rest of the paper introduces capital regulation and examines its impact on equilibrium outcomes. The capital requirement is assumed to be exogenously imposed for regulatory reasons that are not explored in this model. Two capital regimes are considered: one is a flat-rate (or risk-insensitive) capital standard that requires a minimum capital/asset ratio of 8 percent, and the other is a risk-based capital rule that ensures a maximum failure probability of 0.1 percent for all banks (defined in equation (2)).

For space reasons, the analysis focuses on the impact of capital regulation on banks' optimal capital structure and the probability of bank failure, as shown in figure 3.²² In general, the imposition of capital regulation has important effects on the bank's financial

²¹In this paper, it is hard to distinguish between the "buildup" and the "materialization" of business risks because all bank loans mature in one period. The simplified assumption on maturity structure misses the realistic situation that there exists a lag for risks to materialize. Therefore, readers should not link directly the realization of bank failures with the state of the business cycle when fitting model predictions with the data. Moreover, for general discussion on the impact of the timing assumption, see Crockett (2000).

²²The equilibrium outcomes under capital regulation are similar to those under the baseline economy. For instance, the bank's franchise value is concave and is strictly increasing in bank size and the state of the economy. In equilibrium,

Figure 3. Equilibrium Outcomes under Capital Regulation

Note: The figure compares the equilibrium equity/asset ratios and equilibrium probability of bank failure under the baseline economy (without capital regulation), under a flat-rate capital regime (a minimum capital/asset ratio of 8 percent), and under a risk-based capital rule (determined by equation (2)).

decisions, but the channel and the magnitude of the effects are quite different under the two capital regimes.

Under a flat-rate capital regime, the capital rule imposes a lower limit on the equity/asset ratio, and its impact on the probability of bank failure is only indirect. Given the countercyclicality of the economic capital, defined as the capital holding under the unregulated economy (see figure 2), the flat minimum capital requirement

all banks choose a cutoff dividend rule as described in proposition 2(ii). Furthermore, large banks tend to lend more and spend fewer resources on monitoring (per unit of asset). These results are not reported but are available upon request.

is more likely to be binding in the good state. Banks are forced to cut lending during the expansion period, and the lower leverage causes the probability of bank failure to be lower than in the baseline economy. By contrast, during downturns the bank's financial decisions are practically unaffected.

The result that capital buffers, defined as actual capital holding minus the minimum capital requirement, are countercyclical under a flat-rate capital rule is consistent with the prediction in Repullo and Suarez (2007) and empirical findings in Ayuso, Pérez, and Saurina (2004) and Stolz and Wedow (2005). Banks hold large capital buffers during economic downturns out of the safety concern, and they hold small capital buffers during upswings. This countercyclical movement in capital buffers has important implications. First, it suggests that capital regulation is useful in containing excessive risk taking at the top of the business cycle. The average loan quality granted during good times is much better than in the baseline economy (lower probability of bank failure), although the improvement is only partial in that the loan risk is still higher during good times (as will be compared with the risk-based capital rule). Second, the fact that the capital constraint is not binding for most banks at the trough of the business cycle suggests that the procyclicality effect may play only a limited role.

Moreover, the cross-sectional difference in economic capital under the baseline economy suggests that the effectiveness of the flat-rate capital rule is different across banks. In good times, the capital constraint is binding for most banks except for the smallest ones. Accordingly, most banks have to reduce the scale of investment to abide by the regulatory constraint. In bad times, banks are willing to hold capital above the minimum regulatory requirement, and the level of capital buffer reflects the difference between economic capital and regulatory capital. Consistent with the baseline results, the very large and very small banks hold larger capital buffers than the others.

Under a risk-based capital rule, the minimum capital requirement is chosen to ensure a minimum level of bank safety (equation (2)). A risk-based capital rule implies that the level of regulatory capital varies not only with the state of the economy but also with the risk profile of bank assets, the latter of which is endogenously determined. Specifically, the new rule rewards banks that have high-quality assets

by lowering their capital requirements and penalizes banks that have low-quality assets by doing the opposite.

The impact of a risk-based capital rule is most obvious if we compare the right-most panels with the left panels in figure 3. The capital regime appears to be quite efficient in improving the safety of individual banks in that the probability of bank failure is successfully contained below the target level. The disciplinary role of the risk-sensitive capital regime is particularly significant in the case of small banks, which are forced by higher capital charges to reduce the scale of bank lending substantially in good times. In other words, the buildup of excessive risk taking by those banks is substantially mitigated. By contrast, large banks' financial decisions are almost unaffected because their strong equity position and low default risk imply that the risk-based capital rule practically imposes no restriction.

An important difference between the two capital rules is their effectiveness over the business cycle. As mentioned above, under the flat-rate capital rule, the capital constraint is more likely to be binding in the good state. By contrast, under the risk-based capital rule, the capital constraint can also be binding in the bad state. This is because, although banks choose a low leverage ratio (in bad times) in the baseline economy, the economy-wide shock is so severe that the bank's probability of failure is still high. As a result, banks have to adjust their lending decisions to meet the safety target specified by the risk-based capital rule. Not surprisingly, among the three economic environments examined in this paper, the variation in bank loan quality over the business cycle (reflected by the probability of bank failure) is the smallest when a risk-based capital rule is imposed.

4.4 The Transition from a Flat-Rate to a Risk-Based Capital Rule

The calibration results permit a reexamination of the impact of the transition from a flat-rate to a risk-based capital regime, including the level of bank capital and the procyclicality issue. I show that the conclusions to the two issues are quite different once one takes into account banks' strategic responses to changes in capital regime, and the impacts are quite different across banks.

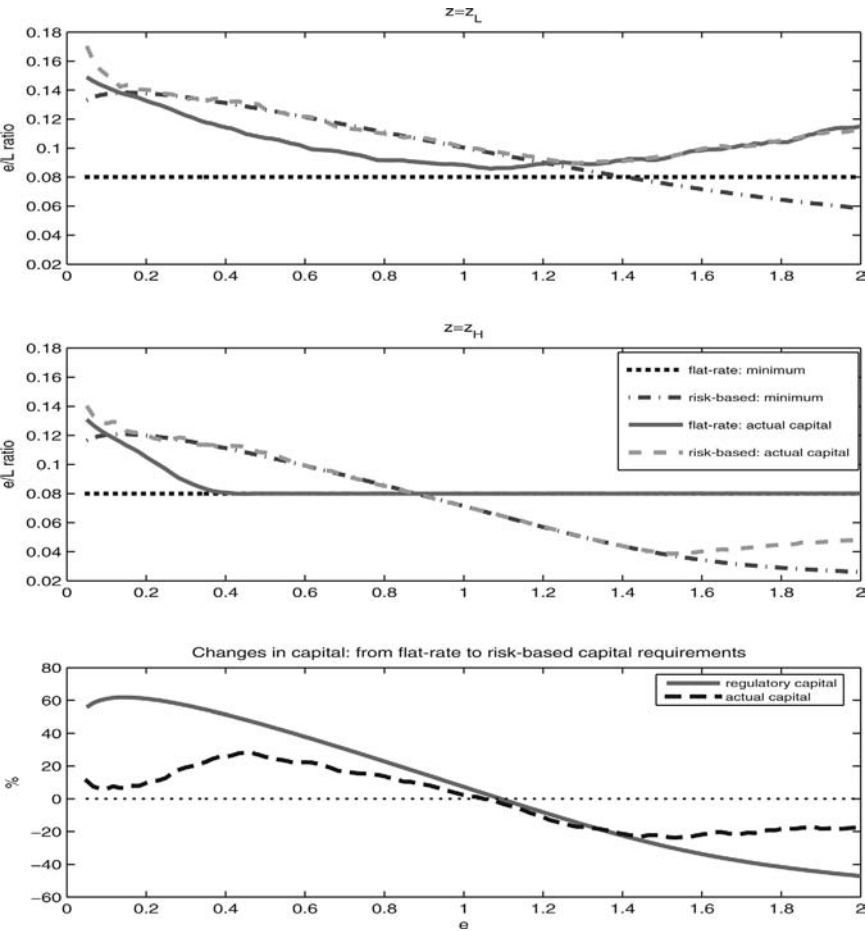
In addressing the first issue, one needs to distinguish between the impact on regulatory minimum requirements and the impact on actual bank capital holdings. The change in regulatory capital is very significant during the transition. The dotted and dash-dotted lines in the top two panels of figure 4, and the solid line in the bottom panel, are indicative of this impact. The risk-sensitive rule favors large (and less risky) banks at the cost of small (and riskier) banks. In particular, it causes the average minimum capital requirement for large banks to be approximately 50 percent lower than under the flat-rate capital regime, and that for small banks to be 50–60 percent higher. These findings are roughly in line with previous studies.

By contrast, the impact of the regulatory rule on actual bank capital holdings is much less pronounced. The solid and dashed lines in the top two panels of figure 4, and the dashed line in the bottom panel, illustrate this impact. Overall, the change in total capital is much smaller than the change in regulatory capital, because of the offsetting effect by endogenous adjustments in banks' capital buffers. Moreover, the impact is less uneven across banks, in that large banks will hold 15–20 percent less capital and small banks will hold approximately 20 percent more capital under the risk-sensitive rule than under the flat-rate one.

The uneven impact across banks is not necessarily undesirable. To some extent, it reflects an important motivation for adopting a risk-sensitive capital regime: to level the playing field. The flat-rate capital requirement does not impose sufficiently high capital requirements on high-risk loans, but at the same time it penalizes high-quality investments. For instance, mid-sized and large banks are forced to forgo high-return, low-risk lending opportunities in good times because of binding capital constraints. By contrast, the risk-based capital regime successfully restores market efficiency by linking capital requirements to asset quality. As the simulation results show, under the risk-based capital regime, the more favorable capital regulation for larger (and safer) banks makes their lending portfolios more similar to those under the baseline economy, while the stricter capital requirement for smaller (and riskier) banks forces them to cut loan extension substantially.

In addressing the second issue, one needs to distinguish between the procyclical movement of bank credit and the procyclicality effect

Figure 4. Comparison of Capital Holdings under Two Capital Regimes



Note: In the top two panels, the dotted and dash-dotted lines plot the minimum capital requirements under two distinctive capital regimes, i.e., a uniform capital/asset ratio of 8 percent or a risk-based capital standard. The solid and dashed lines plot banks' *actual* capital holdings under the two capital regimes. The bottom panel shows the average changes in regulatory capital (solid line) and actual capital (dashed line) if the flat-rate capital rule is replaced by the risk-based capital requirement.

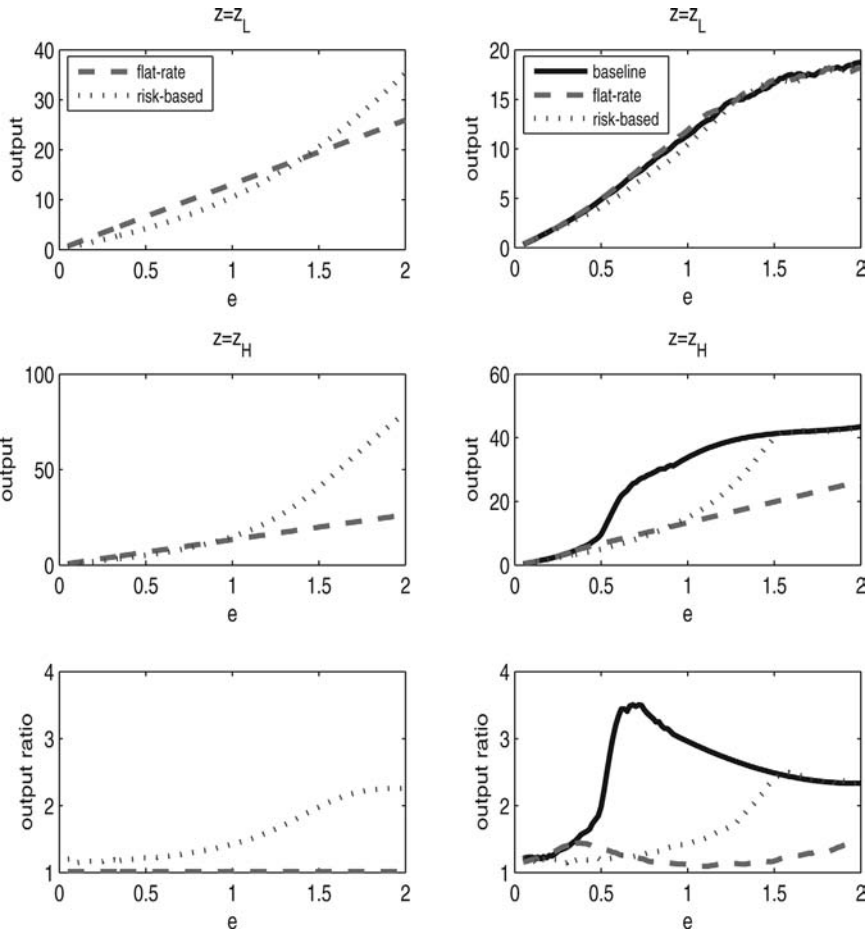
of capital regulation. In particular, the latter refers to *excessive* volatility of bank credit as a result of introduction of capital regulation, compared with the cyclic movement under the unregulated economy. Such distinction is important, because a cyclical movement of bank credit itself could be a manifestation of an optimal allocation of economic resources over the business cycle. By contrast, excessive volatility of bank credit, which in the context of this model is equivalent to an amplified cycle of economic output, could cause potential harm to the real economy and should be avoided.

In this paper, I quantify the magnitude of credit cycle for each bank by calculating the ratio of economic outputs (lending revenue minus the monitoring cost) in good times versus in bad times (shown in the lower right panel in figure 5). Furthermore, under capital regulation, the cycle is further broken down into two components—one attributed to the cyclical movements in capital charges (shown in the lower left panel in figure 5) and the other driven by variation in capital buffers. The comparison of credit cycles between the baseline economy and the one with capital regulation sheds light on the procyclicality effect of capital requirements. Several major observations arise from the calibration results.

First, under both capital regimes, *actual* economic outputs co-move with the business cycle, but the amplitude of the lending cycle is much less volatile than in the baseline economy. This points to an important message that the concern about the procyclicality effect might have been misplaced from the beginning, in that neither the flat-rate nor the risk-based capital regime brings excessive volatility in economic activity. Instead, both capital regimes have a mitigating effect on the cyclical movement in bank credit, because capital regulation constrains banks' excessively risky lending during upswings. It is therefore misleading to label the procyclical movement in bank credit, which is a reflection of the change in investment productivity, as an undesirable effect caused by capital regulation.

Second, the driving forces behind the procyclical movement in bank credit are quite different under the two capital regimes. Under the flat-rate capital requirement (the dashed lines in the upper and middle panels in figure 5), this co-movement is mainly driven by the anticyclical movement in capital buffers, while the capital constraint

Figure 5. Lending Cycles under Two Capital Regimes



Note: The figure plots the level of economic output (lending revenue minus the monitoring cost) for each bank under two distinctive capital regimes, i.e., a uniform capital/asset ratio of 8 percent (dashed lines) or a risk-based capital standard (dotted lines). The left panels refer to the results if banks choose portfolios as indicated by the minimum capital requirements, and the right panels refer to actual outcomes in equilibrium (equilibrium outputs under the baseline economy are also shown in dark lines for comparison purpose). The bottom panels plot the ratio of outputs in the good state versus in the bad state, a measure for the magnitude of lending cycles.

has only a small impact because it merely reflects the difference in the economy-wide shock. By contrast, under the risk-based capital regime, the swing in economic activity is mainly driven by the countercyclicality of capital charges, and the variation in capital buffers has hardly any impact.²³ To some extent, this difference implies that the risk-based capital rule achieves the goal of aligning regulatory capital with actual capital that a bank is willing to hold.

Lastly, the actual cyclic movement of lending activity tends to be bank specific because of distinction in banks' financial decisions. Under the flat-rate capital regime, banks that hold low capital buffers in both states (initial equity ≈ 1) have the smallest credit cycle in the banking industry. Under the risk-based capital regime, the cyclical swing in lending activity is most remarkable for largest banks, as their low risk profile (due to the strong equity position) implies that they are practically not constrained by the capital rule and hence their lending fluctuates freely, reflecting the change in productivity over the business cycle.

Importantly, these results suggest that it is unclear whether the transition toward a risk-sensitive capital rule would amplify or mitigate the credit cycle in the aggregate economy. The calibration results show that the cycle for very small banks is less pronounced under the risk-sensitive capital rule, whereas it becomes more volatile for large banks. The aggregate effect, therefore, depends on the size distribution of the banking industry. For instance, the simulation results in figure 5 suggest that the aggregate credit cycle is less pronounced under the risk-based capital standard than under the flat-rate capital rule if the majority of assets in the banking industry are held by very small banks, and vice versa if large and medium banks play a dominant role.

²³The large swing in minimum capital requirements is driven partly by the short horizon of the portfolio risk assessment. For simplicity reasons, this model assumes that bank assets mature in one period (or one year); therefore, the risk profile of a bank's portfolio is a point-in-time measure. This is consistent with the prevalent practice of using one-year default probability in computing risk-based capital requirements. If risk assessments over a longer time horizon are adopted (e.g., "through-the-cycle" ratings), the variation in regulatory capital over time tends to be smoother. This requires a richer maturity structure of bank assets in a dynamic setting, which is beyond the scope of this paper.

5. The Trade-Off between Efficiency and Safety and the Impact of Capital Regimes

The previous section discusses how the change in capital rules affects banks' equilibrium financial decisions. Nevertheless, there is no indication whether this change is desirable. For instance, lower leverage and higher bank safety are not always appealing if they imply a big sacrifice in lending efficiency and growth potential. The analysis in this section can shed some light on this important issue.²⁴

Although there is no consensus on the objective function of a bank regulator, researchers tend to agree that the introduction of capital regulation can be justified on the basis of improving the safety and soundness of the banking sector and minimizing negative externality associated with bank failure. Following the mainstream of the literature (see, e.g., Dell'Ariccia and Marquez 2006; Hellman, Murdock, and Stiglitz 2000; and Kashyap and Stein 2004), I assume that there are two major components in a regulator's objective function: franchise value and the probability of bank failure, which stand for the profitability (or efficiency) and safety of the banking sector, respectively. For obvious reasons, the regulator favors higher franchise value and lower bank fragility.

There are two ways for a regulator to examine the welfare implications of capital regulation—a *micro* perspective that draws conclusion from the analysis based on individual banks, and a *macro* perspective that focuses on the banking system as a whole. This is in line with the two distinct views regarding the objective of prudential regulation, known as the micro- and macro-prudential perspectives.²⁵ The first view emphasizes the safety of *individual* financial institutions, with an implicit assumption that the safety of individual banks is a sufficient and necessary condition for the

²⁴The following discussion focuses exclusively on two capital regimes examined in this paper. Implicitly, I assume that a regulator only cares about franchise value and bank fragility but nothing else, and he or she is fully aware of banks' responses to various capital rules. The discussion is not intended to address the issue of optimal design of capital regulation.

²⁵See Crockett (2000) and Borio (2003) for further discussion. Nevertheless, the distinction between the two strands of views is not always clear-cut.

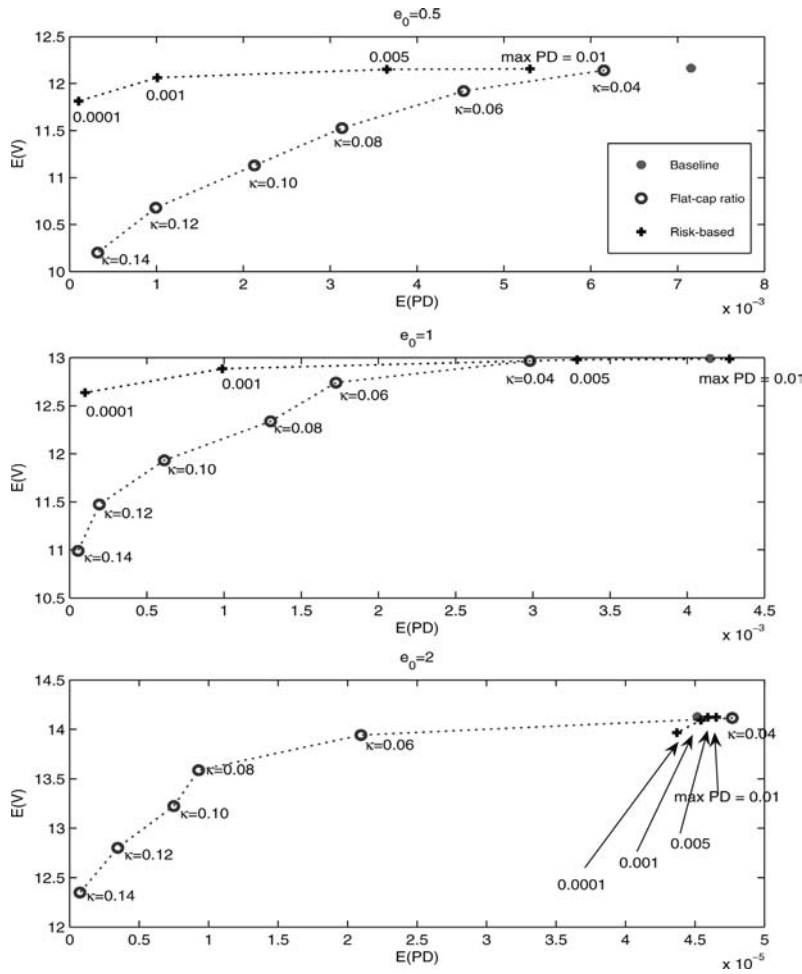
soundness of the banking system. By contrast, the macro perspective analyzes the financial system as a whole. Therefore, systemically more important banks receive more attention from the regulator than small banks that contribute less to the systemic risk.

Figure 6 illustrates the impact of the two capital regimes on the trade-off between efficiency and safety from a micro perspective, using a very small bank (initial equity $e_0 = 0.5$), a small bank ($e_0 = 1$), and a large bank ($e_0 = 2$) as examples. The horizontal axis represents the expected probability of bank failure and the vertical axis represents expected franchise value, using the fact that the good and bad states are equally likely to occur in the long run. The exercise is implemented by studying the flat-rate capital regime with different minimum capital ratio requirements ($\kappa = 2\%, 4\%, \dots, 14\%$) and the risk-based capital regime with different threshold default probabilities (based on the credit VaR definition as specified in equation (2)).

The results are interesting. When the capital standard is very loose, the equilibrium outcome converges to the result under the baseline economy. And the tightening in capital standards (higher uniform capital ratios or lower threshold default probabilities) moves the equilibrium outcome to the lower-left corner. That is, it reduces the probability of bank failure at the cost of banks' franchise value. This is not surprising, since the regulator puts more emphasis on bank safety relative to efficiency. The more interesting finding is that the risk-based capital rule usually yields a higher franchise value for the same degree of bank safety. This implies that the risk-based capital regulation is more appealing from a regulator's perspective. At first glance, this result is not quite intuitive. For instance, comparing the flat-rate capital requirement with $\kappa = 8$ percent and the risk-based capital requirement with the threshold default probability at 0.1 percent, a small bank actually faces a more restrictive capital requirement under the second regime and is forced to choose lower leverage and a lower probability of failure in the current period. However, the bank has a higher franchise value under the risk-sensitive regime than under the flat-rate regime, despite the current-period losses.

The answer to this seeming puzzle lies in the dynamic benefit from the risk-based capital regulation. As I mentioned earlier,

**Figure 6. Comparison between the Two Capital Regimes:
A Micro Perspective**



Note: This figure plots expected franchise value and expected probability of failure for a very small bank (initial equity of 0.5, the top panel), a small bank (initial equity of 1, the middle panel), and a large bank (initial equity of 2, the bottom panel), with the good and bad states equally likely to occur in the long run. The circles refer to equilibrium outcomes under flat-rate capital requirements, with minimum capital ratios varying from 4 percent to 14 percent. The “+” symbols refer to equilibrium outcomes under the credit VaR-based capital regulation, with threshold default rates varying from 1 percent to 0.01 percent.

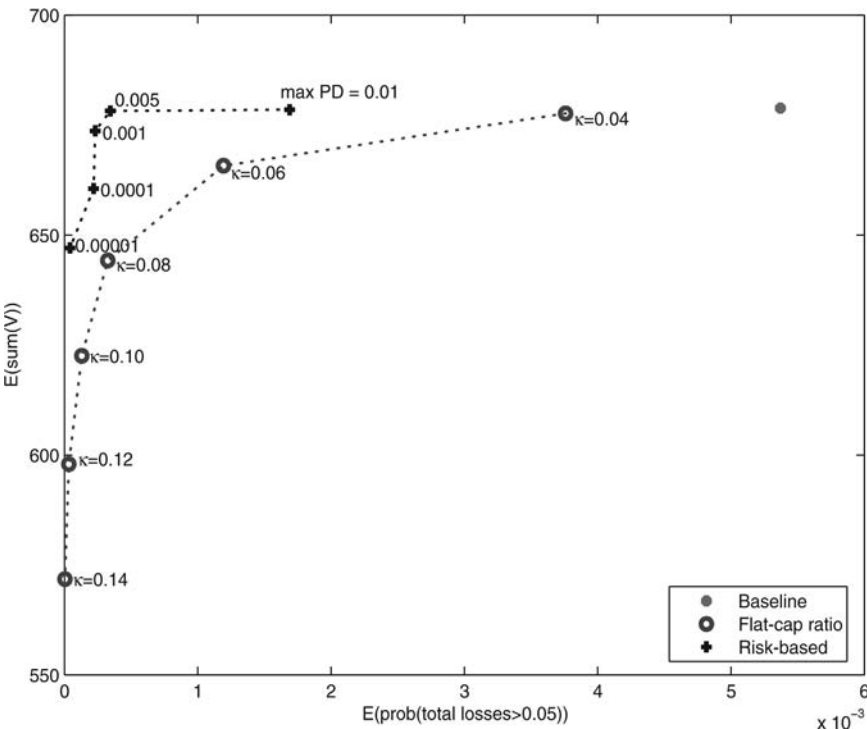
a potential problem with the flat-rate capital rule is that it is not capable of discriminating between sound loans and risky loans; therefore, it tends to hamper efficiency as when it reduces bank fragility. This is most obvious in the case of large banks (the bottom panel), which are forced to forgo sound investment opportunities in order to achieve a very low probability of failure. This conservative lending strategy is no longer appealing, because efficiency losses are too high. From a dynamic perspective, it implies that all banks pay a very high future franchise-value cost under the flat-rate capital regime. This cost tends to overshadow any potential one-period expected gains (as in the small-bank example above) and cause expected franchise value to be significantly lower than under a risk-based capital regime.

In addition, figure 7 illustrates the comparison from a macro perspective, based on a hypothetical banking sector that consists of five large banks ($e_0 = 2$) and fifty small banks ($e_0 = 1$).²⁶ Bank-specific shocks are assumed to be independent across banks. The horizontal axis represents the fragility of the whole banking sector, defined as the probability that more than 5 percent of total bank assets default in the current period. In this example, the threshold is equivalent to the failure of one of the five large banks or the failure of four small banks simultaneously. In this simplified banking sector, the comparison result between the two capital regimes is in line with the micro-based results. The risk-based capital requirement achieves a better balance between lending efficiency and safety, in that (expected) aggregate franchise value is higher for any given level of banking stability.

Clearly, this is only an illustrative example, and the results should not be interpreted as general and conclusive. There are many complicated issues related to a practical implementation of this method, but these issues are beyond the scope of this paper. For instance, the results depend crucially on the size distribution of the banking sector and the interdependence of asset returns across banks, both of which are not explored in this paper.

²⁶In equilibrium, the concentration ratio of the top five banks is about 30–40 percent, which is in line with the survey results in the United States and other industrial countries (see Bank for International Settlements 2004, 131).

**Figure 7. Comparison between the Two Capital Regimes:
A Macro Perspective**



Note: This figure plots expected franchise value and expected probability of failure for a group of banks that consist of five large banks (initial equity of 2) and fifty small banks (initial equity of 1), with the good and bad states equally likely to occur in the long run. The horizontal axis represents the probability that more than 5 percent of total bank assets default in one period, and the vertical axis represents total franchise value. The circles refer to equilibrium outcomes under flat-rate capital requirements, with minimum capital ratios varying from 4 percent to 14 percent. The “+” symbols refer to equilibrium outcomes under the credit VaR-based capital regulation, with threshold default rates varying from 1 percent to 0.001 percent.

6. Concluding Remarks

This paper develops a dynamic equilibrium model and examines the impact of the transition from a flat-rate to a risk-sensitive capital

regulation on banks' financial decisions. Several important policy implications arise. First, in gauging the impact, it is important to take into consideration the potential influence on banks' capital buffers and portfolio decisions. I show that the change in actual capital holdings is much smaller than the change in regulatory capital. Second, the impact of the capital regime differs substantially across banks. The risk-sensitive capital standard leads to a more favorable capital requirement for large (and in this model also less risky) banks. Hence, the cyclical movement in lending behavior is more remarkable for these banks. By contrast, small (and riskier) banks are subject to a stricter capital regulation under the new regime. They have to cut loan extension substantially in good times, when risk is most likely to build up in the absence of regulatory restrictions. As a result, the lending cycle for these banks is actually less volatile under the risk-sensitive regime. Lastly, the analysis suggests that the transition might be welfare improving from a regulator's perspective, in that the risk-based capital regime finds a better balance between safety and efficiency and causes less distortion in loan decisions.

The analysis provides a starting point to examine issues regarding the design of bank capital regulation and the interaction between the banking sector and the real economy. It is by no means a completed task. One caveat of this model is that the asset side of the bank's balance sheet is not explicitly modeled, for the sake of isolating the potential cyclicity coming from the supply side of the bank loan market. This model, however, can serve as a building block for a full dynamic stochastic general equilibrium model with a production sector composed of entrepreneurial firms that rely on relationship bank lending. There are good reasons to believe that the main arguments in this paper will remain valid in the more general setting.

In a dynamic general equilibrium model in which firms do not have alternative funding sources to bank loans, banks' financial decisions will have an important feedback effect on firms' investment decisions. A countercyclical movement in regulatory capital, as specified by a risk-based capital rule, tends to interact with the feedback effect and exacerbate the fluctuations in economic activity. Importantly, this amplification mechanism builds on the assumption that the minimum capital requirement is binding during economic downturns.

The results in this paper do not question the soundness of the above argument. Nevertheless, the results suggest that the above spiral effect can be mitigated for several reasons. First, the assumption that capital constraints are binding during economic downturns is debatable. If banks hold capital buffers at the trough of the business cycle, as shown in this paper, the above amplification mechanism tends to be much smaller. Second, the above argument ignores the positive role of capital regulation in containing excessive risk taking during economic upswings, or the risk buildup period. As shown in this paper, the binding capital constraint in good times may be equally important in containing excessive output fluctuation over the business cycle. Last, but not least, this paper suggests that it is very important to take into account banks' optimal financial decisions. In a general equilibrium model in which the above amplification mechanism is effective, the cost associated with a binding capital constraint during economic downturns is much higher. This will increase banks' incentive to hold more capital buffers. Therefore, it is still important to examine banks' endogenous capital buffer and portfolio decisions in a full-blown dynamic stochastic general equilibrium model.

Appendix

Proof of Proposition 1

- (i) The dynamic program problem is equivalent to

$$V(e_t, z_t) = \max_{(L_t, \mu_t, e_{t+1})} \sum_{z_{t+1}} \left\{ \int_{\bar{\epsilon}}^{\infty} [d(\pi_t, e_{t+1}) + \beta V(e_{t+1}, z_{t+1})] \phi(\epsilon) d\epsilon \cdot P(z_{t+1} | z_t) \right\} \quad (15)$$

subject to equations (6), (7), (8), (10), (11), and (13).

It suffices to prove the existence of a unique solution V^* by showing that the $V(\cdot)$ function satisfies the two Blackwell's sufficient conditions and therefore is a contraction mapping. To check the two conditions, perform the following steps:

- (*Monotonicity*) Take $V_1 \geq V_2$, then

$$\begin{aligned} T(V_1) &= \max E[d(L^*, \mu^*, e^*) + \beta V_1(e^*, z_{t+1})] \\ &\geq \max E[d(L^*, \mu^*, e^*) + \beta V_2(e^*, z_{t+1})] \\ &= T(V_2). \end{aligned}$$

- (*Discounting*) For any positive real number $c > 0$,

$$\begin{aligned} T(V + c) &= E\{d(L^*, \mu^*, e^*) + \beta[V(e^*, z_{t+1}) + c]\} \\ &= E\{d(L^*, \mu^*, e^*) + \beta V(e^*, z_{t+1})\} + c \\ &= T(V) + c. \end{aligned}$$

- (ii) The concavity of the $V^*(e)$ function can be proved based on theorem 4.8 in Stokey, Lucas, and Prescott (1989, 81). For simplicity, I only prove the case in which z is constant, but the proof can easily be extended to the case in which z is random.

First, it is straightforward that the $d(\pi_t, e_{t+1})$ function is continuous and bounded if we can restrict feasible values of e within the range $[e_{min}, e_{max}]$, where $e_{min} = 0$ and e_{max} is sufficiently large so that banks' equity position will never exceed the boundary.

Second, the $d(\cdot)$ function is concave. That is, for any (π_1, e'_1) , (π_2, e'_2) and all $h \in [0, 1]$,

$$d(\pi_3, e'_3) \geq hd(\pi_1, e'_1) + (1 - h)d(\pi_2, e'_2),$$

where $\pi_3 \equiv h\pi_1 + (1 - h)\pi_2$ and $e'_3 \equiv he'_1 + (1 - h)e'_2$. This is because of the following:

- If $\pi_1 \geq e'_1$ and $\pi_2 \geq e'_2$, then $\pi_3 \geq e'_3$ and $d(\pi_3, e'_3) = hd(\pi_1, e'_1) + (1 - h)d(\pi_2, e'_2)$.
- Similarly, if $\pi_1 < e'_1$ and $\pi_2 < e'_2$, then $\pi_3 < e'_3$ and $d(\pi_3, e'_3) = hd(\pi_1, e'_1) + (1 - h)d(\pi_2, e'_2)$.

- If $\pi_1 \geq e'_1$, $\pi_2 < e'_2$, and $\pi_3 \geq e'_3$, then

$$\begin{aligned} d(\pi_3, e'_3) &= h(\pi_1 - e'_1) + (1 - h)(\pi_2 - e'_2) \\ &> h(\pi_1 - e'_1) + (1 - h)(1 + \lambda)(\pi_2 - e'_2) \\ &= hd(\pi_1, e'_1) + (1 - h)d(\pi_2, e'_2). \end{aligned}$$

- If $\pi_1 \geq e'_1$, $\pi_2 < e'_2$, and $\pi_3 < e'_3$, then

$$\begin{aligned} d(\pi_3, e'_3) &= h(1 + \lambda)(\pi_1 - e'_1) + (1 - h)(1 + \lambda)(\pi_2 - e'_2) \\ &> h(\pi_1 - e'_1) + (1 - h)(1 + \lambda)(\pi_2 - e'_2) \\ &= hd(\pi_1, e'_1) + (1 - h)d(\pi_2, e'_2). \end{aligned}$$

Finally, the feasible set of e' is convex if we restrict the feasible value of e within $[e_{min}, e_{max}]$.

Based on the above three properties, the $V^*(e)$ function is strictly concave in e .

Proof of Proposition 2

To determine the optimal monitoring effort in the current period, it is worth noticing that future bank value, $V(e_{t+1}, z_{t+1})$, is independent of μ_t . Therefore, the optimal μ_t should maximize

$$\int_{\bar{\epsilon}}^{\tilde{\epsilon}} (\pi_t - e_{t+1})(1 + \lambda)\phi(\epsilon)d\epsilon + \int_{\tilde{\epsilon}}^{\infty} (\pi_t - e_{t+1})\phi(\epsilon)d\epsilon, \quad (16)$$

where $\tilde{\epsilon}$ is the critical value below which new equity needs to be raised. It is determined by $(\tilde{\epsilon} - \bar{\epsilon})F(L) = e_{t+1}$.

The first-order condition of the maximization problem can be rearranged as

$$\frac{d\bar{\epsilon}}{d\mu} \cdot [e_{t+1}(1 + \lambda)\phi(\bar{\epsilon}) - (1 + \lambda)F(L)(\Phi(\tilde{\epsilon}) - \Phi(\bar{\epsilon})) - F(L)(1 - \Phi(\tilde{\epsilon}))] = 0.$$

Since the bank chooses μ_t at the beginning of period t , at which time the choice of e_{t+1} is still uncertain, the above condition is satisfied if and only if $\frac{d\bar{\epsilon}}{d\mu} = 0$. From equations (4) and (5), it is straightforward to derive that $\frac{d\bar{\epsilon}}{d\mu} = \frac{\theta'(\mu)L - F(L)}{(1 - \Phi(\bar{\epsilon}))F(L)}$. Therefore,

$$\left. \frac{d\bar{\epsilon}}{d\mu} \right|_{\mu=\mu^*} = 0 \quad \Rightarrow \quad F(L) = \theta'(\mu^*)L. \quad (17)$$

In addition, given the strict concavity of V^* , the dividend policy assumes a cutoff rule for a given initial bank equity. Since this is quite straightforward, I leave out the strict proof here, but the intuitions are described in the text.

Computational Procedure

The computational procedure for the dynamic programming is based on value function iteration as introduced in Judd (1998). The value function is solved based on the discretization of the state variable e_t , the quasi-state variable π_t , and control variables L_t and e_{t+1} . The main features of the numerical algorithm are as follows:

- Set the minimum and maximum values of bank equity (e) and end-of-period net profit (π).
- The range of bank lending (L) is set within $[e, e/\kappa]$ under the flat-rate capital requirement or within $[e, L_{max}]$ under the risk-based capital regime (L_{max} is calculated from equation (2) by setting $e_{min} = e$).
- For each pair of (e, L) , calculate the corresponding $\bar{e}$ and μ based on equations (6) and (7) and proposition 2(i).
- For each pair of (e, L) , calculate the probability distribution of end-of-period net asset values.
- Guess initial values of the $V(e)$ function for $z = z_H$ and $z = z_L$.
- Solve for the $W(\pi, z)$ function and the policy function e' for given π , e , and z .
- Based on the W function, solve for the $V(e, z)$ function and the policy function L for given e and z . The found values are the new guesses of the V function. The procedure is then restarted from the previous step until converged.

References

- Acharya, V. 2003. "Is the International Convergence of Capital Adequacy Regulation Desirable?" *Journal of Finance* 58 (6): 2745–81.

- Allen, L., and A. Saunders. 2004. "Incorporating Systemic Influences into Risk Measurements: A Survey of the Literature." *Journal of Financial Services Research* 26 (2): 161–92.
- Ayuso, J., D. Pérez, and J. Saurina. 2004. "Are Capital Buffers Pro-Cyclical? Evidence from Spanish Panel Data." *Journal of Financial Intermediation* 13 (2): 249–64.
- Bank for International Settlements. 2004. "BIS 74th Annual Report." Basel, Switzerland.
- Barrios, V. E., and J. M. Blanco. 2003. "The Effectiveness of Bank Capital Adequacy Regulation: A Theoretical and Empirical Approach." *Journal of Banking and Finance* 27: 1935–58.
- Basel Committee on Banking Supervision. 1988. "International Convergence of Capital Measurement and Capital Standards." BCBS. Basel, Switzerland.
- . 2006. "International Convergence of Capital Measurement and Capital Standards: A Revised Framework—Comprehensive Version." BCBS. Basel, Switzerland. Available at www.bis.org/publ/bcbs128.pdf.
- Berger, A. N., R. J. Herring, and G. Szegö. 1995. "The Role of Capital in Financial Institutions." *Journal of Banking and Finance* 19 (3–4): 393–430.
- Berger, A. N., N. H. Miller, M. A. Petersen, R. G. Rajan, and J. C. Stein. 2005. "Does Function Follow Organizational Form? Evidence from the Lending Practices of Large and Small Banks." *Journal of Financial Economics* 76 (2): 237–69.
- Bernanke, B. S., and A. S. Blinder. 1988. "Credit, Money, and Aggregate Demand." *American Economic Review Papers and Proceedings* 78 (2): 435–39.
- Bikker, J., and P. Metzmakers. 2004. "Is Bank Capital Pro-cyclical? A Cross-Country Analysis." DNB Working Paper No. 009.
- Borio, C. E. V. 2003. "Towards a Macroprudential Framework for Financial Supervision and Regulation?" BIS Working Paper No. 128.
- Borio, C., C. Furfine, and P. Lowe. 2001. "Procyclicality of the Financial System and Financial Stability: Issues and Policy Options." BIS Paper No. 1, 1–57.
- Borio, C., and P. Lowe. 2002. "Asset Prices, Financial and Monetary Stability: Exploring the Nexus." BIS Working Paper No. 114.

- Carter, D. A., and J. E. McNulty. 2005. "Deregulation, Technological Change, and the Business-Lending Performance of Large and Small Banks." *Journal of Banking and Finance* 29 (5): 1113–30.
- Catarineu-Rabell, E., P. Jackson, and D. P. Tsomocos. 2005. "Procyclicality and the New Basel Accord—Banks' Choice of Loan Rating System." *Economic Theory* 26 (3): 537–57.
- Cole, R. A., L. G. Goldbert, and L. J. White. 2004. "Cookie Cutter vs. Character: The Micro Structure of Small Business Lending by Large and Small Banks." *Journal of Financial and Quantitative Analysis* 39 (2): 227–51.
- Cooley, T. F., and V. Quadrini. 2001. "Financial Markets and Firm Dynamics." *American Economic Review* 91 (5): 1286–1310.
- Corcostegui, C., L. González-Mosquera, A. Marcelo, and C. Trucharte. 2002. "Analysis of Procyclical Effects on Capital Requirements Derived from a Rating System." Working Paper.
- Crockett, A. 2000. "Marrying the Micro- and Macro-Prudential Dimensions of Financial Stability." Speech before the Eleventh International Conference of Banking Supervisors, September 21, Basel, Switzerland.
- Dell'Ariccia, G., and R. Marquez. 2006. "Competition among Regulators and Credit Market Integration." *Journal of Financial Economics* 79 (2): 401–30.
- Diamond, D. W., and P. H. Dybvig. 1983. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy* 91 (3): 401–19.
- Elizalde, A., and R. Repullo. 2006. "Economic and Regulatory Capital in Banking: What Is the Difference?" Working Paper.
- Ervin, D. W., and T. Wilde. 2001. "Pro-cyclicality in the New Basel Accord." *Risk* 14: 28–32.
- Estrella, A. 2004a. "Bank Capital and Risk: Is Voluntary Disclosure Enough?" *Journal of Financial Services Research* 26 (2): 145–60.
- . 2004b. "The Cyclical Behavior of Optimal Bank Capital." *Journal of Banking and Finance* 28 (6): 1469–98.
- Gordy, M. 2003. "A Risk-Factor Model Foundation for Ratings-Based Bank Capital Rules." *Journal of Financial Intermediation* 12 (3): 199–232.
- Gordy, M. B., and B. Howells. 2006. "Procyclicality in Basel II: Can We Treat the Disease without Killing the Patient?" *Journal of Financial Intermediation* 15: 393–417.

- Greenspan, A. 2001. "The Financial Safety Net." Speech at the 37th Annual Conference on Bank Structure and Competition of the Federal Reserve Bank of Chicago, May 10, Chicago, IL.
- Group of Ten. 2001. "Report on Consolidation in the Financial Sector." Available at www.imf.org/external/np/g10/2001/01/Eng/index.htm.
- Heid, F. 2005. "Cyclical Implications of Minimum Capital Requirements." Deutsche Bundesbank Discussion Paper No. 06/2005.
- Hellmann, T. F., K. C. Murdock, and J. E. Stiglitz. 2000. "Liberalization, Moral Hazard in Banking, and Prudential Regulation: Are Capital Requirements Enough?" *American Economic Review* 90 (1): 147–65.
- Jokivuolle, E., and K. Kauko. 2001. "The New Basel Accord: Some Potential Implications of the New Standards for Credit Risk." Bank of Finland Discussion Paper No. 2/2001.
- Jordan, J., J. Pek, and E. Rosengren. 2003. "Credit Risk Modeling and the Cyclicity of Capital." Working Paper.
- Judd, K. 1998. *Numerical Methods in Economics*. Cambridge, MA: MIT Press.
- Kashyap, A. K., and J. C. Stein. 1994. "Monetary Policy and Bank Lending." In *Monetary Policy*, ed. N. G. Mankiw, 221–56. Chicago: University of Chicago Press.
- . 2004. "Cyclical Implications of the Basel II Capital Standards." *Economic Perspective* (Federal Reserve Bank of Chicago) 28 (1st Quarter): 18–31.
- Kim, H.-W., and H. Lee. 2006. "Bank Capital Regulation and Procyclicality of Bank Lending: Implications for Basel II Implementation." Working Paper.
- Koopman, S. J., A. Lucas, and P. Klaassen. 2005. "Empirical Credit Cycles and Capital Buffer Formation." *Journal of Banking and Finance* 29 (12): 3159–79.
- Kupiec, P. H. 2004. "Capital Adequacy and Basel II." FDIC Center for Financial Research Working Paper No. 2004-02.
- Office of the Comptroller of the Currency (OCC), Board of Governors of the Federal Reserve System (FRB), Federal Deposit Insurance Corporation (FDIC), and Office of Thrift Supervision (OTS). 2004. "Summary Findings of the Fourth Quantitative Impact Study." Mimeo.

- Peura, S., and E. Jokivuolle. 2004. "Simulation Based Stress Tests of Banks' Regulatory Capital Adequacy." *Journal of Banking and Finance* 28 (8): 1801–24.
- Repullo, R., and J. Suarez. 2007. "The Procyclical Effects of Basel II." Working Paper.
- Rosen, R. J. 2005. "Switching Primary Federal Regulators: Is It Beneficial for U.S. Banks?" *Economic Perspective* (Federal Reserve Bank of Chicago) 29 (3rd Quarter): 16–33.
- Santos, J. A. C. 2001. "Bank Capital Regulation in Contemporary Banking Theory: A Review of the Literature." *Financial Markets, Institutions and Instruments* 10 (2): 41–84.
- Segoviano, M. A., and P. Lowe. 2002. "Internal Ratings, the Business Cycle and Capital Requirements: Some Evidence from an Emerging Market Economy." BIS Working Paper No. 117.
- Stokey, N., R. Lucas, and E. Prescott. 1989. *Recursive Methods in Economics Dynamics*. Harvard University Press.
- Stolz, S., and M. Wedow. 2005. "Banks' Regulatory Capital Buffer and the Business Cycle: Evidence for German Savings and Cooperative Banks." Deutsche Bundesbank Discussion Paper.
- Van den Heuvel, S. 2002. "The Bank Capital Channel of Monetary Policy." Working Paper.

Preference Heterogeneity in Monetary Policy Committees*

Alessandro Riboni and Francisco J. Ruge-Murcia
Département de sciences économiques, Université de Montréal

This short paper employs individual voting records of the Monetary Policy Committee (MPC) of the Bank of England to study heterogeneity in policy preferences among committee members. The analysis is carried out using a simple generalization of the standard New Keynesian framework that allows members to differ in the weight they give to output compared with inflation stabilization and in their views regarding optimal inflation and natural output. Results indicate that, qualitatively, MPC members are fairly homogeneous in their policy preferences, but there are systematic quantitative differences in their policy reaction functions that are related to the nature of their membership and career background.

JEL Codes: E4, E5, D7.

1. Introduction

This short paper uses the individual voting records of the Monetary Policy Committee (MPC) of the Bank of England to study the extent and nature of preference heterogeneity among committee members. Understanding committees is important because many central banks use a committee structure to formulate policy. This is the case, for example, in seventy-nine of the eighty-eight central banks in Fry et al. (2000). At the theoretical level, we use a simple generalization of the standard New Keynesian framework that allows members to differ in the relative weight they attach

*Work on this research project was carried out while the second author was Visiting Scholar at the Dutch National Bank. The financial support of the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged. Corresponding author: Francisco J. Ruge-Murcia, Département de sciences économiques, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal (Québec) H3C 3J7, Canada.

to output versus inflation stabilization and in their views regarding optimal inflation and natural output. Under the assumption of sincere voting, individual reaction functions are derived and estimated for each member and for the committee as a whole. The model implies that, given the economy parameters and private-sector expectations, committee members agree on their reaction to the expected output gap and demand shocks and disagree on their reaction to expected inflation and supply shocks. In addition, members will differ systematically in their preferred interest rate even if they share the same inflation target and estimate of the natural output rate.

A first snapshot of the MPC voting records reveals that disagreement is frequent at the policy meetings of the Bank of England. In about two thirds of meetings between June 1997 and June 2006, decisions were not unanimous. Summary statistics like those computed, for example, by Gerlach-Kristen (2003) and Spencer (2006) are suggestive of heterogeneous preferences over policy. Both studies conclude that external members appear to have different voting patterns than internal ones, tend to dissent more frequently, and—when they dissent—are more likely to vote for a lower interest rate. These results mirror the ones obtained by the literature on the Federal Open Market Committee (FOMC), which usually finds that Federal Reserve Board members appear to prefer more expansionary policy than the Reserve Bank presidents (see Belden 1989, Gildea 1990, Havrilesky and Schweitzer 1990, and Havrilesky and Gildea 1991). The aforementioned literature uses dissenting votes only and analyzes differences in the voting behavior of the two groups by comparing dissenting frequencies. Tootell (1991, 1999), whose approach is followed here, uses instead all votes and investigates heterogeneity within the FOMC by comparing coefficients of individual policy reaction functions. Contrary to the previous literature, he finds no evidence of a systematic difference in voting patterns between Board members and Bank presidents. Individual reaction functions are also estimated by Chappell, McGregor, and Vermilyea (2005). The literature has also considered other factors that may have an effect on FOMC voting behavior, such as career backgrounds (see Gildea 1990 and Havrilesky and Schweitzer 1990) and regional affiliations. For example, Meade and Sheets (2005) find that regional unemployment rates have a statistically significant

effect on the voting patterns of Bank presidents and Board members alike.

The main empirical results are the following. First, committee members behave in agreement with economic theory in the sense that (i) they generally favor higher (lower) interest rates when inflation (unemployment) is expected to increase and (ii) the overidentification restrictions of the model are not rejected by the data. Thus, in qualitative terms, MPC members are fairly homogeneous in their policy preferences. Second, the joint hypothesis that individual reaction-function coefficients are equal to those of the committee is rejected by the data. Hence, in quantitative terms, there is heterogeneity in policy reaction functions. Third, this heterogeneity appears to be systematic in that individuals who favor a strong response to expected inflation or unemployment do so regardless of the forecasts horizon. Fourth, systematic heterogeneity appears to be related to the nature of the membership (whether external or internal) and individual career background (whether academia, private sector, or Bank of England). These two characteristics are not independent, because most academics and individuals with private-sector experience are external members. In general, we find that external members react more strongly to unemployment than internal members, but there is no clear pattern regarding their reaction to inflation.

The paper is organized as follows: section 2 describes the model that motivates the empirical specification, section 3 reports empirical results, and section 4 concludes.

2. The Model

Consider a monetary policy committee composed of N members. Assume that members have heterogeneous policy preferences in the sense that they may attach different relative weights to output versus inflation stabilization and may have different views regarding the optimal inflation rate and the true natural output level. The payoff of member $n \in \{1, 2, \dots, N\}$ in each period is

$$U_n(\pi_t, x_t) = -(1/2)((\pi_t - \pi_{n,t}^*)^2 + \alpha_n(x_t - x_{n,t}^*)^2), \quad (1)$$

where π_t is the inflation rate; x_t is an output measure; $\pi_{n,t}^*$ and $x_{n,t}^*$ are, respectively, the optimal inflation rate and the natural level of x_t according to member n ; and $\alpha_n > 0$ is the relative weight that member n attaches to output stabilization. Normalizing the inflation weight to one entails no loss of generality. The special case where members agree on a (possibly time-varying) inflation target corresponds to $\pi_{n,t}^* = \pi_t^*$ for all n . The difference between member n 's estimate of the natural rate and its true value is assumed to be well approximated by

$$x_t^* - x_{n,t}^* = \nu_{n,t}, \quad (2)$$

where $\nu_{n,t}$ is an exogenous autocorrelated disturbance with zero mean and constant conditional variance. This formulation allows the special cases where all members share the same estimate of the natural rate when $\nu_{n,t} = \nu_t$ for all n .¹

The behavior of the private sector is summarized by

$$x_t = x_t^* - \phi(i_t - E_t\pi_{t+1}) + E_t(x_{t+1} - x_{t+1}^*) + \varepsilon_t, \quad (3)$$

$$\pi_t = \lambda(x_t - x_t^*) + \beta E_t\pi_{t+1} + \eta_t, \quad (4)$$

where i_t is the nominal interest rate; $\phi, \lambda, \beta > 0$ are constant parameters; E_t denotes the public's expectations conditional on information available at time t ; and ε_t and η_t are exogenous autocorrelated disturbances with zero mean and constant conditional variance. The public is assumed to form expectations rationally, and its information set at time t includes current and previous realizations of all model variables and shocks. Equation (3) is an IS curve and equation (4) is the New Keynesian Phillips curve. The IS curve may be derived from a linearized Euler equation for consumption and the resource constraint (see McCallum and Nelson 1999). The New Keynesian Phillips curve may be derived from the linearized pricing decision of a profit-maximizing monopolistic competitor in a sticky-price environment.

¹It is easy to generalize the model to the case where member n 's estimate of the natural rate differs systematically from its true value by introducing a member-specific constant in equation (2). This generalization does not change the reduced form of the reaction function estimated below, but it changes the structural interpretation of its intercept.

Committee members are assumed to sincerely cast their vote. In each period and under discretion,² the policymaker n computes x_t and π_t to maximize

$$E_t \sum_{\tau=0}^{\infty} \beta^{\tau} U_n(\pi_{t+\tau}, x_{t+\tau})$$

subject to the aggregate supply curve and the IS curve. Given the optimal values of π_t and x_t , the interest rate preferred by member n is determined by equation (3). Since private-sector expectations are taken as given by policymakers, future inflation and output are not affected by current decisions, and the first-order condition reads

$$\lambda(\pi_t - \pi_{n,t}^*) + \alpha_n(x_t - x_{n,t}^*) = 0. \quad (5)$$

Using (5), (2), and the aggregate relations (3) and (4), one can write

$$\begin{aligned} i_{n,t}^* = & - \left(\frac{\lambda}{\phi(\alpha_n + \lambda^2)} \right) \pi_{n,t}^* + \left(1 + \frac{\lambda\beta}{\phi(\alpha_n + \lambda^2)} \right) E_t \pi_{t+1} \\ & + \left(\frac{1}{\phi} \right) E_t (x_{t+1} - x_{t+1}^*) + \left(\frac{\lambda}{\phi(\alpha_n + \lambda^2)} \right) \eta_t \\ & + \left(\frac{\alpha_n}{\phi(\alpha_n + \lambda^2)} \right) \nu_{n,t} + \left(\frac{1}{\phi} \right) \varepsilon_t, \end{aligned} \quad (6)$$

where $i_{n,t}^*$ is member n 's preferred interest rate.

Given the economy parameters and private-sector expectations, three empirical implications follow from equation (6). First, committee members agree on their policy reaction to the expected output gap and demand shocks. To see this, note that the coefficients of $E_t(x_{t+1} - x_{t+1}^*)$ and ε_t are both equal to $1/\phi$, which is the inverse of the slope of the IS curve. Second, committee members disagree in their reaction to expected inflation and supply shocks. This follows from the observation that the coefficients of $E_t \pi_{t+1}$ and η_t depend on the idiosyncratic preference parameter α_n . Finally, even if

²The model is solved under discretion because the view taken here is that inflation targeting, as it is actually practiced in the United Kingdom and other countries, is not a rule in the classical sense. Bernanke and Mishkin (1997) argue that "inflation targeting ... contains a considerable degree of what most economists would define as policy discretion."

members share the same inflation target and estimate of the natural rate (meaning that $\pi_{n,t}^* = \pi_t^*$ and $\nu_{n,t} = \nu_t$ for all n), they would disagree in their preferred interest rate because the coefficients of $\pi_{n,t}^*$ and $\nu_{n,t}$ also depend on α_n .

The reason why members agree in their response to demand shocks but disagree in their response to supply shocks is the following. In the New Keynesian model, the effect of demand shocks on the output gap may be offset directly by changes in the nominal interest rate (see equation (3)). However, offsetting the effect of supply shocks on inflation requires a monetary-policy-induced change in the output gap; consequently, committee members face a trade-off between inflation and output stabilization.

3. Empirical Analysis

3.1 Data

The voting data consist of the preferred interest rates stated by each member of the Monetary Policy Committee of the Bank of England in all the meetings held between June 1997 and June 2006. The sample starts with the first meeting of the MPC and ends at the time the data was collected. We use both assenting and dissenting votes. This is possible because the minutes record the names of members in favor of the Governor's proposal and the names and preferred policy options of dissenting members. If a member assents, we attribute to that member the policy adopted by the committee. If a member dissents, we attribute to that member his or her declared policy preference as stated in the minutes.³ The underlying assumption is that committee members cast their votes sincerely.

The MPC consists of nine members, of which five are internal (that is, chosen from within the ranks of bank staff) and four are

³In this sense, we treat assents and dissents symmetrically. The same approach is followed in Tootell (1991, 1999). Chappell, McGregor, and Vermilyea (2005) also use all votes but argue that a dissent favoring a move to, say, tighten is "behaviorally different from an assent that concurs with a committee move to tighten." In their empirical specification, dissents only arise when disagreement is sufficiently acute. Recent game-theoretical models of dissent (see Groseclose and Milyo 2006 and Seidmann 2006) point out that the choice of dissenting or assenting may also depend on strategic considerations (for example, on the probability of being pivotal).

external appointees. Internal members are nominated by the Governor, while external members are appointed by the Chancellor. Meetings are chaired by the Governor of the Bank of England and take place monthly, usually on the Wednesday and Thursday after the first Monday of each month. Decisions are made by simple majority on a one-person, one-vote basis. As of June 2006, twenty-four individuals have been members of the MPC, but six of them have voting records that are too short to allow meaningful empirical analysis.⁴ After excluding these individuals, the sample reduces to 901 votes stated by eighteen MPC members during 109 meetings.⁵

For the estimation of reaction functions, we also use the time series of inflation, unemployment, industrial production, and money growth. Monthly observations of the Retail Price Index (RPI), the Retail Price Index excluding mortgage-interest payments (RPIX), the Consumer Price Index (CPI), the unemployment rate, and the Index of Industrial Production (IIP) were taken from the web site of the UK National Statistics (www.statistics.gov.uk). The stock of money is measured by the end-of-month value of outstanding M4 and was taken from the web site of the Bank of England (www.bankofengland.co.uk).

3.2 *Econometric Strategy*

The reaction function of member n may be written in reduced form as

$$i_{n,t} = a_n + b_n E_t \pi_{t+1} + c_n E_t (x_{t+1} - x_{t+1}^*) + w_{n,t}, \quad (7)$$

⁴Of these six, four are very recent members of the MPC and, as of June 2006, had participated in only twelve (D. Walton), eight (J. Gieve), four (D. Blanchflower), and one (T. Besley) meetings. H. Davies was a member in the original MPC committee but participated in only two meetings on June 6 and July 10, 1997. A. Budd was a member from December 1997 to May 1999, but his voting record consists of only fourteen observations.

⁵Prior to November 1998, minutes reported whether dissenting members favored tighter or looser policy, but not the interest rates they voted for. This means that a total of twenty-three dissenting votes from the period June 1997 to October 1998 could not be used for the estimation of the individual reaction functions of five members. They are M. King (three votes), A. Budd (four votes), W. Buiter (eight votes), C. Goodhart (three votes), and D. Julius (five votes). Except for W. Buiter, the number of missing observations is relatively small compared with the total number of individual observations.

where

$$\begin{aligned} a_n &= - \left(\frac{\lambda}{\phi(\alpha_n + \lambda^2)} \right) \pi_t^*, \\ b_n &= 1 + \frac{\lambda\beta}{\phi(\alpha_n + \lambda^2)}, \\ c_n &= \frac{1}{\phi}, \\ w_{n,t} &= \left(\frac{\lambda}{\phi(\alpha_n + \lambda^2)} \right) \eta_t + \left(\frac{\alpha_n}{\phi(\alpha_n + \lambda^2)} \right) \nu_{n,t} + \left(\frac{1}{\phi} \right) \varepsilon_t. \end{aligned}$$

Since the Bank of England follows an inflation-targeting policy, the definition of the intercept, a , explicitly assumes that committee members share the same inflation target. However, notice that because a also depends on the preference parameter, α_n , the intercept in the reaction function is member specific. For the same reason, the inflation coefficient, b , and the variance of the disturbance term are member specific as well. In principle, the output coefficient, c , should be the same for all members, but the approach that we will follow here is to estimate an unrestricted version of the reaction function that allows c to vary across members and then statistically test whether c is constant or not. Following most of the literature, lagged realizations of the interest rate are added to the right-hand side of (7) in order to capture interest-rate smoothing and the serial correlation of the error term.⁶

The assumption of rational expectations on the part of the public means that

$$\pi_{t+1} = E_t \pi_{t+1} + \zeta_{t+1}, \quad (8)$$

$$x_{t+1} - x_{t+1}^* = E_t (x_{t+1} - x_{t+1}^*) + \xi_{t+1}, \quad (9)$$

⁶An alternative approach would be to explicitly incorporate an interest-rate-smoothing motive in the payoff function (1), as in Woodford (2003). However, this strategy makes the model more difficult to solve because the committee member's problem becomes dynamic and its optimal solution requires him or her to forecast the preferred policies of other members.

where ζ_{t+1} and ξ_{t+1} are forecast errors uncorrelated with information available at time t . Then, substituting (8) and (9) into (7) delivers

$$i_{n,t} = a_n + b_n \pi_{t+1} + c_n (x_{t+1} - x_{t+1}^*) + u_{n,t}, \quad (10)$$

where the reduced-form disturbance $u_{n,t} = w_{n,t} - b_n \zeta_{t+1} - c_n \xi_{t+1}$ now includes the forecast errors. Since π_{t+1} and $x_{t+1} - x_{t+1}^*$ are respectively correlated with ζ_{t+1} and ξ_{t+1} , and hence with $u_{n,t}$, ordinary least squares (OLS) would yield biased and inconsistent parameter estimates. For this reason, we use instead an instrumental-variables (IV) procedure to estimate the reaction function of each member and of the committee as a whole. In particular, we use efficient generalized method of moments (GMM). As it is well known, GMM is statistically more efficient than the standard IV and two-stage least squares (2SLS) because it exploits the heteroskedasticity and serial correlation of the residuals in an initial 2SLS regression to construct the optimal weighting matrix.⁷ We construct the weighting matrix using the Newey-West estimator with a Barlett kernel.

Instruments are the inflation target; one lag of the nominal interest rate, RPI, CPI, and RPIX inflation rates; the rate of money growth; the unemployment rate; and the IIP, the latter two detrended in the manner described below. (We use a one-period lag because publication delays mean that contemporaneous observations may not actually form part of the public's information set at time t .) These instruments are both valid and relevant. The instruments are valid because, as part of the public's information set at time t , they are uncorrelated with the forecasts errors ζ_{t+1} and ξ_{t+1} , and are uncorrelated with the structural shocks in $w_{n,t}$ because the latter are exogenous by assumption. This means that the instruments are uncorrelated with the error term, $u_{n,t}$. The instruments are relevant because

⁷Since the dependent variable is discrete, an alternative strategy would be to use a logit or probit regression. We abstained from following that approach because addressing the endogeneity of expectations is much more difficult in that setup.

they are strongly correlated with the variables they are meant to instrument for.⁸

In the reaction function, inflation is measured by the annual percentage change in the RPIX,⁹ and the output gap is measured by the deviation of the unemployment rate from a Hodrick-Prescott trend computed recursively using a window of 120 observations.¹⁰ The change in the inflation target is modeled by allowing a break in the intercept term after December 2003, but in order to preserve degrees of freedom, we restrict the size of the break to be equal to the actual target change.

3.3 Results

Table 1 reports benchmark results based on a forecast horizon of twelve months. This horizon is motivated by the observation that monetary policy affects output and inflation with a long lag.¹¹ Tables 2 and 3 report results using other forecast horizons—namely, nine months (table 2) and six months (table 3). These two tables

⁸The R^2 s for the first-stage regressions vary from 0.66 to 0.92.

⁹Prior to December 10, 2003, the inflation target applied to the twelve-month change in the RPIX. Thereafter, it applies to the change in the CPI. Results using the RPIX for the sample until December 2003 and the CPI after that are similar to the ones reported below and support the same conclusions, but they unfortunately require a specific assumption as to when the public switched from forecasting RPIX to forecasting CPI inflation.

¹⁰We also considered using other output measures, but few are available at the monthly frequency. The National Institute for Economic and Social Research produces a monthly GDP index dating back to April 1984, but the series before and after April 1995 are not comparable. The Bank of England compiles an index of capacity utilization based on survey data, but it is not available for the complete sample. In preliminary empirical work, we used the detrended IIP as output-gap measure but found the reaction-function coefficients to be statistically equal to zero in almost all cases and frequently of a sign opposite to that predicted by theory. This result is perhaps not surprising, because manufacturing accounts (as of 2003) for only 16 percent of national output and 13 percent of employment in the United Kingdom and, consequently, is a relatively poor proxy for aggregate output.

¹¹Batini and Haldane (1999) evaluate the performance of forecast-based monetary policy rules with different forecast horizons and find that the optimal horizon is around four quarters.

Table 1. Benchmark Results (Twelve-Month Horizon)

A. Reaction-Function Coefficients							
Member	Intercept		Inflation		Unemployment		<i>J</i> -test (<i>p</i> -value)
	Estimate	s.e.	Estimate	s.e.	Estimate	s.e.	
George	−0.493	0.338	0.277	0.280	−0.983	0.667	0.451
King	−1.132 [†]	0.686	0.768	0.505	−2.334 [†]	1.410	0.633
Lomax	−0.233*	0.063	0.199*	0.059	−0.661*	0.215	0.738
Large	−0.294*	0.117	0.305*	0.137	−0.407	0.472	0.433
Tucker	−0.151	0.099	0.049	0.066	−0.129	0.300	0.382
Bean	−0.263	0.201	0.096	0.078	−0.607	0.553	0.150
Barker	−0.442*	0.213	0.182*	0.092	−0.692	0.600	0.429
Nickell	−0.627*	0.260	0.225 [†]	0.131	−1.124 [†]	0.603	0.517
Allsopp	−0.657*	0.147	0.306*	0.095	−0.641	0.436	0.477
Bell	−0.424*	0.132	0.138 [†]	0.074	−0.840*	0.281	0.567
Lambert	−0.251*	0.083	0.179*	0.060	−0.642*	0.221	0.627
Buiter	−1.114*	0.389	0.996*	0.357	−2.592*	1.095	0.152
Goodhart	−0.055	0.298	0.166	0.250	0.003	0.651	0.292
Vickers	−0.971*	0.236	1.088*	0.319	−2.246*	0.689	0.478
Julius	−1.281*	0.335	1.278*	0.403	−2.381*	0.531	0.546
Wadhvani	−0.262	0.236	0.036	0.129	−1.391*	0.308	0.524
Plenderleith	−1.394 [†]	0.788	1.185	0.755	−2.113	1.368	0.932
Clementi	−0.656 [†]	0.368	0.448	0.308	−0.871	0.817	0.524
Committee	−0.927	0.628	0.630	0.467	−1.808	1.308	0.576
B. Individual and Committee Coefficients Are Equal (<i>p</i> -value)							
	Intercept		Inflation		Unemployment		
All Members	0.002		0.003		0.053		
External	0.020		0.034		0.077		
Internal	0.050		0.170		0.284		
Academic	0.262		0.055		0.501		
Private	0.117		0.260		0.088		
Bank	0.419		0.248		0.385		
C. Median Coefficients							
	Intercept		Inflation		Unemployment		
All Members	−0.468		0.251		−0.856		
External	−0.442		0.182		−0.840		
Internal	−0.493		0.305		−0.871		
Academic	−0.814		0.537		−1.685		
Private	−0.424		0.182		−0.840		
Bank	−0.233		0.199		−0.661		

(continued)

Table 1. (Continued)

D. Committee Coefficients Are Equal to the Median (<i>p</i> -value)			
	Intercept	Inflation	Unemployment
	0.464	0.418	0.467
Note: Internal members are George, King, Lomax, Large, Tucker, Bean, Vickers, Plenderleith, and Clementi. The remaining members are external. King, Bean, Nickell, Allsopp, Buiter, and Vickers have academic background. Large, Barker, Bell, Lambert, Julius, and Clementi come from the private sector. The remaining members have been Bank of England staff or government officials. See Spencer (2006) for additional details. The superscripts * and † denote the rejection of the null hypothesis that true coefficient is zero at the 5 percent and 10 percent significance levels.			

Table 2. Robustness to Forecast Horizon
(Nine-Month Horizon)

A. Reaction-Function Coefficients							
Member	Intercept		Inflation		Unemployment		<i>J</i> -test (<i>p</i> -value)
	Estimate	s.e.	Estimate	s.e.	Estimate	s.e.	
George	−0.182	0.127	0.081	0.116	−0.621*	0.279	0.357
King	−0.348*	0.156	0.237†	0.130	−0.888*	0.334	0.347
Lomax	−0.096	0.075	0.137	0.111	−0.301*	0.127	0.307
Large	−0.067	0.109	−0.053	0.098	−0.104	0.201	0.381
Tucker	−0.017	0.100	−0.075	0.064	−0.078	0.200	0.573
Bean	−0.256*	0.100	0.181*	0.072	−0.685*	0.272	0.649
Barker	−0.289*	0.118	0.167	0.124	−0.555†	0.324	0.497
Nickell	−0.336*	0.116	0.175†	0.103	−0.785*	0.286	0.476
Allsopp	−0.645*	0.093	0.410*	0.083	0.317	0.268	0.378
Bell	−0.289*	0.129	0.067	0.095	−0.716*	0.334	0.235
Lambert	−0.081	0.089	−0.043	0.149	−0.403*	0.129	0.338
Buiter	−0.365*	0.178	0.128	0.207	−2.076*	0.729	0.420
Goodhart	−0.216	0.136	0.224*	0.110	−0.712	0.445	0.266
Vickers	−0.365	0.242	0.386	0.287	−1.288*	0.410	0.334
Julius	−0.390*	0.173	0.312	0.248	−1.176*	0.400	0.334
Wadhwani	−0.245	0.185	0.169	0.138	−1.140*	0.203	0.186
Plenderleith	−0.265*	0.127	0.229	0.165	−0.717*	0.267	0.360
Clementi	−0.346†	0.192	0.248	0.204	−0.974†	0.527	0.517
Committee	−0.445*	0.185	0.320*	0.148	−1.000*	0.382	0.594

(continued)

Table 2. (Continued)

B. Individual and Committee Coefficients Are Equal (<i>p</i> -value)			
	Intercept	Inflation	Unemployment
All Members	0.004	0.031	<0.001
External	0.013	0.057	0.002
Internal	0.306	0.812	0.048
Academic	0.150	0.007	0.007
Private	0.342	0.909	0.012
Bank	0.308	0.435	0.159
C. Median Coefficients			
	Intercept	Inflation	Unemployment
All Members	−0.277	0.172	−0.714
External	−0.289	0.169	−0.716
Internal	−0.256	0.181	−0.685
Academic	−0.356	0.209	−0.836
Private	−0.289	0.167	−0.716
Bank	−0.182	0.137	−0.621
D. Committee Coefficients Are Equal to the Median (<i>p</i> -value)			
	Intercept	Inflation	Unemployment
	0.363	0.317	0.454
Note: Internal members are George, King, Lomax, Large, Tucker, Bean, Vickers, Plenderleith, and Clementi. The remaining members are external. King, Bean, Nickell, Allsopp, Buiter, and Vickers have academic background. Large, Barker, Bell, Lambert, Julius, and Clementi come from the private sector. The remaining members have been Bank of England staff or government officials. See Spencer (2006) for additional details. The superscripts * and † denote the rejection of the null hypothesis that the true coefficient is zero at the 5 percent and 10 percent significance levels.			

serve to assess the robustness of the results to different forecast horizons.¹²

Panel A in table 1 reports estimates of individual and committee reaction functions. These results indicate that the voting behavior

¹²The working paper version of this article (Riboni and Ruge-Murcia 2007) reports results using a different measure of inflation—namely, RPI inflation. Those results are very similar to the ones reported here and indicate that our conclusions are also robust to the index used to compute inflation.

**Table 3. Robustness to Forecast Horizon
(Six-Month Horizon)**

A. Reaction-Function Coefficients							
Member	Intercept		Inflation		Unemployment		<i>J</i> -test (<i>p</i> -value)
	Estimate	s.e.	Estimate	s.e.	Estimate	s.e.	
George	−0.171 [†]	0.099	0.119	0.098	−0.244	0.203	0.195
King	−0.125	0.082	0.101	0.076	−0.396*	0.176	0.168
Lomax	−0.001	0.050	0.013	0.041	−0.168 [†]	0.100	0.276
Large	0.118	0.082	−0.207*	0.089	−0.076	0.172	0.560
Tucker	0.052	0.087	−0.127 [†]	0.077	−0.069	0.169	0.547
Bean	−0.079	0.066	0.123*	0.050	−0.163	0.144	0.577
Barker	−0.078	0.067	0.039	0.053	−0.143	0.146	0.347
Nickell	−0.140	0.089	0.037	0.069	−0.372*	0.166	0.576
Allsopp	−0.339 [†]	0.195	0.110	0.147	0.008	0.357	0.122
Bell	−0.188	0.165	0.110	0.139	−0.386	0.319	0.162
Lambert	−0.041	0.070	0.026	0.079	−0.169	0.129	0.433
Buiter	0.082	0.182	0.549*	0.275	−1.735*	0.500	0.325
Goodhart	−0.055	0.104	0.276	0.197	−0.701*	0.246	0.342
Vickers	−0.366	0.304	0.581	0.374	−1.683*	0.567	0.328
Julius	−0.341 [†]	0.205	0.597*	0.166	−2.197*	0.443	0.480
Wadhvani	0.233	0.142	0.186	0.177	−2.026*	0.350	0.715
Plenderleith	−0.140	0.099	0.304 [†]	0.185	−0.670*	0.315	0.132
Clementi	−0.205 [†]	0.122	0.283 [†]	0.172	−0.921*	0.388	0.165
Committee	−0.201*	0.084	0.164*	0.072	−0.377*	0.158	0.220
B. Individual and Committee Coefficients Are Equal (<i>p</i> -value)							
	Intercept		Inflation		Unemployment		
All Members	0.100		<0.001		<0.001		
External	0.199		<0.001		<0.001		
Internal	0.073		0.324		0.114		
Academic	0.618		<0.001		0.085		
Private	0.026		0.040		<0.001		
Bank	0.165		0.182		0.382		
C. Median Coefficients							
	Intercept		Inflation		Unemployment		
All Members	−0.102		0.114		−0.379		
External	−0.078		0.110		−0.386		
Internal	−0.125		0.119		−0.244		
Academic	−0.132		0.116		−0.384		
Private	−0.078		0.110		−0.386		
Bank	−0.055		0.119		−0.244		

(continued)

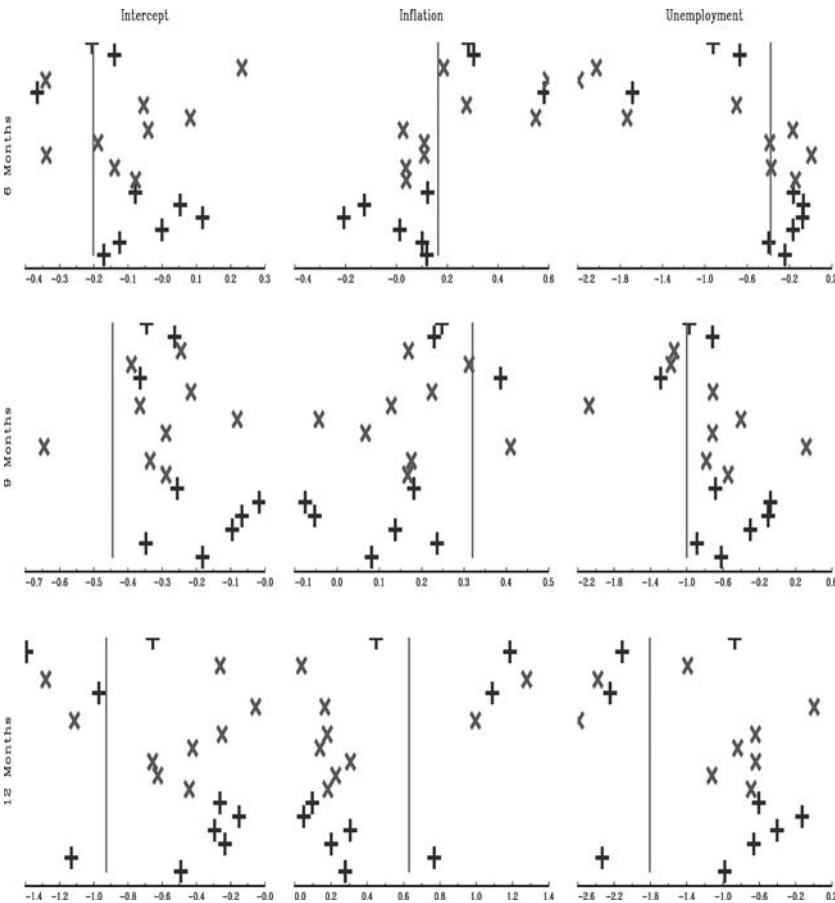
Table 3. (Continued)

D. Committee Coefficients Are Equal to the Median (<i>p</i> -value)			
	Intercept	Inflation	Unemployment
	0.236	0.490	0.994
Note: Internal members are George, King, Lomax, Large, Tucker, Bean, Vickers, Plenderleith, and Clementi. The remaining members are external. King, Bean, Nickell, Allsopp, Buiter, and Vickers have academic background. Large, Barker, Bell, Lambert, Julius, and Clementi come from the private sector. The remaining members have been Bank of England staff or government officials. See Spencer (2006) for additional details. The superscripts * and † denote the rejection of the null hypothesis that the true coefficient is zero at the 5 percent and 10 percent significance levels.			

of individual MPC members is generally consistent with the theory. First, point estimates of the inflation (unemployment) coefficients are usually positive (negative) and statistically significant, meaning that members tend to vote to raise (lower) the repo rate when inflation (unemployment) twelve months ahead is expected to increase. In this and other tables, there are occasional exceptions to this general observation, but exceptions are usually characterized by numerically small and statistically insignificant coefficients. Second, the overidentifying restrictions of the model cannot be rejected for any committee member nor for the committee as a whole. To see this, note that the *p*-values of *J*-tests reported in the last column of panel A in table 1 are all above standard significance levels. Comparing these results with those in panel A in tables 2 and 3 indicates that these conclusions are robust to the forecast horizon. The reaction-function coefficients are also graphically reported in figure 1, where “pluses” are coefficients of internal members, “crosses” are coefficients of the external members, and vertical lines are the coefficients of the committee as a whole. Overall, results in panel A and figure 1 suggest that, in qualitative terms, MPC members are fairly homogeneous in their policy preferences.

However, these results also indicate that, in quantitative terms, the reaction coefficients vary substantially across committee members. In order to examine whether this heterogeneity is statistically significant, we perform Wald tests of the joint null hypothesis that

Figure 1. Reaction-Function Coefficients



the individual coefficients are equal to those of the committee and report results in panel B of table 1. For the complete sample of all committee members, the hypothesis can be rejected at the 5 percent significance level for the intercept and inflation coefficient but not for the unemployment coefficient. Although the latter result is marginal, it is in line with the implication of the model that members are more likely to disagree on their policy reaction to expected inflation than on their reaction to the expected output gap. Recall that in the model this was due to the fact that changes in expected

inflation require a monetary-policy-induced change in the output gap that leads to a trade-off between inflation and output stabilization, while changes in the expected output gap may be offset directly by adjusting the nominal interest rate. From panel B in tables 2 and 3, it is clear that the rejection of the null hypothesis for the inflation coefficients is robust to the forecast horizon, but results concerning the unemployment coefficient are not. Overall, these statistical results suggest that individual policy responses within the MPC are heterogeneous.

Figure 2 plots the relation between the reaction-function coefficients for all members for all possible pairwise comparisons. In all cases there is a positive and statistically significant relation across forecast horizons.¹³ This means that, in general, individuals who favor a strong response to expected inflation or unemployment do so regardless of the forecast horizons. Thus, the heterogeneity in individual policy responses within the MPC appears to be systematic.

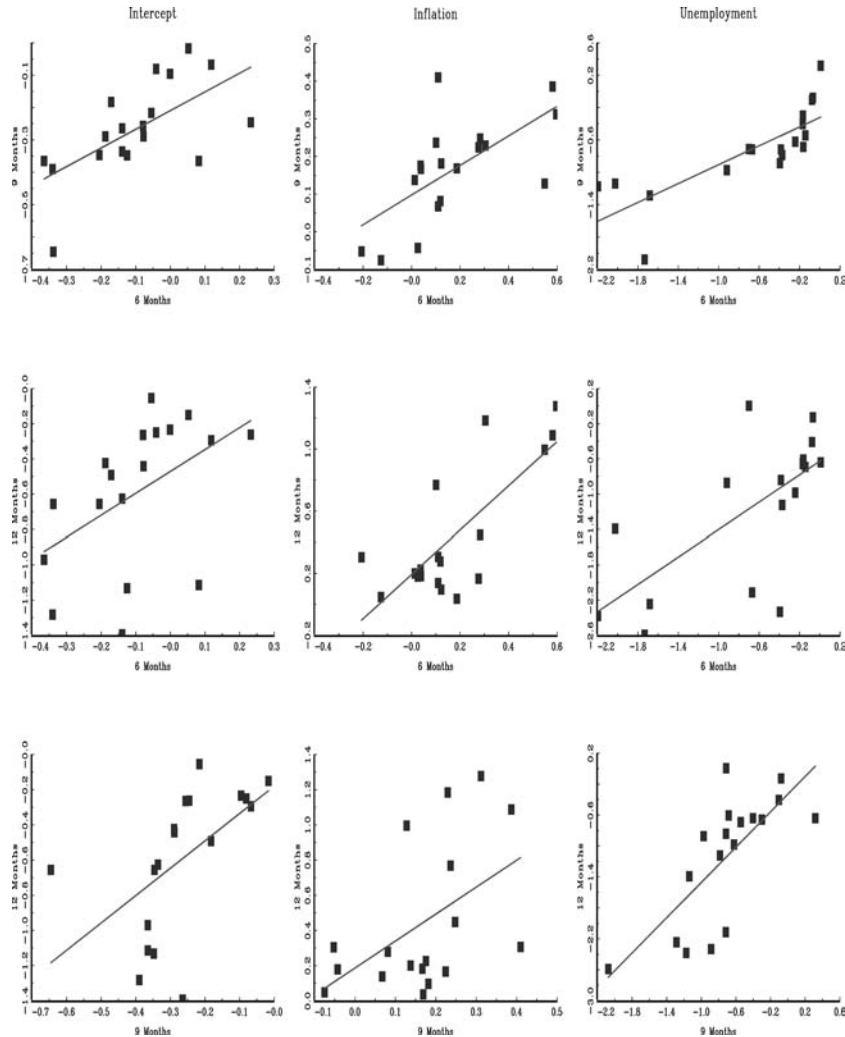
There are at least two dimensions along which MPC members differ systematically. First, the nature of their appointment (whether internal or external to the Bank of England) and, second, their career background (whether academic, private sector, or Bank of England).¹⁴ Thus, we also perform the Wald test described above for subsamples of committee members. These results are reported in panel B of table 1 and show that the hypothesis that the individual inflation and unemployment coefficients are equal to those of the committee cannot be rejected for internal and Bank members, but is frequently rejected for external, academic, and private-sector members.¹⁵ These results are not independent, because most academic and private-sector members in the MPC are also external appointees, while Bank members are internal members. Results are generally robust to the forecast horizon (see panel B in tables 2 and 3) and, as a whole, statistically confirm the idea that one important source of heterogeneity within the MPC is the nature of the

¹³To save space, the slope parameters are not reported, but their *t*-statistics range from 2.21 to 5.27.

¹⁴For the classification of MPC members in terms of career background, we mostly rely on the information contained in Spencer (2006).

¹⁵Note, however, that in some cases results are marginal in that the hypothesis would be rejected at the 5 percent but not the 10 percent significance level.

Figure 2. Relation Between Coefficients at Different Horizons



individual membership.¹⁶ In order to further explore this finding, panel C in tables 1–3 reports the median coefficients for all members and subsamples. Estimates indicate that the median external

¹⁶On this point, see also Gerlach-Kristen (2003) and Spencer (2006).

member tends to react more strongly to unemployment than the median internal member, but there is no clear pattern regarding their reaction to inflation. The same result is observed when comparing academic and private-sector members with Bank members. In general, reaction-function coefficients increase with the forecast horizon, especially for academics.

Finally, panel D in table 1 reports the results of t -tests that the committee reaction-function coefficients are equal to the median of the individual coefficients. Since the p -values are above standard significance levels, the hypothesis cannot be rejected for the intercept, inflation, or unemployment. From panel D in tables 2 and 3, it is clear that this result is robust to the forecast horizon. Since the committee and median coefficients are (statistically) the same but the coefficients of committee and external members are (statistically) different, it follows that the median voter in the MPC is likely be an internal member.

4. Conclusions

This paper exploits the voting records of the Monetary Policy Committee of the Bank of England to study preference heterogeneity among committee members. The analysis is motivated by the empirical observations that disagreement is frequent within the MPC and by summary statistics reported in earlier literature that suggest differences in voting patterns. This research shows that, in qualitative terms, MPC members are fairly homogeneous in their policy preferences. That is, individual members tend to vote for an interest rate raise (decrease) when inflation (unemployment) is expected to increase, as one would expect. However, there are systematic quantitative differences in their policy reaction functions that appear related to the nature of the membership and previous career background. These systematic differences may be important in the decision-making process of the MPC and affect the interest rate level selected by the committee.

References

- Batini, N., and A. Haldane. 1999. "Monetary Policy Rules and Inflation Forecasts." *Quarterly Bulletin* (Bank of England) 39: 60–67.

- Belden, S. 1989. "Policy Preferences of FOMC Members as Revealed by Dissenting Votes." *Journal of Money, Credit, and Banking* 21 (4): 432–41.
- Bernanke, B. S., and F. S. Mishkin. 1997. "Inflation Targeting: A New Framework for Monetary Policy?" *Journal of Economic Perspectives* 11 (2): 97–116.
- Chappell, H. W., R. R. McGregor, and T. Vermilyea. 2005. *Committee Decisions on Monetary Policy*. Cambridge, MA: MIT Press.
- Fry, M., D. Julius, L. Mahadeva, S. Roger, and G. Sterne. 2000. "Key Issues in the Choice of Monetary Policy Framework." In *Monetary Frameworks in a Global Context*, ed. L. Mahadeva and G. Sterne. London: Routledge.
- Gerlach-Kristen, P. 2003. "Insiders and Outsiders at the Bank of England." *Central Banking* 14: 96–102.
- Gildea, J. A. 1990. "Explaining FOMC Members' Votes." In *The Political Economy of American Monetary Policy*, ed. T. Mayer, 211–28. Cambridge: Cambridge University Press.
- Groseclose, T., and J. Milyo. 2006. "A Rational-Choice, Formal-Theoretic Argument Against the Existence of Sophisticated Voting in Legislatures." Mimeo, University of California, Los Angeles.
- Havrilesky, T. M., and J. A. Gildea. 1991. "The Policy Preferences of FOMC Members as Revealed by Dissenting Votes: Comment." *Journal of Money, Credit, and Banking* 23 (1): 130–38.
- Havrilesky, T. M., and R. L. Schweitzer. 1990. "A Theory of FOMC Dissent Voting with Evidence from the Time Series." In *The Political Economy of American Monetary Policy*, ed. T. Mayer, 197–210. Cambridge: Cambridge University Press.
- McCallum, B. T., and E. Nelson. 1999. "An Optimizing IS-LM Specification for Monetary Policy and Business Cycle Analysis." *Journal of Money, Credit, and Banking* 31 (3): 296–316.
- Meade, E. E., and D. N. Sheets. 2005. "Regional Influences on FOMC Voting Patterns." *Journal of Money, Credit, and Banking* 37 (4): 661–77.
- Riboni, A., and F. J. Ruge-Murcia. 2007. "Preference Heterogeneity in Monetary Policy Committees." DNB Working Paper No. 157.
- Seidmann, D. 2006. "A Theory of Voting Patterns and Performance in Private and Public Committees." Mimeo, Nottingham University.

- Spencer, C. 2006. "The Dissent Voting Behaviour of Bank of England MPC Members." Discussion Paper 03/06, University of Surrey.
- Tootell, G. M. B. 1991. "Are District Presidents More Conservative than Board Governors?" *New England Economic Review* (Federal Reserve Bank of Boston), 3–12.
- . 1999. "Whose Monetary Policy Is It Anyway?" *Journal of Monetary Economics* 43 (1): 217–35.
- Woodford, M. 2003. "Optimal Interest-Rate Smoothing." *Review of Economic Studies* 70 (4): 861–86.