# Bank Crisis Resolution and Foreign-Owned Banks

**ROBERT A. EISENBEIS AND GEORGE G. KAUFMAN**

*Eisenbeis is executive vice president and director of research at the Federal Reserve Bank of Atlanta. Kaufman is the John Smith Professor of Banking at Loyola University Chicago and a consultant to the Federal Reserve Bank of Chicago.* *This article was initially prepared for presentation at the conference "Banking Crisis Resolution: Theory and Practice," held June 16–17, 2005, at Norges Bank in Oslo, Norway. The authors are indebted to the participants at the conference for helpful suggestions and comments.*

In practice, resolving insolvent banks efficiently, particularly when they are very large, presents a challenge that has been poorly met in almost all countries in recent years. Insolvent banks have generally been resolved only at a high cost to the country in which the bank is located. The costs are of two types. The first is transfer costs arising from the use of taxpayers' funds to reduce or eliminate losses to some or all claimants of insolvent banks—for example, depositors, other creditors, and, on occasion, shareholders. These claimants usually bear the cost of bankruptcies in most nonbank failures. The second type is real costs associated with the misallocation of resources from the often prolonged operation of insolvent banks; this misallocation reduces a country's aggregate income below potential.[1] Both types of costs are magnified when reluctant bank regulators fail to move swiftly to resolve institutions when they first become insolvent.

Ironically, such regulatory forbearance is motivated in large measure by a fear of high societal costs from officially recognizing and resolving insolvencies sooner. Depositors may perceive themselves as losing when legal closure and official recognition transform unrecognized implicit losses into explicit losses. They may also experience liquidity problems if they lose immediate and full access to the funds when the insolvent bank is legally closed. Borrowers may experience liquidity problems if they are unable to access their credit lines. Finally, the payment system may be interrupted. It is widely feared that the effects of these problems may spill over beyond the banks and have adverse consequences for the economy as a whole.

This article examines the efficiency of the resolution process for foreign-owned banks and considers whether host country regulators face additional potential problems when these institutions experience financial difficulties. As Lastra (2004) has noted, banking is becoming increasingly international, but prudential regulation and insolvency resolution have remained national. This article concludes that the problems resulting from this international/national dichotomy in foreign-owned bank resolution

are real and that their severity depends both on the size of a foreign bank's host country operations and whether the bank operates through branches or subsidiaries. Obtaining useful, timely, and accurate financial data may be difficult for a host country regulator, especially when the foreign bank is operating through branches. But even if accurate financial information were available, host country regulators' ability to move promptly to take corrective action may be hampered by the need to share prudential supervision, regulation, and enforcement with home country regulators. These overlapping responsibilities introduce the possibility of delays not only in sharing relevant information but also in legally closing insolvent institutions.[2]

*Large-scale foreign bank presence in a country could increase the potential confusion both before and after one or more foreign institutions become insolvent.*

Additional problems arise because foreign (home) countries often provide deposit insurance for branches of domestically headquartered banks in all countries and, in some cases, such as the European Union, are required to do so.[3] Because both deposit insurance and resolution policies may differ from home country to home country, the insolvency and receivership ("pulling-the-plug") criteria applied to foreign branches in a host country may differ depending upon where the institution is headquartered. Potential resolution problems may be compounded when a host country regulator deals with institutions from many different home countries operating under different deposit insurance and failure resolution schemes. These variations can result in differences in both the treatment of creditors and in supervisory and resolution policies. For example, the payment of claims is governed by rules that specify when insured depositors at insolvent institutions are paid, the office at which the claims are booked, when and how much uninsured depositors and other creditors are paid, and whether the insurance agency contacts the eligible claimants or the claimants have to file claims individually.

Differences in supervisory and resolution policies also depend upon how strict prudential sanctions are and how the closure rule is structured and enforced. The latter affects how long insolvent banks are kept open and operating before legal closure takes place. As the number of branches from different home countries operating in a given host country increases, so does the possible number of different deposit insurance and resolution systems that will have to be reconciled should institutions fail. Therefore, large-scale foreign bank presence in a country could increase the potential confusion both before and after one or more institutions become insolvent.

Finally, using home country domestic taxpayer funds to make deposit insurance payments to depositors in other countries could, if nothing else, prove to be politically difficult. Indeed, on occasion, the problems that arise in resolving foreign-owned banks, particularly those involving branches, may be so severe that that cross border banking in this form may be too costly to permit as a matter of public policy despite the well-recognized benefits in terms of intensified competition and improved management that may be associated with the entry of foreign banks.[4]

We focus particularly on the less well recognized "dark side" of direct cross border investment in banks that may be encountered when economic times turn bad. We also discuss the costs and benefits of entry in foreign countries via branches versus subsidiary banks. To minimize the costs of bank failures, a four-point program is proposed for resolving insolvent institutions efficiently. The article describes how the program is designed to work in the United States. But the presence of foreign-owned banks in a country may increase that country's difficulties in satisfying the program's points. The article concludes by proposing a number of policies that may mitigate

these problems. Considering the issues involved before a crisis occurs is critical because cross border banking in the form of foreign direct investment in banks is growing rapidly as advances in computer and telecommunications technology reduce the costs of operating both additional offices and across greater distances. Thus, improvements in countries' ability to resolve insolvent foreign banks more efficiently and to lower the costs of bank failures would increase the likelihood that the bright side of cross border banking will outweigh the dark side in both the long and the short run.

## Efficient Bank Insolvency Resolutions in the United States

A bank becomes economically insolvent when the value of its assets declines below the value of its deposits and other debt funding so that the market value of its capital turns negative. At this point, the bank cannot pay out all its debts, including deposits in full and on time, and its depositors and other creditors share in the losses according to their legal priority.

These claimants may experience both credit and liquidity losses in the resolution process. Credit losses may occur when the recovery value of the bank as a whole or in part falls short of the par value of its deposits or other debt on the respective due dates. Liquidity losses may occur for two reasons: First, depositors may not have immediate (next business day or so) and full access to the par value of their de jure insured claims or to the estimated recovery value of their de jure uninsured claims. Second, qualified borrowers may not be able to use their existing credit lines immediately. Insolvent banks are resolved efficiently when the sum of aggregate credit losses and aggregate liquidity losses, or total losses, is at or close to zero (see Kaufman 2004b).

An insolvent bank may be resolved efficiently at the lowest cost to both the bank claimants and the macroeconomy if the process employed by bank regulators in the country in which the bank is chartered or licensed can satisfy the following four rules or principles.[5] Each principle focuses importantly on the term "prompt":

1. prompt legal closure when the bank's capital declines to some prespecified and well-publicized positive minimum greater than zero (legal closure rule);
2. prompt estimate of the recovery values and assignment of credit losses ("haircuts") to de jure uninsured bank claimants when equity is de facto negative;

---

1. Caprio and Klingebiel (1999) provide a list of banking crises and their costs. Hoggarth, Reis, and Saporta (2002) estimate that that the real costs of the forty-seven crises they studied amounted to 1 to 20 percent of annual gross domestic product (GDP). Other estimates have been provided by Barth, Caprio, and Levine (2001) and Kaufman (2000), who provide estimates of both real and transfer costs.
2. Many of these issues have been recognized in a recent paper by the European Financial Services Round Table (2005).
3. In some instances, host countries may also provide insurance for branches of foreign banks operating in the host country or may top off insurance if coverage offered in the host country is superior to that offered by the home country fund.
4. The advantages are discussed in Barth, Caprio, and Levine (forthcoming) and Berglöf et al. (2005). In New Zealand, the central bank requires an arm's-length relationship between bank subsidiaries in New Zealand and their parent companies (see Bollard 2005 and Reserve Bank of New Zealand 2004a, 2004b). These issues are also discussed in Borchgrevink and Moe (2004).
5. See Kaufman (2004a). Similar plans have been proposed by Mayes (2005) and the Reserve Bank of New Zealand (Harrison 2005), among others.

3. prompt reopening (for example, next workday), particularly of larger banks, with full depositor access to their accounts on their due dates at their insured or estimated recovery values and full borrower access to their pre-established credit lines; and
4. prompt reprivatization in whole or in part with adequate capital.

The next section reviews how each principle is or could be satisfied in bank insolvencies in the United States. We argue that the current U.S. system, while not without flaws and not focused on foreign-owned banks, may serve as a useful model for other countries in designing their insolvency resolution policies. The U.S. system was developed largely in response to the widespread and costly bank and thrift institution insolvencies of the 1980s.

**Prompt legal closure.** The Federal Deposit Insurance Corporation Improvement Act (FDICIA) of 1991 introduced a "bright line" bank closure rule that is triggered when the ratio of book-value tangible equity capital to total on-balance-sheet assets declines to a minimum of 2 percent.[6] If this ratio is not corrected within ninety days, the bank must be declared legally insolvent, closed by the appropriate federal or state regulator, and placed in receivership or conservatorship.[7] Its charter is revoked, shareholder controlling interests are terminated, and senior management is typically changed. If the institution can be successfully resolved before its market value capital declines below zero, losses are confined to shareholders. Depositors and other creditors are fully protected and kept whole, and deposit insurance is effectively redundant. Thus, any adverse spillover effects, which occur primarily when capital turns negative and losses are imposed on counterparties, are minimized.

Because the closure rule is specified as book-value rather than market-value capital, there is no guarantee that the institution will be resolved before its economic capital is depleted or that creditors will be fully protected against losses. As a bank approaches insolvency, book values tend to increasingly overstate market values for assets. Using a book-value closure rule could result in de facto forbearance, so a market value–based rule would be preferable.[8] Nevertheless, specifying a closure rule based on a book-value capital ratio that is greater than zero provides some protection against losses due to the deviation of book from market value and to errors in measuring asset values.

Legal closure that proceeds according to a well-specified, publicized, and credibly enforced closure rule has several desirable attributes. It provides no surprises: All players know the rules in advance and base their actions accordingly. It treats all depositors and other creditors in the same priority class more fairly. Because banks tend to have a larger percentage of demand deposits and other short-term deposits and debt than other firms do, bad news, impending insolvency, or uncertainty about how creditors would be treated in the event of insolvency typically increases the incentives of those who can withdraw their funds to do so while assets are still available to satisfy their claims. Uncertainty thus raises the probability of a run, with the initial runners receiving full payment and those unable or unwilling to run receiving less. However, the presence of an enforced closure rule that would close a bank while its capital is still positive would reassure claimants and greatly reduce their incentive to run. All debt claimants, regardless of the date of maturity of their claims, would know that their claims would be kept whole. Reducing the incentive for runs also increases the time available for regulators to act to deter insolvency or bring about an efficient resolution if the closure trigger is breached.

Banks become insolvent in the United States and need to be legally closed when regulators are unsuccessful in implementing another provision of FDICIA—prompt

corrective action (PCA).[9] PCA is designed to provide incentives for financially troubled banks to turn around before insolvency. PCA established a series of five capital tranches ranging from "well-capitalized" to "critically under-capitalized." To discourage insolvency, bank regulators apply progressively harsher and more mandatory sanctions on weak financial institutions as their net worth declines through these tranches (see Table 1). The sanctions are similar to those the market imposes on firms in nonregulated industries. Sanctions include change in senior management; reductions in dividends; restrictions on growth and acquisitions; adoption of capital restoration plans; and, if the bank is a subsidiary of a financial holding company, loss of its parent's status as a financial holding company with the associated wider range of powers (PricewaterhouseCoopers 2003). The tranches effectively serve as speed bumps to slow a bank's deterioration and to force regulators to become more involved with the troubled bank well before insolvency occurs so that they may be ready to close the bank legally when necessary and not be caught by surprise and delayed. Thus, PCA effectively buys time for regulators to act efficiently.

*Reducing the incentive for bank runs increases the time available for regulators to act to deter insolvency or bring about an efficient resolution if the closure trigger is breached.*

PCA also grants regulators some discretion to accelerate the application of appropriate sanctions and actions as a bank's capital position deteriorates. This authority is in contrast to the supervisory actions employed prior to FDICIA, when intervention was less frequent or timely and discretion was often focused on ways to keep institutions in business rather than on resolving them after they had become economically insolvent. The pre-FDICIA policy tended to result in greater losses to both uninsured creditors and the FDIC.

While PCA has not prevented all bank failures, it has contributed significantly to turning troubled banks around before insolvency and reducing both the number and the aggregate cost of failures.[10] It is important to note, however, that PCA and a closure rule at positive capital are not intended to prevent all failures. As in other industries, inefficient or unlucky banks should be permitted to fail and inept management replaced. But, because the adverse externalities of bank insolvencies are widely perceived to be substantially greater than for other firms, such failures should occur only at low cost with minimal losses to creditors.

---

6. Banks and thrift institutions in the United States are not subject to the corporate bankruptcy code but to a special code in the Federal Deposit Insurance Act (FDIA). The bank act is considerably more administrative and less judicial, considerably more creditor friendly, and potentially faster in the declaration of insolvency ousting the shareholders and in-place senior management and in making payments to creditors. Bank and financial holding companies are, however, subject to the general corporate bankruptcy code (Bliss and Kaufman 2005).
7. Two ninety-day extensions are permitted.
8. While regulators in the United States may also declare a bank insolvent for a number of other reasons, such as unsafe and unsound banking, they must do so when the closure-rule capital ratio is breached. Wall and Eisenbeis (2002) and Kaufman (2004a) demonstrate that, on average, institutions in the United States have been legally closed long after the market value of equity became negative.
9. The exception is that institutions can fail because of fraud, which is by definition difficult to detect under most circumstances.
10. See OCC (2003) and Salmon et al. (2003). Kaufman (2004c) and Wall and Eisenbeis (2002) both suggest, however, that losses in individual cases have been significant.

Table 1
**Summary of Prompt Corrective Action Provisions of the Federal Deposit Insurance Corporation Improvement Act of 1991**

| Zone | Mandatory provisions | Discretionary provisions | Capital ratios (percent) Risk-based Total | Risk-based Tier 1 | Leverage Tier 1 |
|---|---|---|---|---|---|
| 1. Well-capitalized | | | >10 | >6 | >5 |
| 2. Adequately capitalized | 1. No brokered deposits, except with FDIC approval | | >8 | >4 | >4 |
| 3. Undercapitalized | 1. Suspend dividends and management fees<br>2. Require capital restoration plan<br>3. Restrict asset growth<br>4. Approval required for acquisitions, branching, and new activities<br>5. No brokered deposits | 1. Order recapitalization<br>2. Restrict interaffiliate transactions<br>3. Restrict deposit interest rates<br>4. Restrict certain other activities<br>5. Any other action that would better carry out prompt corrective action | <8 | <4 | <4 |
| 4. Significantly undercapitalized | 1. Same as for zone 3<br>2. Order recapitalization[1]<br>3. Restrict interaffiliate transactions[1]<br>4. Restrict deposit interest rates[1]<br>5. Pay of officers restricted | 1. Any zone 3 discretionary actions<br>2. Conservatorship or receivership if bank fails to submit or implement plan or recapitalize pursuant to order<br>3. Any other zone 5 provision if such action is necessary to carry out prompt corrective action | <6 | <3 | <3 |
| 5. Critically undercapitalized | 1. Same as for zone 4<br>2. Receiver/conservator within 90 days[1]<br>3. Receiver if still in zone 5 four quarters after becoming critically undercapitalized<br>4. Suspend payments on subordinated debt[1]<br>5. Restrict certain other activities | | | | <2 |

[1] Not required if primary supervisor determines action would not serve purpose of prompt corrective action or if certain other conditions are met.

Source: Board of Governors of the Federal Reserve System

**Prompt estimate and allocation of credit losses.** Because regulators should be scrutinizing a troubled bank under PCA well before the bank approaches the capital ratio closure trigger, they should in most instances be able to quickly estimate the recovery value of the institution as a whole or in part upon it legal closure. If the estimated recovery value falls short of the par value of the deposits and other debts, pro rata losses (haircuts) should be allocated to these claimants in their order of legal priority. In the United States, the FDIC has equal standing with depositors at domestic offices and higher standing than other depositors and creditors.[11] The FDIC stands in the place of the insured depositors at domestic offices and is obligated to cover insured deposits in their entirety. The FDIC also shares proportionally in any losses

with uninsured depositors at domestic offices beyond the losses charged first to other creditors and deposits at foreign offices. FDICIA requires the FDIC to share any losses with uninsured claimants and resolve the institution at least cost to the insurance fund. The only exception is when doing so is likely to "have serious adverse effects on economic conditions and financial stability." Requiring parties besides the FDIC to share in any losses is necessary to minimize moral hazard excessive risk-taking behavior by banks and to enhance market discipline by reinforcing the ex post at-risk nature of de jure at-risk claimants. This requirement should, in turn, reduce the number of bank failures.

**Prompt reopening of large banks.** Liquidity losses to depositors can occur when access to their deposit accounts is delayed or their accounts are frozen. These actions transform demand deposits involuntarily into longer-term time deposits or bonds. Liquidity losses also result when credit lines cannot be relied upon or drawn down to meet business needs. Loss of liquidity thus impairs the efficient operation of the payment system. When regulators close a bank legally, they often also effectively close it physically, at least partially, until funds are recovered from the sale of assets to start paying depositors on their claims. In many countries the lack of access to deposits and credit lines is more feared than actual losses to depositors and generates as great, if not greater, adverse externalities. The more likely depositors are to receive their funds promptly, the less likely they are to engage in runs.

Regulators often are unable or unwilling to avoid, at least briefly, closing banks physically when they close them legally—for example, because of insufficient information on depositors or recovery values. Thus, regulators are under considerable pressure to avoid legally closing banks promptly. By delaying legal closure, regulators temporarily avoid liquidity losses. But delay also postpones, at least temporarily, explicitly recognizing underlying implicit credit losses and provides additional time in which the bank may try to regain solvency and thereby avoid altogether the unpleasantness of legal closure. Evidence in many countries strongly suggests that, on average, such forbearance increases the costs in the long run over what they would have been had the insolvent institution been legally closed promptly. To reduce the incentive for regulators to forbear, FDICIA made prompt legal closure mandatory and, to increase the efficiency of the resolution, required that it be at least cost to the FDIC.

Liquidity losses may be minimized or eliminated entirely by legally reopening the insolvent bank the next business day. Reopening would provide insured depositors immediate access to the par value of their accounts, would give uninsured depositors and other general creditors access to the estimated recovery value of their accounts on due dates, and would allow borrowers to access their credit lines.[12] Thus, legal closure is separated from physical closure.

Potential payments to depositors and other debt claimants, either directly or through assumption of these claims by another bank, require an immediate sale of

---

11. Under the Depositor Preference Act of 1993, claims of general bank creditors, including sellers of fed funds, and deposits at foreign offices are subordinated to deposits at domestic offices. See Kaufman (1997) and Marino and Bennett (1999).
12. Fear of the adverse consequences of liquidity as well as credit losses have at times induced regulators not to give haircuts to uninsured debt claimants, particularly at large banks, after failing the institution by revoking its charter and ousting shareholders and management. This practice is often incorrectly termed "too-big-to-fail" (TBTF). The bank has already failed. This practice has proved highly costly and inefficient. Losses tend to increase, and ultimate resolution is typically only postponed, at which time losses are borne by the FDIC or the taxpayer.

the bank by the FDIC or access to a source of funds. The FDIC may also operate the bank temporarily through a newly chartered bridge bank that assumes most or all of the failed bank's assets and liabilities, generally at market values. The bridge bank is either capitalized with equity by the FDIC or its deposits are fully guaranteed by the FDIC during its operation until it is reprivatized.

The FDIC usually pays insured deposits at a failed bank at par the next business day either through a transfer of the deposits to another solvent bank, which assumes the liabilities with an offsetting financial payment or, less frequently, through a payout.[13] The FDIC can make such speedy payments because it has been monitoring the problem bank carefully under PCA and has access to the bank's records on eligible insured deposits. In contrast, uninsured depositors and other creditors generally are given receivership certificates and are paid in order of their legal priority as proceeds are received from the sale of the bank assets. Unless there is an active secondary market for these certificates, uninsured creditors receiving the certificates may suffer liquidity difficulties. To maximize efficiency, these depositors should share in any credit losses but not suffer liquidity losses. To help minimize liquidity losses, the FDIC has the authority to make advance payments to these claimants on the basis of estimated or historical average recovery amounts (Kaufman 2000; Kaufman and Seelig 2002). If payments are made at the time of legal closure, this procedure is essentially equivalent to not having physically closed the institution. Advance dividends also permit the estimated recovery value of uninsured deposits to be transferred to a newly chartered bridge bank with immediate access by depositors. In bridge banks, borrowers generally maintain access to their existing credit lines, further reducing any liquidity losses.

*While prompt corrective action has not prevented all bank failures, it has contributed significantly to turning troubled banks around before insolvency and reducing both the number and the aggregate cost of failures.*

Estimates of the recovery value of the funds advanced as dividends tend to be on the conservative side because the FDIC absorbs the loss if it overestimates the recovery amounts. If it underestimates the recovery amounts, it makes additional payments to the claimants later. The FDIC, in its capacity as receiver, can borrow the necessary funds to make advance dividend payments from its corporate capacity, which has access to the FDIC accumulated fund.

The FDIC used advance dividends briefly in a number of resolutions in the early 1980s and early 1990s, when it did not fully protect most or all uninsured debt claimants. But probably because most bank failures in the United States since the mid-1990s have involved small banks and have been largely the result of major fraud, the FDIC has not used advance dividends often.

Use of bridge banks and advance dividends to minimize liquidity losses, especially in combination with the previous principle of preventing, or at least minimizing, credit losses, should eliminate much of the fear of bank failures. It should permit efficient resolutions of large banks without strong negative reactions by the affected depositors and having to invoke "too-big-to-fail," which is now known as the "systemic risk exemption" to the prohibition on the FDIC against protecting uninsured claimants against credit losses in bank resolutions. This practice permits the FDIC to partially or totally protect de jure uninsured claimants in order to mitigate "any serious adverse effects on economic activity and financial stability."[14]

**Prompt reprivatization and recapitalization.** FDICIA requires that insolvencies be resolved at least cost to the FDIC. This requirement also reduces losses

to depositors at domestic offices that share the same priority and encourages rapid sale of bank assets after legal closure. Reprivatization can be more difficult when banks are publicly owned, including bridge banks. Public ownership of banks is not always rooted in the desire to allocate resources efficiently. Nor do publicly owned institutions necessarily seek to maximize profits. Rather, the intent may be to reallocate funds for socially desirable or political purposes. Thus, when a government-sponsored bank becomes insolvent, the government is likely to keep the institution in operation regardless of its financial condition, and its return to solvency is likely to be slower. The consequence is that losses are likely to continue, and the ultimate cost of resolution to the taxpayer is likely to be larger than it otherwise would.

To minimize government forbearance and its attendant costs, insolvent banks should be sold to the private sector in whole or parts as soon as this can be done efficiently. Indeed, in the United States, the maximum life of a bridge bank is specified by law to be no longer than two years, with three one-year extensions (which is probably longer than necessary). Moreover, the sale should be on terms that provide sufficient private capital to ensure that, after adjusting for any guarantees to the buyers, the resulting institution will attain, at minimum, "adequately capitalized," if not "well capitalized," status, to guard against a quick return to insolvency. Again, because under PCA the FDIC is aware of most pending insolvencies, it can begin the bidder search process for most banks before legal closure and the actual bidding at closure. As noted, larger banks may need to be bridged to give the FDIC additional time to sell to the highest bidders without having to resort to fire-sale losses or otherwise being forced to unwind the bank inefficiently.

## Potential Problems in Efficient Insolvency Resolution Introduced by Foreign-Owned Banks

Several characteristics of foreign-owned banks can make it more difficult for host countries to achieve some or all of the above four objectives for efficient resolution of insolvencies of foreign-owned banks. The problems arise primarily because foreign-owned banks are likely to be subject to more than one bank regulator, more than one deposit insurance agency, and more than one insolvency resolution agency—the home and host country entities—with likely different rules and procedures.[15] In addition, the bank offices in the host country are generally not fully independent of their headquarter office or parent holding company in the home country and thus are not fully under the scope of a single country. As a result, the potential for conflict and confusion is heightened.[16] For example, who is the primary prudential regulator, and

13. A recent survey of deposit insurance practices indicates that few countries (only about 15 percent) pay insured depositors within three months (see Demirgüç-Kunt, Karacaovali, and Laeven 2005). In large part, this payment delay reflects that the insurance agency has insufficient information on the identity of the insured depositors and the amount of deposits insured and requires the claimant to file a claim. In the United States, the FDIC generally has the necessary information and typically does not require depositors to file claims.
14. The systemic risk exemption is discussed at greater length in Kaufman (2004d).
15. See, for example, Garcia and Nieto (2005), European Commission (2005), Financial Stability Forum (2001), Eisenbeis and Kaufman (forthcoming), and Goodhart (2005) as well as a number of the essays included in Evanoff and Kaufman (2005). See Eisenbeis (2004) for a discussion of the conflicts that can arise for regulatory agencies, especially in situations where multiple regulators exist. See also Mayes (2004b).
16. Holthausen and Rønde (2004) and Kane (2005b) discuss models of imperfect information flows among regulators.

how does it regulate? Who provides deposit insurance and how? Who declares legal closure and how? And who resolves insolvencies and how?

These problems can be exacerbated because it may reasonably be assumed that regulators of all types operate primarily in the best interests of the citizens of their own country and not necessarily in the best interests of the host countries, particularly during bad economic times when insolvencies threaten (see Mayes 2005; Borchgrevink and Moe 2004; and Bollard 2005). At such times, home-host country conflicts can become more serious and move to the foreground and must be considered. The remainder of this section considers some potential problems with implementing the four efficient resolution principles articulated above in situations involving cross border banking, particularly when foreign-owned banks operate in host countries through branches rather than through separately chartered subsidiaries.[17]

*Obtaining useful, timely, and accurate financial data may be difficult for a host country regulator, especially when the foreign bank is operating through branches.*

As noted earlier, foreign subsidiaries of parent holding companies located in another country are either chartered or licensed by the host country, similar to domestic banks. But foreign branches may or may not be chartered, licensed, or approved by the host country. On the one extreme, the European Union (EU) has introduced a single bank charter that permits a bank chartered in any member country to establish branches in any other member country without further permission. Branches of European banks, regardless of where they operate, are subject to prudential regulation, basic deposit insurance, and insolvency resolution procedures of the home country.

On the other extreme, foreign branches in the United States are subject to national treatment, which is similar but not identical to treatment of domestic branches, and must be approved by both the Federal Reserve and either the Comptroller of the Currency or the state in which the branch is to be located. The Fed will approve a branch (or any other foreign office) only if it determines that the parent holding company is subject to comprehensive consolidated supervision by its home country comparable to that in the United States. The branch is subject to prudential regulation and supervision by the Fed and the other approving agency. Except for a small number of branches established before 1991 and grandfathered, foreign branches may not accept retail deposits (under $100,000) and their deposits may not be insured by the FDIC. They may be insured by the home country. Because it is not a separate legal entity, a foreign branch is required by U.S. regulation to maintain minimum capital equivalency deposits (CEDs) calculated as a percent of third-party deposits at third-party banks and may initiate prompt corrective action if these deposits decline to or below the specified minimum. If an institution becomes of regulatory concern, it may also be required to pledge assets of up to 105 percent of third-party deposits at the branch. Because these requirements help protect third-party deposits at the branch, foreign branches in the United States may be treated not very differently from foreign subsidiaries when they encounter financial difficulties. To focus attention on areas of greatest current concern, the analysis below assumes primarily unrestricted cross border branching as is permitted in the EU.

**Prompt legal closure.** Prompt legal closure at positive capital according to a prespecified and well-publicized closure rule requires timely, accurate, and meaningful information and data about the financial condition of the bank. Such information is more difficult for host country regulators to obtain or interpret for foreign-owned

banks. For branches of foreign banks, meaningful data apply only to the banking organization as a whole, not to individual offices. The Danmarks Nationalbank also notes that "the EU regulator does not generally grant the authorities of the host country insight into the risks associated with the activities of a branch" (Danmarks Nationalbank 2005, 65). Moreover, a bank may be legally closed for reasons of insolvency only by the home country. Therefore, except for reasons other than insolvency, legal closure of unlicensed branches is generally outside the control of the host country regulators. From the host county regulators' perspective, the closure decision is effectively outsourced. Different countries are likely to have different legal closure and PCA rules and may enforce these rules differently. Some countries, such as the United States, have special bankruptcy codes for banks, and others include banks in their general corporate bankruptcy codes.

Home country regulators may be influenced in both the decisions to officially declare a bank insolvent and on how to resolve the institution by the relative importance of the bank in its home country regardless of its relative importance in the host countries. Regulators may act differently in both timing and enforcement of legal closure, as well as in imposing corrective sanctions, for a banking organization that is small in the home country but large in the host countries relative to one that is large at home but relatively small abroad or is large in all countries. As a result, branches of different foreign-owned banks in host countries may be legally closed at different speeds depending on the home country. This difference may affect both the size and distribution of losses borne by depositors and other creditors at branches of banks in the same country but chartered in different countries. The greater the number of banks in a country that are chartered in other countries, the more confusing the situation is to both bank customers and regulators in that country. The closure rule and its enforcement may be either more or less efficient in the home country than in the host country.

Subsidiaries of foreign banks also present problems, but on a smaller scale and of a different kind. The subsidiary is a legal entity that is chartered or licensed in the host country (which is the home country for that subsidiary), and the host country has jurisdiction over it but not its parent. Increasingly, bank holding companies are being managed on a consolidated companywide basis. So while the data collected from the domestic subsidiary relate only to the subsidiary, the data may still be difficult to interpret for purposes of assessing solvency. For example, if the regulator perceives that an insolvent, or near-insolvent, subsidiary may be rescued by a solvent parent, then there may be less concern about an apparent weakness in that subsidiary. Conversely, if the regulator perceives that a healthy subsidiary may have its assets stripped by an insolvent or near-insolvent foreign parent, then it may be prudent for the regulator to be more concerned about the subsidiary's risk exposure. Indeed, in New Zealand, for example, authorities have imposed constraints on large subsidiary banks of foreign parents in an attempt to limit their dependence upon the parent or other affiliates outside of New Zealand (see Bollard 2005).

These considerations raise the practical question of when the host country regulator should legally close a possibly insolvent subsidiary and place it in receivership. What is the probability that the parent will rescue a subsidiary bank by injecting additional capital or that it will not inject capital and will walk away? The decision depends on a large number of factors such as size of the parent and the subsidiary, the countries involved, and so on. (Some of these possibilities are summarized in

17. Some of these issues were recently discussed in European Commission (2005) and Mayes (2004a).

Table 2
**Likely Implications for Host Country Treatment of Foreign Bank Subsidiaries of Insolvent Parent or Subsidiary Banks by Relative Size of Bank in Country**

| | | | | Home country (parent) | | | |
|---|---|---|---|---|---|---|---|
| | | | | Large bank | | Small bank | |
| | | | | Solvent | Insolvent | Solvent | Insolvent |
| Host country (subsidiary) | Large bank | Solvent | | NP | RR | NP | RR |
| | | Insolvent | | PC* | R | R** | R |
| | Small bank | Solvent | | NP | RR | NP | RR |
| | | Insolvent | | PC* | R | R** | R |

Notes:
NP No problem
RR Reputation risk/asset protection
PC Parent choice of rescue or walk and resolution with asset protection
R Resolution with asset protection
* Parent likely to rescue
** Parent likely to walk

Assumptions:
- Parent bank likely to attempt to "repatriate" assets at foreign subsidiaries in anticipation of official insolvency so host needs to protect subsidiary assets.
- Abstracts from functionality concerns re computer/records/senior management availability for operating subsidiary as independent (stand-alone) facility after insolvency and legal closure of either the subsidiary or parent.
- Abstracts from capital maintenance agreements between parent and subsidiary banks or host countries.

Table 2.) Should the host regulator move more aggressively if the parent but not the subsidiary is insolvent? The parent may try to transfer good assets quickly to itself or other subsidiaries in other countries. How well does a closure rule apply in such situations? How should regulators react if both the parent and the subsidiary are solvent but a separate operating subsidiary in a third country that provides critical services to the subsidiary is insolvent and may not be able to provide the services? Finally, some countries may mandate capital maintenance agreements that require foreign parents to come to the rescue of their subsidiaries. But how enforceable are such agreements across national borders?

**Prompt estimate of recovery values and assignment of losses.** Branches are not separate entities, and their assets are part of the entire banking organization. Thus, a host country regulator may have difficultly determining branches' financial condition. Because of interbank transfers, there may be little relationship between the assets and the liabilities booked at a branch in a host country. Any branch may have third-party assets greater or less than its liabilities to third parties. For this reason, it may be difficult to estimate losses in branch offices, and such estimates may not be very meaningful even if they could be done. Reliable loss estimates can be prepared only for the banking organization as a whole and, at best, may be able to be prorated to each depositor at each office. The promptness and reliability of these estimates depend largely on the home country, not the host country.

Loss estimates for insolvent separately capitalized subsidiaries are made by the host country and thus present problems that are not significantly different from those for insolvent domestically owned banks. However, last-minute asset-liability shifting across borders in anticipation of insolvency, although generally illegal, is more difficult to reverse than within a country, and the probability of solvent parent assistance should be factored into the estimates. Regulators have a strong tendency, however, to overestimate the probability of such parent rescues. As noted, parents

generally have the freedom to walk away from troubled subsidiaries and are likely to evaluate each case on its own merits.[18]

**Prompt reopening of large banks.** The timing of the physical reopening of foreign branches in host countries is often in the hands of the bank's home country. Likewise, both the deposit insurance scheme for foreign branches of a cross border branch banking organization and the rules for resolving the branches of insolvent organizations are often established by the bank's home rather than host country. If so, this situation may lead to serious problems (see Borchgrevink and Moe 2004). For example, the timing of the ability of both insured and uninsured depositors and other creditors to access their accounts and the value of their claims may not necessarily be determined in the best interest of the host country.

*Loss estimates for insolvent separately capitalized subsidiaries are made by the host country and thus present problems that are not significantly different from those for insolvent domestically owned banks.*

Deposit insurance schemes vary greatly across countries. Differences can exist in account coverage, premium assessments, solvency, administration, and private or taxpayer involvement. The speed with which payments are made to insured and uninsured claimants at insolvent institutions may also differ. The speed depends in part upon whether the insurance agency identifies the eligible claimants for payment or the claimants are required to file claims individually. In many countries, the deposit insurer, be it the host or home country, does not have current or complete data on insured deposit ownership and coverage. Thus, claimants need to file specific claims. This method is usually considerably more time consuming and would delay prompt reopening. (Some of these potential differences are shown in Table 3. See also European Commission 2005 and Garcia and Nieto 2005.) Most cross country studies of deposit insurance schemes have emphasized differences that affect the likelihood of solvent banks becoming insolvent. In contrast, this study emphasizes differences after banks become insolvent that affect the speed at which they may be reopened.

Home countries may also encounter political backlash when they are required to make promised insurance payments to depositors at insolvent banks in host countries, especially when the bank's foreign operations are large relative to its home country presence. Such cross border payment liabilities are effectively equivalent to foreign debt. This problem may be particularly acute when some of the payments are funded by home country taxpayers and the host country uses a different currency so that the foreign debt also involves exchange-rate risk. Thus, in host countries, branches of foreign-owned banks from different home countries may be physically reopened at different speeds after being legally closed depending on the practices of the respective home countries.

Importantly, the host country regulator may wish to see a legally closed branch office reopened quickly, independently of when offices are opened in the home country. But reopening a branch might not be feasible even if the host country were to operate the office temporarily because the branch may not have "functionality"—the senior management, the records, or the computer facilities necessary to operate on a stand-alone basis (Bollard 2005). Thus, at best, some physical closure of a foreign branch might not be avoidable. If the branch offices in the host country are

18. In the United States, the Federal Reserve has argued that parent holding companies should be a "source of strength" to their subsidiary banks and finance any losses that may occur from insolvency. This argument is not universally accepted.

Table 3
**Possible Deposit Insurance System Differences in Different Countries**

Account coverage
- Maximum amount
- Type of account, e.g., interbank
- Foreign currency deposits
- Coinsurance

Ownership
- Private vs. public (government)

Funding (premiums)
- Ex ante vs. ex post
- Magnitude
- Risk-based vs. flat
- Regular vs. "topping up"

Reserve fund
- Minimum magnitude
- Voluntary or required

Government support
- Explicit (official) vs. implicit
- Credibility of private funding (premiums)

Speed of payment if insolvency
- Insured depositors—to par value
- Uninsured depositors—to market (recovery) value
- Advance dividends vs. as assets sold

Claim filed
- Automatically
- By claimant

Pre-insolvency intervention
- Prompt correction action (PCA)

Declaration of insolvency
- Private creditors or government agency
- Insurance agency vs. other
- Closure rule vs. discretion (forbearance)

Insolvency resolution
- Administered by insurance agency, other agencies, or bankruptcy court
- Least-cost resolution (LCR)
- Insurer serves as receiver/conservator
- Too big to fail

Membership
- Mandatory or voluntary

Powers of insurance agency
- Payer only
- Supervisory and regulatory

Other
- Coinsurance
- Offsetting/netting

sufficiently important in that country, the host regulators may be able to minimize or avoid physical closure by requiring the branches to maintain sufficient redundancy in these functional areas or to credibly guarantee immediate availability. But such measures would increase costs and reduce whatever efficiencies might be associated with branching.

In contrast to branches, subsidiary banks are separate, stand-alone legal entities. The host country provides deposit insurance on the same basis as to domestically owned banks and resolves insolvencies, so many of the conditions for prompt reopening are in the host country's hands. Nevertheless, as holding companies are increasingly managed on a companywide consolidated or integrated basis, functionality may still be a problem. The necessary management, records, and computers to keep the facility functioning seamlessly may be at the parent or at another operating subsidiary in another country. Thus, host countries may also wish to require as a condition of chartering a subsidiary of a foreign bank that some redundancy be maintained domestically even though this may reduce benefits that may be derived from economies of scale or scope.[19]

**Prompt reprivatization and recapitalization.** Prompt reprivatization and recapitalization of insolvent foreign-owned banks may be more difficult than for insolvent domestic banks, especially to the extent that foreign-owned banks are resolved at a higher cost (greater negative net worth) for the reasons discussed above. There may be insufficient domestic capital, and a search for foreign capital may take longer. A host country may also have difficulty reprivatizing branches of

insolvent foreign banks independently of the actions of the home country. To do so would require the regulator to seize the branch offices and supercede the jurisdiction of the home country. Such seizure could delay the process and raise international legal issues. Finally, there may be cultural and language differences that may hamper a smooth and seamless transfer to potential owners in another country.

In sum, because of the increased difficulties in achieving at least some of the four principles for efficient resolution, foreign ownership of banks may increase the cost of insolvencies at such banks by increasing both the credit and liquidity losses in host countries relative to the losses experienced in resolving insolvent domestically owned banks. Higher resolution costs are more likely in cross border banking through branching on a single-license basis than for banking through subsidiaries.

## Possible Remedies for Foreign-Owned Bank Resolution Problems

It is apparent from the preceding analysis that the difficulties with efficient resolution of foreign-owned bank insolvencies lie primarily in the heterogeneity of both the closure rule and the deposit insurance structure across countries. These difficulties include differences in both provisions and enforcement; overlapping of legislation, regulation, and supervision between home and host countries; and inherent incentives for regulators to favor the welfare of their home countries, possibly at the expense of the host country. These problems are complex and do not lead to easy or simple lasting solutions. Moreover, they become increasingly significant as more and more banks operate banking offices in foreign countries.

Coordination and cooperation among home and host countries is a necessary but not sufficient condition to solve the problem.[20] What appears to be required is greater harmonization and homogeneity, particularly in closure policies and claims resolution. Indeed, centralized multinational regimes for deposit insurance and insolvency declaration (closure rules) and resolution, in terms of both provisions and enforcement, appear to be the most promising way to ensure that bank failures are resolved efficiently and without creating undue uncertainty.[21] Centralized regimes would eliminate the differences that make multiple individual country regulatory regimes and cross border enforcement and payments a severe problem.

But such a system raises numerous questions: Which countries should be included in the arrangement? How would those excluded be dealt with? How would the governing board be organized, and how would countries be represented on the board? What authority and enforcement power would such a board have? What funding would be available? And would the conflicts discussed above be eliminated by a single structure or merely internalized and hidden from view? These issues are significant enough that it is unlikely that a single, multinational structure for either deposit insurance or insolvency resolution could be adopted in the near future. Nevertheless, now is the time to begin the thinking process.

In the meantime, how should cross border banking be operated? Is there an intermediate or transition position that would encourage efficient resolution? A number of countries—most notably the Nordic countries, where cross border, single-license branch banking is scheduled to begin shortly—have established formal memoranda of understandings on joint supervisory policies and agreements governing the treatment of institutions in financial distress (Mayes 2005). Likewise, the major financial

19. Again, New Zealand has made such attempts (Bollard 2005; Borchgrevink and Moe 2004).
20. This issue has also been addressed by Danmarks Nationalbank (2005).
21. Kane (2005a) considers another alternative that involves the sale of options on insolvency losses.

public policy authorities in the European Union recently signed a Memorandum of Understanding on Cooperation in Financial Crisis Situations, which was announced in a press release, although the contents of the memorandum were not made public.[22] But, as argued above, it is unlikely that such cooperation will work effectively at all times, particularly not just when it is needed most—when a large cross border bank experiences solvency difficulties. Prudential regulators will have a hard time not putting their own country's interests first and cooperation second. Moreover, by not publicizing the details of the agreement, regulators are not fully revealing the rules of the game to all participants, whose actions are thus less efficient. Adoption of a common resolution scheme, possibly enforced by an independent multinational enforcement agency by EU countries permitting free cross border banking, particularly branching, may solve many of these problems. However, until a satisfactory long-term solution to the problem can be developed, it appears to be in the best interest of countries to limit physical cross border banking presence to subsidiary banks even though such limitations merely reduce and do not eliminate the problems and may not promote maximum efficiency in the majority of times when banks and banking are strong.[23]

*The difficulties with efficient resolution of foreign-owned bank insolvencies lie primarily in the heterogeneity of both the closure rule and the deposit insurance structure across countries.*

## Conclusion

This article explores some special problems in the efficient resolution of insolvent banks raised by cross border banking through foreign-owned banks and the forms that expansion may take. We suggest that, despite the many gains that may attend such expansion when times are good, potential problems may be both significant and more daunting when times are bad and bank insolvencies occur.

The article proposes four principles for the efficient resolution of insolvent banking offices against which to evaluate existing resolution structures for foreign-owned banks. Some characteristics of foreign-owned banks can make achievement of these principles by the host country more difficult. We conclude that the potential costs of such insolvency resolutions in the host country, especially when entry takes place by way of unrestricted branching, can be large not only in terms of larger credit and liquidity losses but also in terms of the potential confusion introduced by different closure rules and resolution policies across countries. In the absence of central multinational deposit insurance, a single insolvency resolution agency, and common or harmonized laws regarding insolvency resolution and enforcement, we suggest that entry by way of subsidiaries presents the lesser, but certainly not insignificant, set of problems for the host country and its residents. For all forms of entry, resolution costs can be most effectively controlled through the universal adoption of well-designed and enforced PCA-type policies and legal closure rules based on market values of assets and liabilities.

22. Press release, "Memorandum of Understanding on co-operation between the Banking Supervisors, Central Banks and Finance Ministries of the European Union in Financial Crisis situations," May 18, 2005 <www.ecb.int/press/pr/date/2005/html/pr050518_1.en.html> (November 11, 2005).
23. This solution is currently preferred by New Zealand.

## REFERENCES


Barth, James R., Gerard Caprio, and Ross Levine. 2001. Banking systems around the globe: Do regulation and ownership affect performance and stability? Paper presented at the NBER Conference on Prudential Supervision, Islamorada, FL, January.

———. Forthcoming. *Rethinking bank regulation and supervision: Till angels govern.* Cambridge: Cambridge University Press.

Berglöf, Erik, Paolo Fulghiere, Jordi Gual, Colin Mayer, Pedro Luis Pita Barros, and Xavier Vives. 2005. *Integration of European banks: The way forward*. London: Centre for Economic Policy Research.

Bliss, Robert R., and George G. Kaufman. 2005. U.S. corporate and bank insolvency regimes: An economic comparison and evaluation. Federal Deposit Insurance Corporation Center for Financial Research Seminar Series, June.

Bollard, Alan. 2005. Being a responsible host: Supervising foreign-owned banks. In *Systemic financial crises: Resolving large bank insolvencies*, edited by Douglas Evanoff and George Kaufman, 3–15. Singapore: World Scientific Publishing.

Borchgrevink, Hendrik, and Thorvald Grung Moe. 2004. Management of financial crises in cross-border banks. Norges Bank *Economic Bulletin* 75, no. 4:157–65.

Caprio, Gerard, and Daniela Klingebiel. 1999. Episodes of systemic and borderline financial crises. World Bank, photocopy.

Danmarks Nationalbank. 2005. Financial stability 2005. Copenhagen: Danmarks Nationalbank.

Demirgüç-Kunt, Asli, Baybars Karacaovali, and Luc Laeven. 2005. Deposit insurance around the world: A comprehensive database. World Bank Policy Research Working Paper No. 3628, June.

Eisenbeis, Robert A. 2004. Agency problems and goal conflicts. Federal Reserve Bank of Atlanta Working Paper No. 2004-24, October.

Eisenbeis, Robert A., and George G. Kaufman. Forthcoming. Cross-border banking: Challenges for deposit insurance and financial stability in the European Union. Federal Reserve Bank of Atlanta Working Paper.

European Commission. 2005. Review of the Deposit Guarantee Schemes Directive. Brussels, July 14.

European Financial Services Round Table. 2005. *On the lead supervisor model and the future of financial supervision in the EU: Follow-up recommendations of the EFR*, June.

Evanoff, Douglas D., and George G. Kaufman, eds. 2005. *Systemic financial crises: Resolving large bank insolvencies*. Singapore: World Scientific Publishing.

Financial Stability Forum. 2001. Guidance for developing effective deposit insurance systems, September.

Garcia, Gillian G.H., and Maria J. Nieto. 2005. Banking crisis management in the European Union: Multiple regulators and resolution authorities. *Journal of International Banking Regulation* 6, no. 3:206–26.

Goodhart, Charles A.E. 2005. Multiple regulators and resolutions. In *Systemic financial crises: Resolving large bank insolvencies*, edited by Douglas Evanoff and George Kaufman. Singapore: World Scientific Publishing.

Harrison, Ian. 2005. The Reserve Bank of New Zealand's Creditor Recapitalization (BCR) Project: An option for resolving large banks? In *Systemic financial crises: Resolving large bank insolvencies*, edited by Douglas Evanoff and George Kaufman. Singapore: World Scientific Publishing.

Hoggarth, Glenn, Ricardo Reis, and Victoria Saporta. 2002. Costs of banking system instability: Some empirical evidence. *Journal of Banking and Finance* 26, no. 5:825–55.

Holthausen, Cornelia, and Thomas Rønde. 2004. Cooperation in international banking supervision. European Central Bank Working Paper No. 316, March.

Kane, Edward J. 2005a. Can the European Community finesse the need for a multinational safety net? Boston College, May 24.

———. 2005b. Confronting divergent interests in trans-Tasman regulatory arrangements. Boston College working paper, July.

Kaufman, George G. 1997. The new depositor preference act: Time inconsistency in action. *Managerial Finance* 23, no. 11:56–63.

———. 2000. Banking and currency crisis and systemic risk: A taxonomy and review. *Financial Markets, Institutions and Instruments* 9, no. 2:69–131.

———. 2004a. Bank regulation and foreign-owned banks. Reserve Bank of New Zealand *Bulletin* 67, no. 2:65–74.

———. 2004b. Depositor liquidity and loss sharing in bank failure resolutions. *Contemporary Economic Policy* 22, no. 2:237–49.

———. 2004c. FDIC losses in bank failures: Has FDICIA made a difference? Federal Reserve Bank of Chicago *Economic Perspectives* 28, no. 3:13–25.

———. 2004d. Too big to fail in U.S. banking: Quo vadis? In *Too big to fail: Policies and practices in government bailouts,* edited by Benton E. Gup. Westport, CT: Praeger Publishers.

Kaufman, George G., and Steven A. Seelig. 2002. Post-resolution treatment of depositors at failed banks: Implications for the severity of banking crises, systemic risk, and too big to fail. Federal Reserve Bank of Chicago *Economic Perspectives* 26, no. 2:27–41.

Lastra, Rosa Maria. 2004. Cross-border bank insolvency. *Journal of International Banking Regulation* 5, no. 2:201–27.

Marino, James A., and Rosalind L. Bennett. 1999. The consequences of national depositor preference. *FDIC Banking Review* 12, no. 2:1–38.

Mayes, David G. 2004a. Preparing for cross-border bank failure. *Financial Regulator* 8, no. 4:58–63.

———. 2004b. Who pays for bank insolvency? *Journal of International Money and Finance* 23, no. 3:515–51.

———. 2005. The role of the safety net in resolving large financial institutions. In *Systemic financial crises: Resolving large bank insolvencies*, edited by Douglas Evanoff and George Kaufman. Singapore: World Scientific Publishing.

Office of the Comptroller of the Currency (OCC). 2003. Annual report: Fiscal year 2003. Washington, D.C., November.

PricewaterhouseCoopers. 2003. *A regulatory guide for foreign banks in the United States*. Washington, D.C.: Regulatory Advisory Service.

Reserve Bank of New Zealand. 2004a. RBNZ gives go-ahead to ANZ National Bank merger. <www.rbnz.govt.nz/news/2004/0153454.html> (November 11, 2005).

———. 2004b. Statement of principles: Bank registration and supervision. <www.rbnz.govt.nz/finstab/banking/regulation/1504882.pdf> (November 11, 2005).

Salmon, Richard, Lisa Anston, Jeanne McBride, Dennis Trimpea, Elvis Nelson, Debbie Barr, Bent Armond, Gwen Hudson, Donna Kinser, and Vicki Robinson. 2003. Costs associated with bank failures. Federal Deposit Insurance Corporation working paper, October.

Wall, Larry D., and Robert A. Eisenbeis. 2002. The major supervisory initiatives post-FDICIA: Are they based on the goals of PCA? Should they be? Federal Reserve Bank of Atlanta Working Paper No. 2002-31, December.

# How Should Banks Account for Loan Losses?

**GEORGE J. BENSTON AND LARRY D. WALL**

*Benston is the John H. Harland professor of finance, accounting, and economics at Emory University's Goizueta Business School, a professor of economics at Emory's College of Arts and Sciences, and a visiting scholar in the Atlanta Fed's research department. Wall is a financial economist and policy adviser in the Atlanta Fed's research department and an adjunct professor at Emory University's Goizueta Business School.* *The authors thank Randall Benston and participants at a seminar sponsored by the Federal Deposit Insurance Corporation for helpful comments and questions. They especially appreciate excellent suggestions and comments from Ray DeGennaro, Jerry Dwyer, and Scott Frame. A similar version of this paper appeared in the March–April 2004 issue of the* Journal of Accounting and Public Policy.

How should banks and other lenders account for the risk that a borrower will default on a loan? This accounting determines when and to what extent a bank recognizes changes in expected credit losses on its income statement and the value at which loans are reported on the bank's balance sheet. Should these amounts differ when the bank's financial reports are prepared for its shareholders and for the banking authorities, or should the authorities manage the numbers directly or indirectly through the capital requirement?

These questions are the source of an ongoing debate among banks and the agencies that regulate them: the Securities and Exchange Commission (SEC), the Financial Accounting Standards Board (FASB), and bank supervisors.[1] The FASB is a private-sector entity delegated by the SEC to set financial accounting standards, called generally accepted accounting principles (GAAP). Bank supervisors have responsibility for setting regulatory accounting principles (RAP). The Federal Deposit Insurance Corporation Improvement Act (FDICIA) mandates that RAP generally follow GAAP but allows exceptions that assist supervisors in performing their duties. However, any such exception must be "no less stringent than generally accepted accounting principles."[2] Bank supervisors are also granted an unusual role in setting GAAP by Section 241 of the Gramm-Leach-Bliley Act of 1999, which mandates that the SEC "shall consult and coordinate comments" with the appropriate federal bank supervisors before "taking any action or rendering any opinion with respect to the manner in which any insured depository institution or depository institution holding company reports loan loss reserves."

Wall and Koch (2000) discuss the philosophy underlying the different positions of the FASB and the SEC versus that of the bank supervisors. The authors conclude that the primary users of financial statements, investors and bank supervisors, are likely to form their own estimates of banks' loan losses. Thus, the philosophy underlying the reported figures may be less important than the transparency of the process

required by the accounting authorities and by the bank supervisors for their respective estimates of loan losses. Nevertheless, the debate over appropriate procedures continues, with Davenport (2004) reporting that the FASB intends to take up the issue again.[3]

This article seeks to answer the question of how the value of loans on a bank's balance sheet should be adjusted for expected credit losses. Underlying the analysis is the assumption that the value most useful to bankers, investors, and bank supervisors is the economic value of loans as of the balance-sheet date. This value is the present (discounted) value of the cash payments the bank expects to receive from the borrowers as a group, which is less than the promised amount, because the bank cannot perfectly predict which loans will default. When loans are recorded at their economic values, there should be no reduction in this amount with an allowance for loan losses because the interest rate charged should be sufficient to cover expected default losses. Our analysis leads us to conclude that even when loans are accounted for at historic cost, under most situations no allowance for loan losses should be made. Loan-loss expense for a period, then, is the loss incurred on loans that defaulted during the period. However, if cash flow expectations change so that loans decrease in value, the decrease is an addition to loan-loss expense.

*Present accounting principles are largely based on a system that values assets and liabilities at their historic cost rather than at their current market value.*

Although historic-cost GAAP can provide useful information, economic value clearly would be superior if measurement of the numbers were reliable and cost effective. The problem is that the economic values of loans are not readily observable unless the loans are traded in sufficiently liquid markets. Many loans, though, are not traded because of information asymmetry between the bank and potential buyers. A bank's knowledge of its customers includes strengths and weaknesses that may not be apparent from documents that describe the lending situation. This basic attribute of loans results in banks often placing a greater value on a loan than buyers are willing to pay or in buyers discounting loans because they fear the seller may be holding back important negative information (adverse selection). Consequently, loan values and the related losses must be estimated. Both accountants and bank supervisors are concerned about the results of this estimation, but each group has a different perspective on the question.

The accounting authorities' concern is for general-purpose users of financial statements, particularly investors. Investors may pay too much for bank stocks if banks' reported losses are understated, or they may sell their stock too cheaply if the losses are overstated. Thus, the accounting authorities are worried about biased estimates in either direction, particularly when the bias is intentional. The authorities recognize that managers, who prepare the financial statements, sometimes have incentives to use loan-loss accounting to manipulate the numbers reported. In some situations managers have an incentive to understate expected losses in order to boost reported net income and capital in the current period. In other situations managers have an incentive to overstate losses in the current report when earnings are high so that they can understate losses in a later period when other earnings are low, thereby smoothing the reported net income.[4]

In contrast, bank supervisors are concerned about banks being inadequately capitalized and possibly failing. Banks should maintain loan-loss allowances sufficient to cover expected losses and maintain sufficient equity capital to absorb unexpected

losses. Bank supervisors argue that reasonable approaches to estimating loan losses are likely to yield a range of estimates. Given supervisors' focus on safety, they want banks to report a loan-loss allowance that is on the high end of those estimates.

This article considers these somewhat different approaches to banks' loan-loss accounting by reviewing existing GAAP for loan-loss accounting in the following section. The next section compares the economic value of a loan with its reported value under current GAAP. The article then reviews the GAAP "reliability" and "relevance" criteria and analyzes existing GAAP and proposals to value loans and other financial assets and liabilities at their "fair values," particularly with respect to the criteria.[5] We then analyze bank supervisors' concerns about loan-loss accounting. The last section summarizes the results of prior discussion and presents conclusions that are relevant for policy.

## Current Accounting Standards

Existing accounting standards as determined by the FASB are specified in its Statements of Financial Accounting Standards (FAS).[6] The two statements most relevant to loan-loss accounting are FAS 5, *Accounting for Contingencies*, and FAS 114, *Accounting by Creditors for Impairment of a Loan.* FAS 5 (paragraph 8) sets two standards, both of which must be met in order to recognize a loss contingency in a firm's financial statement:

a. Information available prior to issuance of the financial statements indicates that it is probable that an asset has been impaired . . . at the date of the financial statements. It is implicit in this condition that it must be probable that one or more future events will occur confirming the fact of the loss.
b. The amount of the loss can be reasonably estimated.

FAS 5 defines "probable" as "the event or events are likely to occur" (paragraph 3). "Probable" is generally interpreted as "more likely than not," or having at least a 50 percent chance of occurring.

FAS 114 provides additional guidance on accounting for individual loans that are impaired in terms of both defining impairment and measuring the extent of

1. These bank supervisors are the Office of the Comptroller of the Currency (OCC), the Board of Governors of the Federal Reserve System, the Federal Deposit Insurance Corporation (FDIC), the Office of Thrift Supervision (OTS), and the National Credit Union Administration (NCUA). Although state banking agencies also supervise banks they charter, we are not aware of any that have taken an interest in this issue.
2. Section 121 of FDICIA, codified to 12 *U.S. Code* 1831n(a).
3. The FASB's indication that it would reconsider loan-loss accounting follows the abandonment of a proposal by the American Institute of Certified Public Accountants (AICPA) to provide guidance on the implementation of existing GAAP, according to Davenport (2004). Bank supervisors had recommended abandonment in a letter to the chairman of the AICPA dated October 6, 2003, available at <www.bdbonline.net/pdfs/RegulatorsLetter.pdf>.
4. Managers also might want to manipulate expenses generally to reduce income taxes. However, since 1986 banks may deduct only realized rather than estimated credit losses. See Walter (1991, especially 24–25) for a discussion of the historic role of taxes in determining loan-loss accounting.
5. The FASB explains that by "fair value" it means "an estimated market exit price, that is an estimate of the amount that would have been realized if the entity had sold the asset" (1999, paragraph 12). Fair value includes, but is not limited to, market value because appraisals, models, and present-value calculations may be used when relevant market values are not available.
6. The texts of FAS may be obtained from the FASB's Web site, <www.fasb.org/st/>.

impairment.[7] FAS 114 amends FAS 5 to indicate in paragraph 8 that "a loan is impaired when, based on current information and events, it is probable that a creditor will be unable to collect all amounts due according to the contractual terms of the loan agreement." Several measures of impairment are permitted under FAS 114, including "the present value of expected cash flows discounted at the *loan's effective interest rate, . . .* observable market price, or the fair value of the collateral if the loan is collateral dependent" (paragraph 13, emphasis added). FAS 114 specifies that "the effective interest rate . . . is based on the original contractual rate, not the rate specified in the restructuring agreement" (paragraph 14). This requirement drives a wedge between economic and accounting values because restructured loans usually are riskier than when originally made. Hence, the interest rate GAAP requires tends to be lower than the market rate applicable to such loans, thereby overstating the actual (economic) value of the loans.

*The philosophy underlying the reported figures may be less important than the transparency of the process required by the accounting authorities and by the bank supervisors for their respective estimates of loan losses.*

The treatment of loans in FAS 5 and FAS 114 does not apply in two special cases. First, if loans have been securitized and are traded in financial markets, they must be accounted for according to FAS 115, *Accounting for Certain Investments in Debt and Equity Securities*. FAS 115 requires classification of debt securities into one of three categories: held to maturity, available for sale, or trading securities. Securities classified as held to maturity are accounted for using historic cost; securities that are available for sale and trading securities are carried at their fair values. Contemporaneous changes in fair value of trading securities are included in current earnings. For securities held for sale, these changes are reported in a supplementary statement of comprehensive income and are included in the income statement only when the changes are realized (for example, the securities are sold).

The second exception to FAS 5 and FAS 114 is loans hedged by a derivative contract. FAS 133, *Accounting for Derivative Instruments and Hedging Activities,* requires that when a derivative is designated as a hedge for a specified asset or liability, the fair-value gain or loss on both the derivative and the hedged item should be recognized. Thus, if a derivative instrument hedges the fair value of a loan, the loan will be reported at its fair value.

## Economic and Accounting Values of Loans

The economic value of a loan (or any asset) is the present value of the expected cash receipts.[8] A banker's estimate of the economic value of a loan need not equal the value the banker would report for the asset from a straightforward application of GAAP.[9] The banker would be expected to use all available information, taking into consideration the cost of obtaining and using the information, to estimate the loan's economic value to the bank. However, the banker is allowed to use only a subset of the available information to report the GAAP value of the loan unless default is probable.

The differences between economic and accounting value may be more precisely identified and analyzed by expressing economic value in notational form. In the equations that follow, uppercase letters represent dollar values and information sets, lowercase letters represent rates and probabilities, and subscripts represent a time period.

The basic equation for calculating the economic value at time 0 of any set of cash flows over the interval from time 0 to time $N$ is

$$(1)\quad EV_0 = \sum_{t=0}^{N}\left(C_t(I_0)/(1+dr)^t\right),$$

where $EV_0$ is the economic value of the loan at time 0; $C_t(I_0)$ is the expected cash received at time $t$, given the information $I$ available at time 0; and $dr$ is the discount rate demanded by investors to compensate for the time value of money, any risks associated with the loan, and the illiquidity of the loan. Equation (1) says that the economic value of an asset at time 0 is the sum of the expected cash flows given the information at time 0, discounted at the appropriate interest rate.

The issues in loan-loss accounting may be illustrated by making two changes in equation (1). First, the expected cash flow at time $t$ may be separated into the promised cash flow less the expected loan losses due to default on the part of the borrower. Second, the promised cash flow may be further separated into promised interest payments and promised principal payments. The economic value of a bullet (or nonamortizing) loan with these modifications to equation (1) is[10]

$$(2)\quad EV_0 = \left[\sum_{t=0}^{N}\left(lrP/(1+dr)^t\right) + P/(1+dr)^t\right] - \sum_{t=0}^{N}\left(pd_t(I_0) * LGD_t(I_0)/(1+dr)^t\right),$$

where $lr$ is the rate of interest paid on the loan; $P$ is the principal value of the loan; $pd_t(I_0)$ is the probability that default will occur at time $t$, given the information available at time 0; and $LGD_t(I_0)$ is the loss given default at time $t$, conditional on the information available at time 0.

The first term (inside the square brackets) of equation (2) represents the discounted promised payments on the loan—the discounted value of the promised interest payments on the loan ($lrP$) plus the discounted value of the promised principal repayment.[11] The second term is the discounted value of the expected loan losses given the information at time 0—the probability that the loan will have defaulted at time $t(pd_t)$ multiplied by the expected loss given default at time $t(LGD_t)$. Equation (2) may be simplified without changing our conclusions by assuming that the value of $LGD$ is constant for every time $t$ and does not change with the arrival of new information.[12]

7. FAS 114 specifically excludes from its scope "large groups of smaller-balance homogenous loans that are collectively evaluated for impairment" (paragraph 6). Such loans are often called retail loans and include credit card and automobile loans.
8. The economic value of a loan after it has been made also depends on the bank's cost of administering the loan. We do not include a term for this cost in our analysis because doing so would complicate the notation without providing any additional insight on loan-loss accounting.
9. In this section we will assume that the banker would honestly report her best estimate of a loan's value in accordance with GAAP. The following section considers the possibility that other considerations may influence the reported value of a loan in a bank's financial statements.
10. Equation (2) could be easily modified to analyze an amortizing loan (a loan with periodic principal prepayments) at the cost of somewhat more complicated notation. However, doing so would not change our qualitative results.
11. We could allow the discount rate on the loan to vary through time and with the arrival of new information, but doing so would not change our qualitative insights on accounting for credit losses. The issues associated with accounting for time variation in the discount rate are discussed below.
12. We realize that the loss given default could change through time or with changes in information, such as the appreciation or depreciation in the value of collateral. While such changes would affect the magnitude of the difference between economic and accounting values, these changes would not alter our qualitative conclusions.

The economic value of a loan may change from one period to the next as a direct consequence of the passage of time.[13] The value may also change because of the arrival of new information that alters the expected values of the probability of default in future periods. Thus, the economic value of the loan at the start of the next period, period 1, is

$$(3)\quad EV_1 = \left[\sum_{t=1}^{N}\left(lrP/\left(1+dr\right)^t\right)+P/\left(1+dr\right)^t\right] - \sum_{t=1}^{N}\left(pd_t\left(I_1\right)*LGD/\left(1+dr\right)^t\right).$$

The change due merely to the passage of time is reflected in the change in the period over which the cash flows are summed—from time 0 to time $N$ in equation (2) and from time 1 to time $N$ in equation (3). The new information is reflected by updating the information set used to calculate the probability of default from $pd_t(I_0)$ in equation (2) to $pd_t(I_1)$ in equation (3).

In contrast, accounting values respond to a change in $pd_t$, and hence accord with economic values, only if a loss is probable, the loan is hedged by a derivative, or the loan has been securitized and is not categorized as held to maturity. Loans that do not meet any of these criteria are hereafter called a "typical loan" in recognition that most loans in banks' portfolios do not meet these criteria. The GAAP value of a loan may differ from its economic value for three important reasons, discussed next.

**GAAP value cannot exceed the principal value of the loan.** The accounting, or GAAP, value of a typical loan may differ from its economic value both at the time the loan is made and in future periods in part because a loan's maximum GAAP value is its principal value. In contrast, a loan's economic value would be expected to exceed its principal value if the market for loans is less than perfectly competitive. In addition, under GAAP the cost of acquiring the loan and putting it on the books (the initial cost) is almost never capitalized (added to the face value of the loan). The loan's economic value is greater than the face value by the present value of these amounts, which the bank collects over time as part of the borrower's interest payments.

**GAAP value does not incorporate all new information about *pd*.** In the periods after a loan is first recorded, the GAAP value may differ from the economic value because GAAP permits consideration of changes in the probability of default (*pd*) only if default is probable. This limit on the recognition of changes in default probabilities would result in the GAAP value of a loan being less than its economic value when new information indicates that the probability of default has declined. However, the GAAP value of a loan could be more than its economic value when new information indicates an increase in the probability of default (*pd*) in future periods but the cumulative probability of default remains below 50 percent (is not probable).

**GAAP value generally ignores anticipated changes in *pd*.** A loan's GAAP value may also differ from its economic value if the hazard rate—the probability of default conditional on the loan not defaulting in a prior period—is not constant through future periods, given the information available at any point in time.[14] A large set of possible alternative variations exists in the probability of default. One common pattern with important regulatory implications is one in which the probability of default is very low for several periods after the loan is made and then increases. For example, experience might indicate that borrowers tend to maintain their promised payments for several years, both because the projects for which they borrowed the funds

take time to fail (if they do) and because they can hide the failure for a while by holding back some of the borrowed funds and then using the funds to meet scheduled interest and principal payments. Experienced bankers probably recognize this pattern of risk and behavior. The contract rate of interest paid on the loan, *lr,* though, is almost never structured to change with expected changes in credit risk during the life of the loan. Rather, the rate is set to compensate the bank for the net present value of expected credit losses over the life of the loan.[15] As a result, in this situation the contract rate of interest will overcompensate for credit risk in some earlier periods and undercompensate for credit risk in some future periods.[16]

Consider the simplest case in which the probability of default increases every period from initiation to maturity. Then, stated in terms of our notation, for some periods from the initiation of the loan at 0 through some later point $M$, the expected loss due to default is less than its average value over the life of the project (assumed to be $N$ periods):

$$(4)\quad pd_i(I_0)LGD < \sum_{t=0}^{N} \frac{pd_t(I_0)LGD}{N}$$

for $i \leq M$. Then in subsequent periods

$$(5)\quad pd_j(I_0)LGD > \sum_{t=0}^{N} \frac{pd_t(I_0)LGD}{N}$$

for $M < j \leq N$.[17] Note that all values of the probability of defaults in equations (4) and (5) are based on the information set $I_0$. Given that the contract rate on the loan, $lr$, is set at time 0 to compensate the firm for the average loss over the period, the implication is that the bank collects a credit risk premium early in the life of the loan that exceeds the required compensation for default losses for the first $m$ periods, but in subsequent periods the credit risk premium will be less than the required compensation.

Although predictable time variation in $pd$ drives a wedge between earnings reported on a GAAP basis and earnings based on economic values for individual loans, the wedge often will not be material at the level of portfolios of similar loans. Portfolios often contain loans of varying maturities with the overstatement of earnings on relatively recent loans being more or less offset by understatement on older loans. Indeed, if the portfolio contains approximately equal amounts of loans at every

13. An exception for which the economic value of the loan does not change arises if the probability of default is constant, the discount rate is constant, and the market for loans is competitive.
14. The interest paid on a loan in any given period compensates for its expected losses in that period—the hazard rate times the expected loss given default.
15. This analysis would not hold if $lr$ varied in response to changes in the borrower's credit quality. However, most variable-rate loans do not adjust to changes in the borrower's credit quality but rather in response to changes in a market rate, such as the London Interbank Offer Rate (LIBOR). The primary reason for movement in these market rates is changes in the default-free risk rate. Risk-related movements in that rate are the result of changes in overall credit risk and not to the risk of a specific borrower or loan.
16. This situation is similar to a level payment on a whole-life insurance policy, where the early-period payments exceed the expected death payout in those periods.
17. This formulation assumes that the probability of default is increasing over the entire life of the loan. While we could model more complicated time patterns, doing so would complicate the analysis without adding any important new insights.

Table 1
**Comparison of GAAP and Economic Values**

| Situation | GAAP value | Economic value |
|---|---|---|
| **A: Value at loan initiation** | | |
| Competitive market for individual loans | Loan principal | Loan principal plus the costs of making the loan |
| Lender has market power | Loan principal | Greater than the loan principal |
| **B: Reported loan value after initiation** | | |
| Default is not probable | Loan principal | Economic value |
| Default is probable | Economic value | Economic value |

maturity between origination and maturity, the difference between GAAP earnings and earnings based on economic value would be about zero. Thus, predictable time variation in $pd$ is unlikely to have a material effect on the difference in earnings unless the size of the loan portfolio is significantly increasing or decreasing.

**Summary of economic and accounting treatment.** The differences between GAAP values and economic values are summarized in Table 1. GAAP requires the reporting of the principal value of the loan in all cases except where default is probable, as shown in panel B. For example, even if new information suggests that the cumulative probability of default over the remainder of a loan's life increased from 1 percent to 49 percent, GAAP would not allow any recognition of the greatly increased probability of default. In contrast, the economic value of the loan at its initiation is likely to be greater than the principal of the loan, as shown in panel A. Thereafter, the economic value of the loan will respond to changes in the discounted value of the promised payment less the discounted value of the expected loan losses, as shown in panel B. Thus, unlike GAAP reported values, the economic value of a loan would decrease in response to new information, suggesting the probability of default had increased from 1 percent to 49 percent.

The differences between the definition of GAAP and economic values for individual loans imply that the two values respond differently to new information and the passage of time if default is not probable, as summarized in Table 2. The GAAP value remains at the principal value of the loan. In panel B, the arrival of new information about the probability of default results in economic value moving inversely with the change in the probability of default.

Panel C of Table 2 summarizes the changes in economic value if the probability of default increases throughout the life of the loan. Absent any news after initiation, the loan's value declines as time proceeds when the probability of default in the period is less than its per-period average probability. Assuming the loan does not default, its value increases later in its life when the average probability of default is above its per-period average amount. Note that panel C gives changes in the economic value of an individual loan and not a portfolio of loans. If a bank's portfolio includes individual loans of approximately equal amounts and maturities, the probabilities of default in any one period tend to average out and be approximately equal among periods.

One overall implication that may be taken from this analysis is that the reported GAAP value is likely to understate the economic value of most banks' portfolios most

Table 2
**Comparison of Changes in GAAP and Economic Values[1]**

| Situation | Change in GAAP value | Change in economic value |
|---|---|---|
| **A: No new information, probability of default is equal in every remaining period** | | |
| No new information, probability of default is equal in every remaining period | No change | No change |
| **B: New information, per-period probability of default is constant over the remaining life of the loan** | | |
| New information suggests an increase in the probability of default | No change | Economic value decreases |
| New information suggests a decrease in the probability of default | No change | Economic value increases |
| **C: No new information, per-period probability of default is increasing over the remaining life of the loan** | | |
| Probability of default is less than average probability of default | No change | Economic value decreases |
| Probability of default is greater than average probability of default | No change | Economic value increases |

[1] Comparison of changes in GAAP and economic values assumes that the economic value equals the loan value at initiation, the loan principal is due at maturity, and default is not probable.

of the time. Consider a loan portfolio consisting of the same amount of loans at every possible maturity and the same probability of default on every loan. If the realized probability of default is equal to the portfolio's expected value, the portfolio should be worth more than its principal value because the initial costs of making and booking loans are not capitalized. If the bank can charge a higher-than-competitive interest rate, perhaps because it has market power, the portfolio would be worth even more. The decline in the value of loans when default becomes more likely would be offset by the increase in the value of loans when default becomes less likely. However, GAAP accounting allows recognition of decreases in the value of loans only in cases in which default is probable and does not recognize increases in loan value. Thus, GAAP accounting requires the bank to report its loans at less than their economic value or their principal value even when overall defaults in the portfolio exactly match the bank's expectations.

## Analysis of Present and Proposed Accounting Rules

Present accounting principles are largely based on a system that values assets and liabilities at their historic cost rather than at their current market value. Historic-cost accounting reflects an emphasis on providing reliable financial information even if the information is not the most relevant to the problem facing the decision maker. The FASB's move toward use of fair-value accounting, particularly for financial instruments, reflects its belief that fair values could and would be measured sufficiently reliably by managers and be audited effectively by independent public accountants and, consequently, would provide more relevant information to decision makers. This section begins by discussing the concepts of "relevance" and "reliability" as expressed by the FASB and then shows how present GAAP has systematically selected options that have greater reliability at the cost of decreased relevance. The next subsection considers the issues raised by the FASB's proposed move to fair-value accounting.

**Concepts of reliability and relevance.** The FASB has published seven statements on Financial Accounting Concepts (FAC) that explain the concepts it seeks to implement in providing guidance on specific accounting issues. In FAC 2, the FASB explains how it evaluates alternative accounting choices: "The better choice is the one that . . . produces from among the available alternatives information that is most useful for decision making" (1980, paragraph 30).[18] Paragraph 15 explains that "the qualities that distinguish 'better' (more useful) information from 'inferior' (less useful) information are primarily the qualities of *relevance* and *reliability*" (emphasis in original). Paragraphs 60 and 61 employ an analogy to medical drugs to distinguish the two qualities. A drug is reliable if the contents of the bottle conform to the formula shown on the label. A drug is relevant if it is effective, that is, if it cures or alleviates the underlying condition. The analogy, though, is incomplete. In many instances the choice is between a reliable drug of limited curative value (historic cost) and a drug that might be effective if it were reliable (fair value based on actual market values) but would most likely be harmful if it were believed to be reliable but actually was unreliable (fair value based on management's subjective estimates). The problem is that financial statements are the responsibility of a firm's management, who provide the estimates of market values.[19] Although managers sometimes have an incentive to underestimate economic values, more typically they may be expected to overestimate these values because their performance evaluations and compensation often are heavily influenced by reported values.

*Recognizing larger expected loan losses in financial statements does not give banks added resources to absorb losses.*

**Present GAAP rules.** GAAP often encompasses a trade-off between reliability and relevance, using values based on historic costs to measure some items and economic values to measure others. Contemporary loan-loss accounting reflects such a mix, with loans recorded at historic cost but with the loan-loss allowance based on estimated market values under certain conditions. Both the use of historic cost to record loans (excluding the initial cost) and the conditions placed on the use of market values serve to increase the reliability of the financial statements.

The analysis of the economic values presented here indicates that a loan should be expected to have a positive economic value to the bank at the time the loan is made whenever the relevant market is not perfectly competitive. Even when the market is perfectly competitive, the loan's value to the bank is greater than the amount the borrower receives because the interest payments include amounts that compensate the bank for its operating costs and return on capital. However, current accounting principles require that the loan be recorded at the amount loaned to the borrower, with initial operating costs (for example, acquisition, credit check, and administration) charged off as current-period expenses. Although the loan amount is a historical number that tends to understate the loan's economic value to the bank, it has the advantage of having been measured reliably and inexpensively. In contrast, the loan's economic value when it is initiated is an estimate that must be made by the preparer of the financial statements—the bank's management.

If loans were not recorded at historic cost, a bank could increase its reported net income merely by making additional loans near the close of the accounting period and recognizing its management's expectations of the discounted profits from the new loans as additional loan value and current-period income. Not only might such profits never be realized, but also such an accounting procedure would create an incentive for some managers to book new loans solely to record estimated and overoptimistic

profits. This situation is more likely to occur when managerial bonuses depend on recorded profits or when managers want to offset losses that must be recognized.[20]

As described earlier, loan-loss accounting requires recognition of estimated losses in certain circumstances, particularly when a loss is probable (FAS 5). This requirement should discourage (but probably does not prevent) banks and other lenders from adjusting their estimates of expected losses as a tool for smoothing net income.[21] Absent this rule, given a sufficiently large portfolio of loans, a bank could materially reduce and later increase reported net income by making changes in expected probabilities of default or loss-given-default of a magnitude that would be difficult, if not impossible, for any outside party to disprove or even verify. Although the requirement that a loss be probable still requires the use of an estimate, it limits the application to those loans that are or soon will clearly be in distress. Moreover, the common practice of banks working with distressed borrowers further reduces the opportunity for a banker to assert incorrectly that a loss is probable on a loan with elevated risk. As a part of the loan workout, the bank often will relax terms that the borrower may have difficulty meeting. For example, the bank may lower interest payments but require additional collateral. These changes reduce the expected loss to the bank and thus tend to reduce the probability of default to less than probable, thereby obviating the loan's being written off or down.

Another FAS 5 requirement increases the reliability of reported losses. FAS 5 maintains that the probability of a loss must be based on information that is known or knowable as of the financial statement date. Although reliability is not the primary reason for these requirements, they increase reliability by requiring that losses be based solely on past events and not in anticipation of future events.[22] However, because an asset's economic value is based on all available information (including the present value of future events that have a non-negligible but less than 50 percent probability of occurring), as described earlier, this GAAP rule tends to overstate some loan values.

When a bank expects that default losses in a portfolio of loans are not constant over time but will, say, increase, to be consistent with GAAP, a bank should recognize that the early-period loan interest payments include compensation for expected future losses. To account for this situation, a bank should record that portion as a deferred credit in the liability section of the balance sheet, with an offsetting reduction of loan-interest revenue. When the higher loan loss is incurred as expected, the accounting entry would reduce (debit) this deferred credit. But this accounting is almost never done because it involves difficult-to-make estimates and cumbersome bookkeeping

18. The decisions are presumably made by "present and potential investors and creditors and other users making rational investment, credit, and similar decisions" (FAC 2 1980, paragraph 22).
19. Although independent public accountants review management's estimates, their role is not to provide substitute estimates but rather to verify that management's estimates were derived by the consistent application of generally accepted procedures and that the numbers presented conform to GAAP.
20. This situation occurred when Enron adopted fair-value accounting for a substantial portion of its activities, as shown in Benston (2005). Overvaluation to obtain bonuses is also given by the FDIC (2000) as an example of the accounting abuses at Pacific Thrift and Loan Company.
21. Note that if banks' use of the allowance for loan losses were substantially eliminated, as our analysis leads us to suggest, this form of income smoothing would be obviated.
22. The primary reason for this requirement in FAS 5 is that GAAP is concerned with the measurement of periodic net income. The intent is that current income should reflect only those events that occurred during or before the reporting period and that expenses incurred to earn the income should be reported (matched) in the same period.

entries. Furthermore, when loan portfolios are stable, the credits and debits tend to wash out because overstatements of current-period loan-loss expense approximately equal overstatements of current-period interest-income revenue.

**Fair-value accounting.** The FASB's interest in implementing fair-value accounting for financial instruments dates back at least to 1991, with FAS 107, *Disclosures about Fair Value of Financial Instruments.* FAS 115, *Accounting for Certain Investments in Debt and Equity Securities* (enacted in 1993), requires the inclusion of fair values in the balance sheet and income statement (rather than only disclosure in footnotes) for securities not held to maturity for which reliable market prices could be determined by reference to securities regularly traded on recognized securities exchanges. FAS 133, *Accounting for Derivative Instruments and Hedging Activity* (enacted in 1998), expresses the FASB view forcefully: "*Fair value* is the most relevant measure for *financial instruments* and the only relevant measure for derivative instruments" (emphasis in original).[23] The primary benefit of fair-value accounting, according to the FASB, is discussed in its *Preliminary Views*:

> The major conceptual advantage of fair value as a measurement attribute is that because it is a market-based notion, it is unaffected by:
>
> a. *The history of the asset or liability*. Fair value does not depend on the date or cost at which an asset or liability is acquired or incurred.
> b. *The specific entity that holds the asset or owns the liability.* Fair value is the same no matter which entity has an asset or liability if both entities have access to the same markets and, for a liability, if they have the same credit standing.
> c. *The future of the asset or liability.* That is, fair value does not depend on the intended disposition of an asset or liability. (1999, paragraph 3)

Thus, fair values are measured as values in exchange, the amounts for which an asset can be sold or a liability extinguished. These exit values necessarily understate the values to investors in companies that do not expect to dispose of their assets. For these "going concerns," the value of assets is their value in use, that is, their present values. If assets could be sold for more than their value in use (including additional value from related business and net of transactions costs), they should and usually would be sold. Hence, assets that are kept would almost always have greater values in use than in exchange, and fair values understate the economic value of those assets to the owners of an enterprise.

**Application of fair values to loan-loss accounting.** If loans could be reported reliably at fair value, where fair value is value in use, there would be no need for a loan-loss provision or allowance. The fair value of the loan portfolio would be reported as an asset, and the change in the fair value of the portfolio, positive or negative, would be recognized on the income statement. The problem with applying fair value is that no market exists for many loans because banks obtain information about borrowers' credit quality that cannot be credibly conveyed to potential buyers of the loan. As a consequence, potential loan buyers are concerned about adverse selection, which here means the possibility that the seller is selling a particular loan because the seller has adverse private information about the loan. The result is that the market for selling some types of loans either breaks down or exists only because the seller retains part of the credit risk. A market for the full transfer of credit risk does not exist because banks place a greater value on the loan than potential buyers are willing to pay.

The problem of information asymmetry arises even in cases in which a market exists for low-risk standardized loans, such as the market for mortgage loans. The contracts between mortgage originators and securitizers impose penalties on the originators if the credit losses exceed some specified minimum amount (Frame 2003). Consequently, a bank may choose to hold rather than securitize such loans. Since loans that are securitized are reduced by a discount reflecting investors' concerns about adverse selection, application of those prices to value loans that are retained understates their value to the firm because such loans are not subject to this concern.

Fair values of loans held by a bank could be approximated, however, by taking account of changes in such observable variables as changes in market rates of interest that apply to loans of a given kind. For example, if the interest rate on conventional mortgage loans with similar terms (such as down payment and maturity) changed, loan values could be determined by discounting the expected cash flows by the current interest rate. (See the sidebar on page 32 for additional discussion of this issue.)

## Analysis of Bank Supervisors' Position

The comparison of economic value and reported accounting values in the second section suggests that adherence to existing accounting principles will result in reported values that are generally less than the loans' economic value (with some important exceptions).[24] Thus, even though loan losses reported in accordance with GAAP reflect only probable losses, net loan values on average will be conservative, as desired by the bank supervisors. Nevertheless, bank supervisors continue to press the case for even more conservative valuations.[25] Are the supervisors seeking valuations that are excessively conservative?

Even though reported loan valuations are conservative on average, they are not conservative in all cases. As we mentioned earlier, one important exception can occur when the expected probability of default ($pd$) is low in the first few periods after the loan is made but increases substantially over time. A portfolio of these loans will appear very profitable in the early years and only later will reflect large losses. Consequently, financial statements based on current GAAP could materially overstate profits and loan values if these types of loans are a large and rapidly growing part of a bank's asset portfolio, assuming that these overstatements are not offset by the GAAP understatement of economic values when loans are initially recorded. Supervisors may reasonably be concerned that allowing a bank to expand its loan

23. The expression of an interest in fair-value accounting in FAS 133 is not surprising, given the difficulty in a hybrid historic-cost and fair-value environment to fairly present the financial position of a firm that uses derivative contracts to hedge. The problem is determining how to account for positions when an instrument valued at fair value is hedging a financial position valued at historic cost. A substantial portion of FAS 133 is devoted to methods of reconciling the two valuations. The problem of fairly presenting a firm's position is eliminated, at least conceptually, if both the position and the hedge were recorded at their fair values.
24. We discuss the exception of nonconstant probability of default next. Changes in the market discount rate could also result in changes in economic values that are not recognized in accounting net income. However, the effect of GAAP's not recognizing the effect of unexpected changes in market interest rates on the values of fixed–interest rate loans affects many assets and liabilities and is not directly related to loan-loss accounting. We discuss this situation in the sidebar on page 32.
25. Bank supervisors have acknowledged existing GAAP, which mandates recording a loss only when it is probable, as the basis for regulatory accounting. However, they argue that the existing rules provide room for recognizing the inherent imprecision in loan-loss estimates, and they would prefer a bias towards overstating the allowance for loan loss (see AICPA 2003).

## Fair-Value Accounting for Interest Rate Changes

Loan-loss accounting recognizes that some borrowers will not fully honor the promised interest and principal payments. However, loan-loss accounting does not and is not intended to recognize changes in loan value due to interest rates. Thus, the question arises as to whether banks should recognize changes in the fair value of loans due to interest rate changes.

The biggest obstacle to fair-value accounting for loan losses is that of estimating cash flows. The bank has valuable information about future cash flows from most types of loans that cannot be readily verified by third parties. In contrast, the most important determinant of fair value for interest rate changes—changes in market interest rates—are routinely collected by several parties, including the Federal Reserve, and these data are widely available.

Although fair-value accounting to adjust asset values for interest rate changes does not face the same fundamental problem as loan-loss accounting, fair value for rate changes is not problem-free. One problem is that if this accounting were limited to financial assets, it would likely be misleading because banks may use both liabilities and derivative contracts to hedge interest rate exposures, such as funding long-term, fixed-rate loans with long-term, fixed-rate deposits. Requiring a bank to measure assets but not liabilities by adjusting for changes in interest rates would typically be less informative about the bank's financial condition than merely reporting assets and liabilities at historic cost. Moreover, many banks measure and hedge their interest rate exposure at the level of portfolios of financial claims or even at the bankwide level. Thus, fair-value accounting should apply the effect of rate changes to all interest rate–sensitive instruments in a bank's portfolio. Although doing so adds complexity, it does not raise insurmountable or even difficult conceptual problems for developing reliable fair-value measures.

More difficult problems with applying fair-value accounting for interest rate changes arise from estimating concurrent, related changes in the cash flows. Cash flows may change when many financial contracts offer options to one or both parties, where the value of the option depends in part on interest rate changes. For example, residential mortgage loan contracts often give the borrower the option to prepay the loan at no extra charge, an option that increases in value as interest rates decline. Another way in which cash flows may change is that banks may administer the rates charged on certain loans and paid on certain liabilities rather than allow the rates to change automatically in response to changes in market rates. An example of such a loan contract is a small-business loan for which the rate varies with changes in a bank-determined prime rate. We next discuss these two cases.

### Options Whose Value Depends in Part on Interest Rates

A bank's portfolio may contain both stand-alone options contracts and a variety of other types of financial contracts that embed options whose value depends in part on interest rates. The value of most stand-alone options can be reasonably reliably measured without any conceptual problems because market prices exist for many types of interest rate options contracts. Some contracts, such as options on Treasury securities, trade on exchanges while others, such as options on interest rate swaps, trade in over-the-counter markets where price quotes are frequently available. Moreover, even where market prices are unavailable, values can be calculated using models that depend solely on readily observable data.[1]

The more difficult options-related problem is that of options embedded in other contracts, such as loans that allow the borrower to prepay part or all of the loan principal at no charge. While the value of these options is largely determined by interest rates, interest rates are not the sole determinant of when the options are exercised. The holders of the options may rationally exercise (or fail to exercise) the options for reasons that are unrelated to interest rates. For example, mortgage borrowers may fail to exercise their option to prepay a mortgage loan and

refinance it with lower-cost debt because their credit quality has declined to the point where they cannot obtain another loan. Additionally, some borrowers, especially consumers, may fail to exercise options because, considering the cost of evaluating the situation and the transactions costs of refinancing, the perceived gain is simply too small to matter.

If an embedded interest rate option is a significant part of a financial instrument's value, market participants may be expected to develop sophisticated quantitative models to value the options. The values obtained from these models could be supplemented by judgmental adjustments based on factors outside the model. However, the starting point in estimating value will almost always come from a quantitative model, given the complexity of valuing interest rate options. The valuations obtained from these quantitative models could be used to provide reliable estimates of embedded options' value for financial accounting.

One potential problem with relying on model estimates is that from time to time managers will want to adopt a new model that they perceive is more accurate. If management could change models every period without disclosing the change, the reliability of the valuations could be compromised. However, informative disclosures about the changes in the models are possible precisely because the estimates come from quantitative models. The bank could be required to disclose that the model used to prepare the financial statements has changed, what values would have been reported under the prior model, the monetary effect of the changes, and the rationale for the change if it has a material impact on loan values.

### Administered Rates

The rate paid on a substantial fraction of bank loans and deposits varies through time, with the rate determined by the bank and not directly by financial markets. The primary examples are loans to consumers and small businesses based on the prime rate, such as a rate of prime plus 1 percent. A bank may base its prime rate on the published industry average, but in many cases the bank's prime rate is whatever the bank says it is. Similarly, the bank sets the rate it will pay on deposits held in the form of negotiable-orders-of-withdrawal accounts (interest-bearing checking accounts) and money market deposit accounts. The bank may set these rates at whatever level it deems appropriate.

One way of valuing these accounts is to estimate expected future interest rates and take account of the potential gain to the bank from its administration of the rate charged on loans and paid on deposits. However, deposit and loan accounts with such floating-rate features typically also provide the bank's customers with the option of withdrawing their deposits or repaying their loan at par. Thus, at any given time the par value of the loan or deposit reflects an implicit market transaction, wherein the bank sets the rate at a level it finds acceptable and customers indicate their willingness to accept the rate by maintaining the deposit or loan relationship. Moreover, this continual implicit recontracting suggests that the bank earns economic rents not as a result of the bank's and customers' initial decisions to enter into the contracts or as a result of likely future rate changes. Rather, the bank earns its rents in each period as customers maintain their loans or deposits with the bank even though the difference between the bank's administered rate and market rates on comparable instruments may have moved in ways that are favorable to the bank. This analysis suggests that loans and deposits with administered rates should be valued at par, with any gains or losses due to rate changes recognized as they are realized.

---

1. Models to value interest rate options are not free from error, but their consistent application can yield valuations that tend to reduce opportunities for earnings manipulation.

portfolio without recognizing the likely increase in loan losses in future periods could result in an overstatement of the bank's current reported profitability and of the value of equity (retained earnings). If banks are allowed to recognize interest intended to cover future loan losses as current income, the bank can experience greater growth while remaining within the supervisor's capital adequacy standards, and the book value of equity will overstate the amount of equity available to absorb losses that may emerge in the loan portfolio.

*A market for the full transfer of credit risk does not exist because banks place a greater value on the loan than potential buyers are willing to pay.*

Supervisors may also be concerned about the GAAP limitation of information that can be used to determine the probability and amount of default. FAS 5 specifies that a loss expense may be recorded only if it is "probable that one or more future events *will* occur confirming the fact of the loss" (paragraph 8b, emphasis added). Losses that result from a possible (less likely than not) change in macroeconomic conditions from expansion to recession are not accounted for. GAAP does not take account of that risk.

Thus, bank supervisors have some legitimate concerns that should be addressed. To the extent that their goal is merely to make loan-loss accounting more conservative, however, the supervisors' use of loan-loss accounting is inappropriate and, from a bank-safety perspective, ineffective. How a bank accounts for loan losses does not change the cash flow it receives from loans, which is determined by borrowers' payments. Recognizing larger expected loan losses in financial statements does not give banks added resources to absorb losses. The loan-loss provision and allowance are merely entries in financial statements. If bank supervisors want more capital relative to assets, they should require it.

If loan-loss accounting is to influence the viability of a bank, it must be through its influence on the bank's investment, funding, and dividend policies. Higher reported loan losses will indeed reduce a bank's reported earnings and equity capital. This reduction may induce a bank to undertake some combination of issuing new equity capital, reducing its dividends, and reducing the growth rate of its risky assets—conservative actions that it otherwise might not undertake. Each of these measures will reduce the probability that a bank will become insolvent. But it is the changes in dividends, equity issuance, or investment policies, and not the change in reported loan losses, that reduce the risk of insolvency.

Nevertheless, if supervisors believed that loan-loss accounting were a more efficient method of obtaining their supervisory goals, they could replace GAAP with the more stringent RAP accounting for loan losses. However, this substitution would be a poor way to deal with this problem because the difference between RAP and GAAP could create confusion for investors. Furthermore, the problem could be dealt with more effectively by adjusting bank capital requirements. Bank supervisors could require risky banks to increase their equity (by reducing dividends or issuing capital) and demand a reduction in the growth rate of risky assets through capital adequacy requirements. Thus, the principal safety and soundness benefits that the supervisors might derive indirectly from higher GAAP loan-loss allowances could be obtained directly with their existing powers.[26]

Bank supervisors and others might object that, absent an allowance for loan losses and a provision for estimated loan-loss expense, relatively more and less risky loans will appear to be the same on the balance sheet. But this outcome merely reflects economic reality, as the following example illustrates.

Consider two investments that a bank might make—a $100,000 risk-free government bond or a $100,000 risky loan, each of which matures in five years. Each promises quarterly interest payments, with the principal repaid at maturity. The interest rates on the bond and the loan are 4 percent and 12 percent, respectively. (For purposes of this illustration, assume that the return on capital is included in operating expenses, no income taxes exist, and no operating expenses are required for the bond investment.) On the loan, 5 percent of the yield is compensation to the bank for operating expenses. Thus, the bank gets an additional 3 percent over the bond rate and operating expenses to compensate it for expected losses given default (assuming the market is perfectly competitive, the bank is risk neutral, and no associated additional benefits or costs exist). Through years one, two, and three the loan does not default and the expectation that the borrower might default does not change.

Hence, in each of these years, if the bank invested in the government bond, it would record $4,000 as interest income. For the loan, it would record interest income of $12,000 less $5,000 of operating expenses, or a net of $7,000, in each year. At the end of the first three years, the bond and the loan have the same economic and accounting value—$100,000. But in the middle of year four the borrower declares bankruptcy and the loan defaults, and the bank can recover, say, only $40,000 after collection costs. If it had invested in the government bond, it would have recorded in that year and the following year $4,000 as interest income. But, because it invested in the loan, in year four it records $6,000 as interest income, $2,500 in operating expenses (for a half year), and $60,000 in net loan losses, or a reduction in net income of $56,500. Thus, the two net income streams from these assets are very different. The government bond earns $4,000 a year over the five years. The annual earnings (losses) from the loan are $7,000, $7,000, $7,000, $(56,500), and $0 for the loan. However, assuming the bank collects on the loan at the start of year five, it will have $40,000 in the last year with which it can invest in an interest-earning asset, reduce its liabilities, or both.

A bank supervisor's position might be that the bank should have increased its allowance for loan losses for the possible loss by reducing its interest income by, say, $3,000 a year and putting this amount into a contra-asset account (allowance for loan losses). A securities regulator might say that this accounting procedure would warn readers of the bank's financial statement that the loan might default. And when default did occur, the accounting would soften the blow in year four by $10,500, the accumulated amount in the allowance for loan losses through mid–year four. This procedure would be bad accounting, however, because it would misrepresent objectively determined economic reality and would be a very poor way to warn financial statement readers of the loan risk. First, the economic value of the risky loan and riskless bond are equal at the end of the first, second, and third years.[27] This situation exists because the risky loan promises a higher return (risk premium) that compensates the bank for the additional possible (expected) cost of default. If the loan does not default in a particular period, then the bank has earned the risk premium for that period. Second, should the loan default, the bank suffers the loss at that time, not earlier. Third, the variability of the loan's income stream is a reality that should not be hidden, but its variance can be reduced if the bank holds a portfolio of loans such that about the same percentage defaults in any one year.

---

26. See Wall and Koch (2000) for a further discussion of this issue.
27. The value of the risky loan remains at par as the credit risk premium earned in each year exactly equals the required premium to cover the possibility that the loan would default in that year.

## Synthesis and Policy Conclusions

Useful financial information is both relevant and reliable, and these qualities are interrelated. We have previously shown that GAAP based on historical data, which initially are derived from market transactions, emphasizes the reliability of reported data. Economic-value estimates, by contrast, could be more relevant for investment and supervisory decisions, but only if the valuations can be trusted, since the estimates are subject to managerial discretion and manipulation.

*If loan-loss accounting is to influence the viability of a bank, it must be through its influence on the bank's investment, funding, and dividend policies.*

The FASB has begun a process that may lead to loans being carried on banks' books at their fair, or estimated market, value, which would incorporate all available public information about a loan's value to a bank. This change could produce more relevant loan valuations if these valuations can be validated by independent public accountants. The problem is that estimates of market value by management are required because most types of bank loans either do not trade in liquid markets or trade only to the extent the seller (explicitly or implicitly) retains a substantial part of the credit risk. However, such estimates may not be reliable, given the many incentives management may have to adjust loan valuations so that the bank can attain its target equity capital and earnings. Hence, the FASB's move toward replacing historical costs with fair values is likely to result in banks (and other lenders) reporting untrustworthy numbers that also fall short of the desired value-in-use numbers.

Bank supervisors are important users of financial statements. Unlike investors, though, they have an asymmetric loss function: Overstated loan values may increase the probability of a bank's failing and increase the loss to the FDIC in the event of failures, but understated loan values impose no cost to supervisors. Thus, supervisors would prefer that loans be valued at the lower end of the range of reasonable estimates calculated using all available information (net of costs). However, supervisors do have other mechanisms that may be used to reduce the probability of a bank failure due to loan losses, such as requiring banks to hold sufficient capital to absorb those losses.

Our analysis of present GAAP, the proposed moves to fair-value accounting, and supervisors' preferred approach reveals significant weaknesses in loan valuation and accounting. However, such accounting difficulties are not unique to loans. Indeed, the economic value to an enterprise of many assets, such as buildings, equipment, and intangibles, rarely can be reliably measured. The approach that has been taken in accounting is to base financial statements on numbers that can be reliably measured without creating an opportunity for the reporting firm's management to overstate earnings. One important qualification is that economic values not based on verifiable actual market transactions do not replace historic costs unless the economic values are lower. This approach limits management's opportunities to report inflated values for their firm's earnings and assets but also provides for disclosure of the more relevant fair values when economic value has declined.[28]

We conclude that using the lower of historic cost or economic value for valuing the credit risk of loans is the most appropriate procedure for both investors and bank supervisors. In most cases this approach would result in values similar to those currently required by GAAP. However, using economic value for loans when these are less than historic cost would change the criteria for recognizing loan losses in several important ways. First, this procedure would require use of the full range of increases in the probability of default, even when a loss is still less than probable. Second, the

procedure would also recognize that reported income might be overstated for rapidly growing loan portfolios for which the hazard rate is expected to increase rather than remain constant over time. Although the traditional GAAP matching concept requires deferral of a portion of interest income during those periods when expected probabilities of default are low until the periods in which the probabilities of default are expected to increase, this deferral rarely is made (perhaps because this economic property of loans has not been recognized or has been neglected).

While the application of the lower of historic cost or economic value to loans would prevent managers from accelerating the reporting of uncertain future earnings to the current period, it could provide them with a tool for deferring current income to future periods. Managers who sought to defer income, perhaps to smooth reported income, could adjust downward their estimates of individual loans' fair value to reduce current income. The underestimate of individual loan values could then be reversed in a subsequent period when the manager sought to report higher income. One mechanism for reducing managers' ability to manipulate reported net income would be to apply the lower-of-cost-or-economic-value rule to portfolios of similar loans rather than to individual loans. If this rule were applied at the portfolio level, the manager would have to make the case that the economic value of the entire portfolio was below its historic-cost value, a more difficult task to the extent that new loans have a positive net present value and some loans have appreciated in value as a result of lower expected future losses. Furthermore, applying the lower-of-cost-or-economic-value rule to a portfolio of related loans rather than to the entire loan portfolio is appropriate because banks report disaggregated results by portfolios of similar loans. Because applying this rule to loan portfolios is likely to understate (and certainly does not overstate) loan amounts, bank supervisors should be pleased.

Valuation of loan portfolios net of loan losses using a lower-of-historic-cost-or-economic-value rule would provide the most relevant adequately reliable measure of loan value. Unlike current GAAP, this rule does not require a reduction in loan portfolios' value when their value is already understated by historic cost. Unlike fair value, our recommended valuation method is less reliant on prices or estimates that are likely to be systematically biased. Finally, unlike the bank supervisors' preferred position, our rule would lead to loans that are less understated and would provide procedures that are less likely to be abused to facilitate earnings management. Supervisors may object that higher loan-loss allowances provide a valuable cushion to absorb unexpected losses and reduce the probability of bank failure. If they believe that banks would be undercapitalized without loan-loss reserves, supervisors have other tools to force banks to hold more capital. Indeed, capital serves to protect creditors (particularly depositors and the FDIC) from all sorts of losses, including losses on securities, real estate, derivatives, foreign exchange, and operations. Loan-loss accounting, therefore, should return to its original function (providing useful information to investors) and not be unnecessarily distorted to accomplish other goals.

28. The FASB's move towards requiring fair-value accounting based on estimates, appraisals, and pricing models, if adopted, would alter the traditional approach.

## REFERENCES

American Institute of Certified Public Accountants (AICPA). 2003. In Re: Proposed statement of position—allowance for credit losses (SOP), October 6. <www.aicpa.org/download/acctstd/cmtltr_recd/clsl/cl279.pdf> (December 2, 2005).

Benston, George J. 2005. Fair-value accounting: A cautionary tale from Enron. Emory University, Goizueta Business School, unpublished paper, March.

Davenport, Todd. 2004. A switch to FASB keeps loan-loss issue on track. *American Banker* 169, no. 125: 1, 4, 8.

Federal Deposit Insurance Corporation (FDIC). Office of Inspector General. 2000. *Material loss review—the failure of Pacific Thrift and Loan Company, Woodland Hills, California*. Audit Report No. 00-022, June 7. <www.fdicig.gov/reports00/00-022.pdf> (December 2, 2005).

Financial Accounting Standards Board (FASB). 1999. Preliminary views on major issues related to reporting financial instruments and certain related assets and liabilities at fair value. Financial Accounting Series No. 204-B, December 14. <www.fasb.org/draft/pvfvalu1.pdf> (December 2, 2005).

Frame, Scott W. 2003. Federal Home Loan Bank mortgage purchases: Implications for mortgage markets. Federal Reserve Bank of Atlanta *Economic Review* 88, no. 3:17–31.

Wall, Larry D., and Timothy W. Koch. 2000. Bank loan-loss accounting: A review of theoretical and empirical evidence. Federal Reserve Bank of Atlanta *Economic Review* 85, no. 2:1–19.

Walter, John R. 1991. Loan loss reserves. Federal Reserve Bank of Richmond *Economic Review* 77, no. 4:20–30.

# Macroeconomic Models with Heterogeneous Agents and Housing

**KARSTEN JESKE**
*The author is a research economist and assistant policy adviser in the macropolicy group of the Atlanta Fed's research department.* *He thanks William Roberds and Ellis Tallman for helpful comments.*

It is hard to miss the important role the housing sector plays in the macroeconomy. In 2004 the value of real estate held in household portfolios amounted to over $17 trillion, or 143 percent of annual nominal gross domestic product (Board of Governors 2005). This amount is larger than all corporate equity held directly, through mutual funds and pension funds combined. Closely linked to housing is the mortgage market, which now totals over $7.5 trillion. On the average household's balance sheet, a home mortgage is the largest item on the liability side, easily dwarfing consumer credit, which includes credit card debt and car loans, by a factor of almost four to one.

Macroeconomists, therefore, have spent considerable effort on incorporating housing and housing finance into macroeconomic models. In understanding the effects that housing and the various related government policies have on households in particular and the economy in general, recognizing the economic and demographic diversity among households is crucial. Net worth, housing value (gross and net of mortgages), and income vary substantially among households not only by age but within age groups. Macroeconomists thus employ models—called heterogeneous agent models—that can accommodate this diversity.

This article is a progress report on where this line of research currently stands. At the Federal Reserve Bank of Atlanta conference "Housing, Mortgage Finance, and the Macroeconomy" in May 2005, several papers shared a common framework: namely, using macro models with heterogeneous agents.[1] This article discusses four of the conference papers, the issues the researchers explored, the progress they made, and the challenges that lie ahead.

The paper by Fang Yang (2005) deals with the life-cycle behavior of housing versus nonhousing consumption. Her objective is to build a model to account for two peculiar features of consumption over the life cycle.

The paper by Wenli Li and Rui Yao (2005) studies the effect of house price changes on the macroeconomy. They point out that while house price changes may

Table 1
**Homeownership Rates by Age Group**

| Age group | Homeownership rates (percent) |
|---|---|
| 0–29 | 30.1 |
| 30–39 | 58.4 |
| 40–49 | 74.0 |
| 50–59 | 79.8 |
| 60–69 | 81.7 |
| 70+ | 79.3 |

Note: The unconditional mean across all ages is 67.7 percent.

Source: Board of Governors of the Federal Reserve System, Survey of Consumer Finances 2001

not have a huge effect on the economy as a whole, different households—in particular those of different age and asset holdings—will be affected very differently by house price changes. An economic model with heterogeneous agents is therefore necessary to measure the effects a price increase might have on agents with different incomes and asset holdings.

The paper by Matthew Chambers, Carlos Garriga, and Donald Schlagenhauf (2005b) asks how different households are affected by the availability of different mortgage contracts: most importantly, which mortgage contract is most successful in allowing younger households (that tend to have less savings and thus might have trouble coming up with the necessary down payment) to purchase a house.

The paper by Karsten Jeske and Dirk Krueger (2005) addresses the important policy question of whether it is desirable to subsidize mortgage interest rates. One important subsidy is the implicit federal guarantee for government-sponsored enterprises (GSEs). Jeske and Krueger examine the distributional effect of such a subsidy. They find that mostly high-income and high-net-worth households benefit while, both in terms of welfare and homeownership, low-income and low-net-worth households will not benefit at all.

## Microlevel Data on Housing and Mortgage Finance

Trying to answer the questions mentioned above requires models that are detailed enough to accommodate a housing market and, most importantly, heterogeneity among households. This heterogeneity is the crucial ingredient of any model attempting to determine the distributional effects of housing policy. Before examining the models, I first introduce some empirical facts coming out of microlevel data to demonstrate how the aggregate numbers on income, real estate, and mortgage debt are distributed across the population. Aggregate real estate values are approximately 143 percent of aggregate annual income. But income, real estate wealth, mortgage debt, and net worth are very unevenly distributed both across and within age groups. The data set used in this article's computations is the Survey of Consumer Finances (from the Federal Reserve Board) for the year 2001, which, among other things, collects data on consumer income and balance sheets.

Table 1 shows homeownership rates by age group compared to the unconditional homeownership rate (that is, over all age groups) of 67 percent. The homeownership rate is the lowest, at about 30 percent, for households headed by persons below the age of twenty-nine. The rate increases steeply over the next two age cohorts, peaking at more than 80 percent for the sixty to sixty-nine age group and moderating slightly after that. Not all homeowners are equal, as Table 2 shows. Not only do fewer young households own a home, but their homes tend to be smaller than those of older households. The average homeowner increases real estate holdings from about $100,000 until reaching his or her fifties, when he or she holds about $216,000 in real estate, and then gradually downsizes the value of the primary residence. Mortgage debt also displays a hump-shaped profile, though the peak occurs in the thirty to thirty-

nine age group. This pattern seems to be consistent with the following scenario: Households upgrade the size and obviously the value of the house they are living in. Early in life most of the upgrading is financed with larger mortgages. Homeowners aged forty and older then tend to finance upgrades out of savings rather than larger mortgages.

In homeownership rates, even within age groups there is substantial variation. The first panel of Figure 1 plots rates by income quintiles within each age group. The chart shows that homeownership is associated with high income, with households in the bottom quintile having below-average ownership rates across all age groups. The second panel of Figure 1 yields the same qualitative pattern for net worth, but the differences between households with low and high net worth households are even more severe. Households in the bottom quintile have only a 30 percent homeownership rate even in the sixty to sixty-nine age group. In contrast, for all ages after forty, the second through fifth quintiles display homeownership rates above the national average of 67 percent.[2]

Table 2
**Home Values and Mortgage Debt by Homeowners' Age Group**

| Age group | Value of primary residence | Mortgage on primary residence |
|---|---|---|
| 0–29 | 103.6 | 65.5 |
| 30–39 | 157.1 | 91.4 |
| 40–49 | 190.0 | 82.1 |
| 50–59 | 216.4 | 68.5 |
| 60–69 | 193.5 | 35.4 |
| 70+ | 171.7 | 11.2 |

Note: Values and debt shown in thousands of dollars.

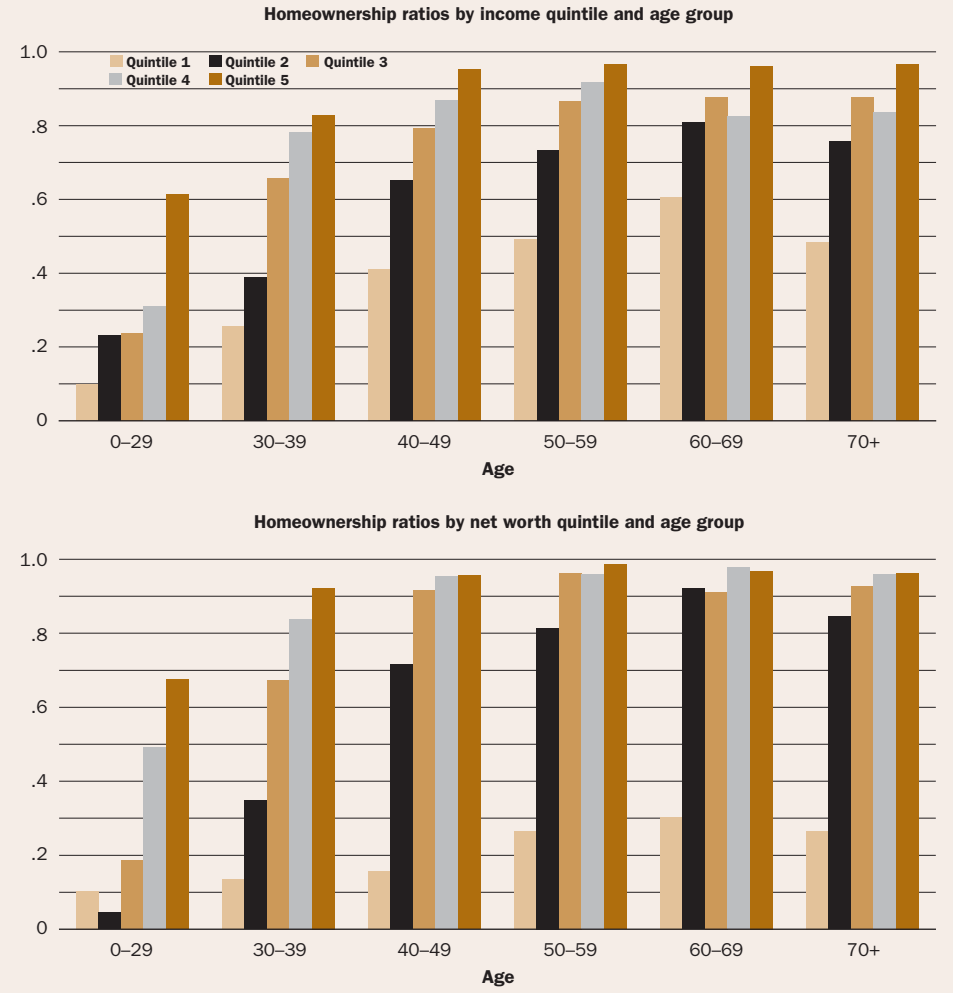
Source: Board of Governors of the Federal Reserve System, Survey of Consumer Finances 2001

Vast differences also exist between homeowners and renters. The first panel of Figure 2 plots median income by age group for the whole population and for homeowners versus renters. For the population as a whole, household income increases from just below $30,000 for the youngest age cohort to almost $60,000 for fifty to fifty-nine age group, after which it declines to about $25,000 for the seventy and over age group. Homeowners tend to have much larger median incomes than renters; the median homeowner makes roughly twice as much as the median renter. Moreover, the hump in the income process is more pronounced for homeowners. The same general pattern holds for average incomes (see the second panel of Figure 2).

The contrasts between homeowners and renters are even more obvious in terms of net worth positions. The third and fourth panels of Figure 2 plot median and average net worth, respectively, by age group. Households accumulate savings during their working years, moreso than does the population as a whole. The median homeowner accumulates more than $200,000 during his or her lifetime. Because the net worth distribution is so skewed, average net worth is considerably larger: The average homeowner accumulates more than $800,000 over a lifetime. The median renter, however, has negligible net worth—consistently under $10,000—for all age groups. In fact, 50 percent of all renters have essentially zero net worth. Even average net worth for renters is below $100,000 in all age groups.

1. The four papers introduced here as well as the program and the remaining papers are available on the Atlanta Fed's Web site at www.frbatlanta.org under "News & Events," "Conferences," "2005."
2. The cutoff for the lowest income quintile is $15,400 in the youngest age group, increases to $26,700 for the fifty to fifty-nine age group, and then declines to $11,300 for households aged seventy and above. The cutoff for net worth is zero for the youngest households and increases to $39,100 for households aged seventy and above.

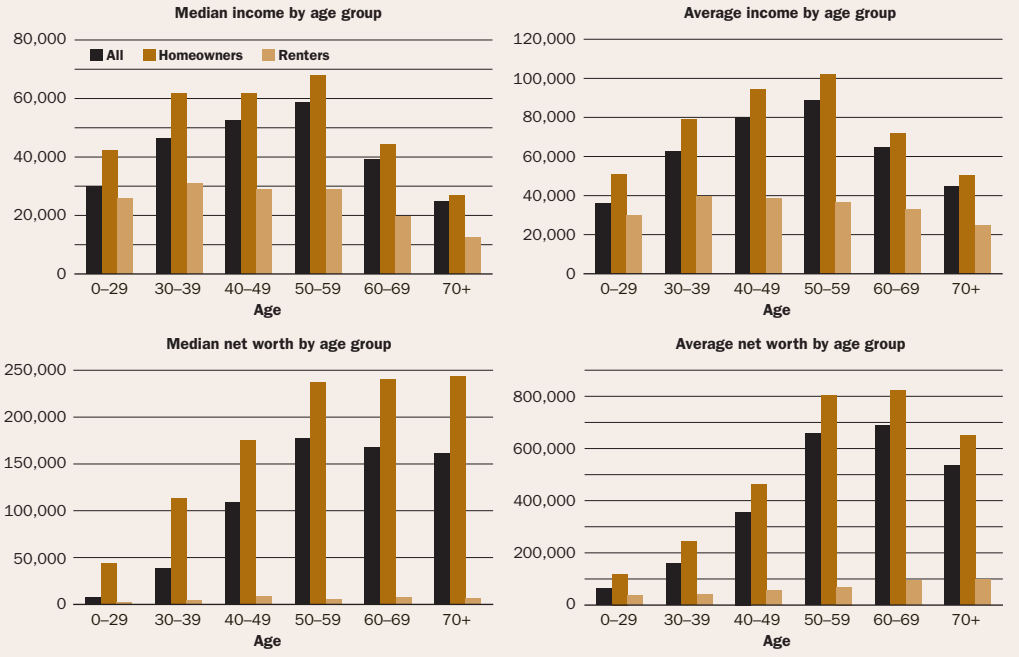
Figure 1
**Homeownership Patterns by Age, Income, and Net Worth**




Source: Board of Governors of the Federal Reserve System, Survey of Consumer Finances 2001

In summary, households differ substantially, both between and within age groups, in their net worth positions and asset allocations. As one would expect, when households are young and middle-aged they accumulate savings for retirement, and one form of savings is real estate. Even within age groups, there is considerable heterogeneity of households. Incomes and especially net worth vary substantially across age groups and between renters and homeowners. When trying to answer questions such as "What is the effect of increasing house prices?" or "Should we subsidize mortgage interest rates?" one should take into account that different households will be affected very differently by changes in house prices or government policies. An increase in house prices might be beneficial to existing homeowners, but renters may not be affected at all or, even worse, might suffer if rental rates increase. Likewise, subsidizing mortgage interest—for example, through mortgage interest tax-deductibility or

Figure 2
**Income and Net Worth by Age Group**




Source: Board of Governors of the Federal Reserve System, Survey of Consumer Finances 2001

government guarantees on mortgage lenders—might mostly benefit existing homeowners, who are well off already because they tend to have above-average income and net worth. This consideration reinforces the importance of studying models with heterogeneous agents to account for the effects that policy has on agents with different demographic and economic backgrounds.

## A Generic Model with Housing

A key ingredient in the model is the life-cycle pattern, which implies two important features the model must include. First, it must include the hump-shaped earnings profile that the average household experiences, as shown in Figure 2. Second, the model must generate a realistic life span, which means households of different age groups face different mortality risk. Around the trend path of life-cycle earnings, a household receives an uncertain stream of income over its life cycle and then tries to maximize its lifetime discounted utility subject to a budget constraint.

A household receives utility from consumption of both nonhousing goods $c$ and housing goods $h$. Income comes both from labor and investment income. The maximization problem of a household of age $t$ in recursive form is then defined in the following way: The value of starting period $t$ with assets $a$ and labor productivity $y$ is the maximum value of current utility plus expected discounted period $t + 1$ value as a function of future assets and future labor productivity.

$$V_t(a,y) = \max_{c,h,b'} u(c,h) + \rho_t \beta E V_{t+1}(a', y'),$$

subject to a number of constraints:

$$c + h + b' = wy + a;$$
$$a' = (1 - \delta^h)h + (1 + r)b';$$
$$b' \geq -\bar{B};$$
$$c, h \geq 0.$$

In the objective function I use a discount factor $\beta$ common to all age groups as well as additional discounting in the form of $\rho_t$, which is the conditional survival probability that depends on age. The first constraint is the budget constraint. Labor income measured as wage rate $w$ times labor productivity $y$ plus assets $a$ can be spent on nonhousing consumption $c$, housing $h$, and other savings $b'$. The second constraint specifies the next period's assets $a'$ as the sum of housing stock net of depreciation and maintenance cost $(1 - \delta^h)h$ and other savings times the gross interest rate $(1 + r)b'$. The third condition is a borrowing constraint that specifies a lower bound on asset holdings—in other words, an upper bound on borrowing. Finally, the fourth condition states that both types of consumption have to be nonnegative. This model closely resembles that of Aiyagari (1994), with some distinguishing features: two types of consumption, a very particular borrowing constraint, and a life-cycle component.[3]

Denote $u_c$ and $u_h$ the derivatives of the utility function with respect to nonhousing consumption $c$ and housing $h$, respectively. Assuming for now that the borrowing constraint does not bind, we can immediately derive two optimality conditions, also called Euler equations:

$$(1)\quad 1 = \rho_t \beta E \frac{u_c(c', h')}{u_c(c, h)}(1 + r), \text{ and}$$

$$(2)\quad 1 = \rho_t \beta E \frac{u_c(c', h')}{u_c(c, h)}(1 - \delta^h) + \frac{u_h(c, h)}{u_c(c, h)}.$$

The first condition is completely standard: The bond return, weighted by the intertemporal marginal rate of substitution (also called the pricing kernel) and discounted by $\rho_t\beta$, equals one. The second condition has a similar structure but also incorporates today's marginal rate of substitution between housing and nonhousing consumption. Both conditions guarantee optimality by ensuring that no utility-improving substitution exists between either consumption and savings or consumption and housing. Next, combine the two Euler equations into

$$\frac{u_h(c, h)}{u_c(c, h)} = \frac{r + \delta^h}{1 + r}.$$

This condition states that the marginal rate of substitution between consumption and housing equals the cost of one unit of housing computed as the depreciation rate plus the opportunity cost of housing measured as the bond rate $r$. I also divide by $1 + r$ because the payoff from housing is in the current period while the bond pays off in the next period. The right-hand side is also called the user-cost formula for

housing, and if there were a rental market in this model, $(r + \delta^h)/(1 + r)$ would be the rental rate per unit.

To put some structure into the model, assume that the utility function takes the following form:

$$u(c,h)=\frac{\left[\theta c^{\eta}+(1-\theta)h^{\eta}\right]^{\frac{1-\sigma}{\eta}}-1}{1-\sigma}.$$

This is the form normally used in the literature. The parameter $\sigma$ is the risk aversion coefficient, $\eta$ determines the substitutability between housing and nonhousing consumption, and $\theta$ is a weighting parameter for how much the household values nonhousing consumption. How good is this model in accounting for the observed data? From the model, one can deduce two implications. First, the user cost formula for housing can be written as

$$\frac{h}{c}=\left(\frac{r+\delta^h}{1+r}\frac{1-\theta}{\theta}\right)^{\frac{1}{\eta-1}},$$

that is, housing over nonhousing consumption stays constant over the life cycle. Next, from the first Euler condition (equation [1]) and the fact that the *h/c* ratio stays constant over time one deduces that

$$E\left(\frac{c'}{c}\right)^{-\sigma}=\frac{1}{\rho_t\beta(1+r)}.$$

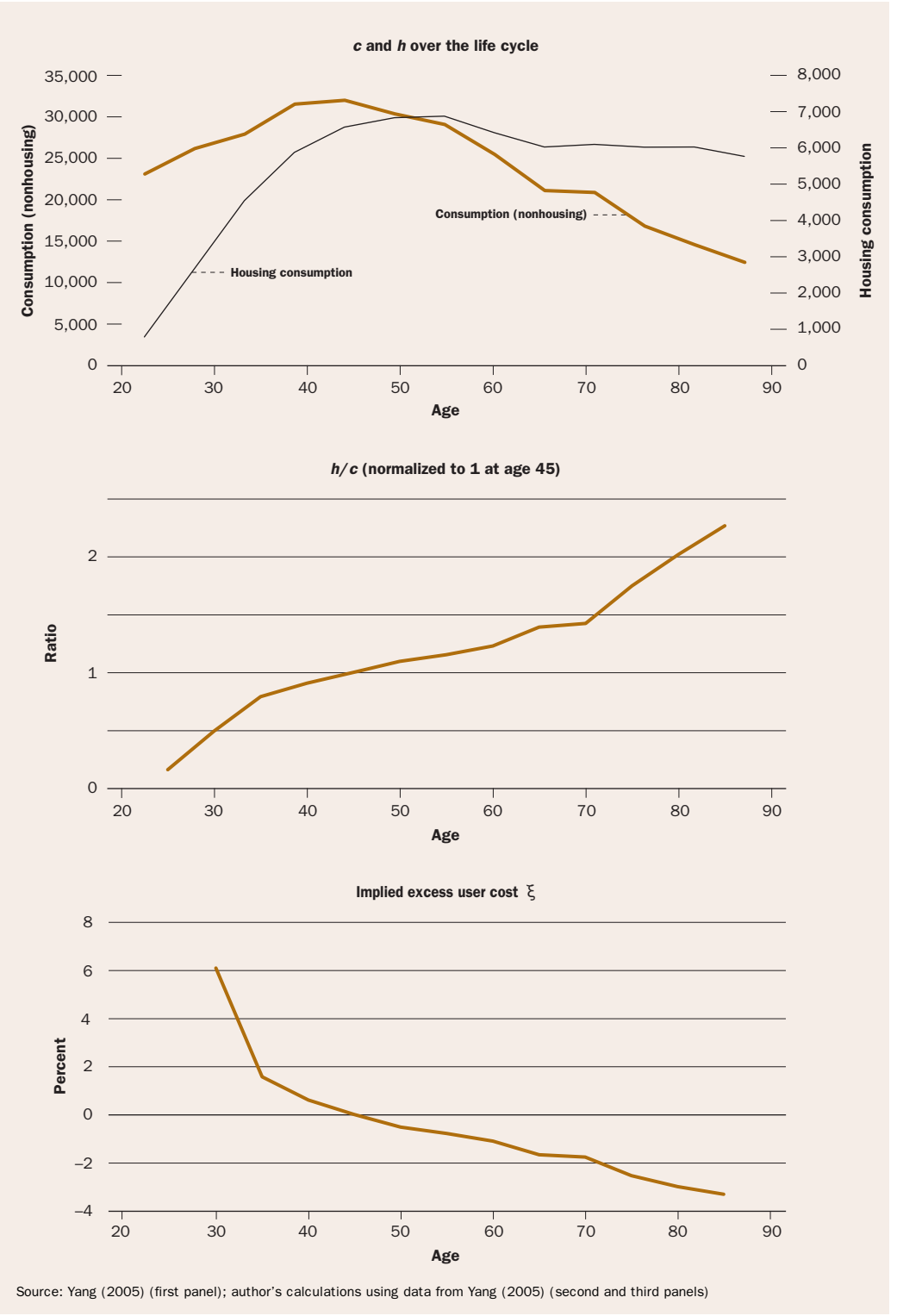
Normally it will be the case that $\beta(1 + r) < 1$.[4] Since the survival probability $\rho_t$ is less than one, the right-hand side is greater than one. Since $\sigma > 0$, the second implication of the model has now been shown: Without borrowing constraints, consumption decreases over the life cycle, at least in expected terms. This result makes perfect sense. According to the permanent income hypothesis, the mean path of earnings is irrelevant in the absence of borrowing constraints; all that counts is the discounted expected path of earnings. If the interest rate is lower than the rate of discounting, households will choose a decreasing consumption path though there may be fluctuations around this decrease because of changes in labor productivity.

## Life-Cycle Patterns of Housing Consumption

The four conference papers mentioned in the introduction are all extensions of this same basic model. Yang (2005) observes that both implications of the model,

3. Note that the timing convention here is slightly different from that in the Yang (2005) paper. Specifically, I assume here that the housing choice *h* enters today's utility instead of tomorrow's. This assumption makes the problem slightly easier to handle because agents pick housing goods for the same period when their income uncertainty is revealed. Therefore, housing is a riskless asset, which makes the optimality conditions slightly more tractable. Jeske and Krueger (2005), for instance, use this timing convention. In their paper housing investment will nevertheless be risky because of uncertain depreciation. In the Yang paper, on the other hand, households pick *h* for the following period when the future productivity shock $y'$ is not known yet, a choice that makes housing a risky asset.
4. Economists usually pick the discount factor $\beta$ so as to match the model's capital-to-output ratio to that in the data. One can show that for the realistic range of the remaining parameters of the model the equilibrium interest rate *r* implied by this capital-to-output ratio satisfies $\beta(1 + r) < 1$.

Figure 3
**Consumption over the Life Cycle**




Source: Yang (2005) (first panel); author's calculations using data from Yang (2005) (second and third panels)

a decreasing consumption profile and a constant housing versus nonhousing consumption ratio, are at odds with the data. The first panel of Figure 3 shows average household consumption on both housing and nonhousing goods by age group (of the head of household). Notice that nonhousing consumption increases from about $23,000 per year for households aged twenty to an average of $32,000 per year for age forty-five and then decreases after that.[5]

Contrast this observation to the behavior of housing consumption, which displays an even sharper increase up to about age fifty-five but then does not drop as sharply as nonhousing consumption. Not surprisingly, then, the *h/c* ratio is not constant but rather increases over the life cycle, as shown in the second panel of Figure 3.

Can the increase in the *h/c* ratio be accounted for without radical changes to the model? One could imagine that younger agents face a higher depreciation rate of housing or a higher interest rate.[6] Formally, assume that the consumption ratio at age $t$ now takes the following shape:

$$(3)\quad \frac{h_t}{c_t} = \left[ \left( \frac{r+\delta^h}{1+r} + \xi_t \right) \frac{1-\theta}{\theta} \right]^{\frac{1}{\eta-1}},$$

where $\xi_t$ generates differences in households' subjective user cost because of differences in depreciation and interest rates during the life cycle. Is it possible to quantitatively match the observed ratio with realistic values of $\xi_t$? How positive would $\xi_t$ have to be for young agents and how negative for old agents in order to generate the *h/c* function observed in the data? To compute this cost, I normalize the implicit cost of a forty-five-year-old household to zero and back out the $\xi_t$ from equation (3) using the observed *h/c* ratios. The third panel of Figure 3 shows this implicit cost $\xi_t$ over the life cycle. To account for the observed consumption patterns, older households would have to find housing around 3 percentage points cheaper than those age forty-five do, and, likewise, younger households would have to find housing 6 percentage points more expensive. It is difficult to justify differences of this magnitude—almost 10 percentage points difference between age eighty-five and age thirty—entirely through different borrowing costs or depreciation rates.

Yang's paper can be thought of as finding other explanations to generate this large difference in subjective user costs. Specifically, Yang introduces two crucial ingredients into the model to account for the features observed in the data—binding borrowing constraints and transaction costs—and shows that they can quantitatively account for the increase in the consumption ratio.[7] One can easily add these two features into the model above. First, assume that $\bar{B} = -(1-\gamma)h$, that is, the household can borrow only up to a maximum of a certain fraction of the house value. Parameter $\gamma$ can be viewed as the minimum down payment on a house. This constraint ensures that with a hump-shaped labor income profile (see the first and second panels of

5. Yang is not the first economist to point out this feature. Earlier work on the consumption profile over the life cycle includes Blundell, Browning, and Meghir (1994); Attanasio and Browning (1995); Attanasio and Weber (1995); and, more recently, Gourinchas and Parker (2002) and Fernandez-Villaverde and Krueger (2002, 2004).
6. For example, one could justify this assumption with the fact that households with a younger head tend to have a larger family size, and thus more persons per square foot cause a higher depreciation rate.
7. Yang introduces even more ingredients, such as a bequest motive and a social security system; however, to illustrate the key intuition, a borrowing constraint and transaction costs are sufficient.

Figure 2) agents will be constrained in their borrowing early in life. In other words, households cannot raise their consumption to the desired level derived from the permanent income model without violating their borrowing constraint. This restriction binds because income is low early in life.

Second, adjusting the level of housing incurs transaction costs. Assume that if a household had a housing level of $h_{old}$ last period and wants to change it to $h$ today, it has to pay a transaction cost of $\psi(h_{old}, h)$. Assume for now that the transaction cost satisfies the following conditions:

$$\psi(h_{old}, h) \geq 0;$$
$$\psi(h_{old}, h_{old}) = 0;$$
$$\frac{\partial^2 \psi\left(h_{old}, h\right)}{\partial h^2} \geq 0.$$

That is, the transaction cost function takes on nonnegative values, there is no cost of leaving the size of the house unchanged,[8] and the transaction cost function is convex.[9] In this economy the optimality condition between housing and nonhousing consumption becomes

$$(4)\quad \frac{u_h\left(c,h\right)}{u_c\left(c,h\right)} = \left(1 - \frac{\mu}{u_c\left(c,h\right)}\right)\frac{r+\delta^h}{1+r} + \frac{\mu}{u_c\left(c,h\right)}\gamma + \psi_2\left(h_{old},h\right) + \rho_t \beta E \frac{u_c\left(c',h'\right)}{u_c\left(c,h\right)}\psi_1\left(h,h'\right),$$

where $\mu$ is the Lagrange multiplier on the borrowing constraint and $\psi_1$ and $\psi_2$ are derivatives of the transaction cost function with respect to the first and second argument, respectively. This formula is the same one as before, linking the marginal rate of substitution between housing and nonhousing consumption with the cost of purchasing housing. Instead of the user-cost formula of housing, one now has a weighted average between the previous user cost $(r + \delta^h)/(1 + r)$ and the down payment ratio $\gamma$, where the weight on $\gamma$ is proportional to the Lagrange multiplier of the borrowing constraint.[10]

One can interpret the first two terms on the right-hand side of equation (4) as the subjective user cost of housing (net of transaction cost). The more the borrowing constraint binds, the higher the housing cost for the household. This result is intuitive: If the constraint binds, the only way to afford more housing is to reduce current consumption, which comes at a cost of $\gamma$ units of consumption for every unit of housing purchased. The remaining two terms come from the transaction cost function. They determine the marginal cost of housing changes—that is, adjusting housing consumption from $h_{old}$ to $h$ in the current period and adjusting from $h$ to $h'$ in the following period.

Using the same elasticity of substitution between the two consumption types as in the Yang paper—that is, $\eta = 0$—one then obtains the following consumption ratio:

$$(5)\quad \frac{h}{c} = \frac{1-\theta}{\theta}\left[\left(1 - \frac{\mu}{u_c\left(c,h\right)}\right)\frac{r+\delta^h}{1+r} + \frac{\mu}{u_c\left(c,h\right)}\gamma + \psi_2\left(h_{old},h\right) + \rho_t \beta E \frac{u_c\left(c',h'\right)}{u_c\left(c,h\right)}\psi_1\left(h,h'\right)\right]^{-1}.$$

One way to interpret Yang's paper is that the additional terms in the square brackets tend to decrease in age; thus, inversely, the consumption ratio increases in age. For young agents there are two channels: the borrowing constraint and the transaction cost. A binding borrowing constraint has the same effect as a higher interest rate.

Since the subjective user cost for young agents (before transaction costs) is higher than the cost $(r + \delta^h)/(1 + r)$ for unconstrained agents, the $h/c$ ratio will increase just as in the data because the subjective user cost decreases when the borrowing constraint is relaxed over the lifetime.

So far, the larger implicit housing cost causes young households to substitute away from housing and into nonhousing goods, even without any transaction costs. Transaction costs add to this effect. Since the borrowing constraint induces a hump-shaped profile for consumption (both housing and nonhousing), young agents have the tendency to increase their $h$. Consequently, the following will be true for any three consecutive housing consumption values: $h_{old} \leq h \leq h'$. In that case $\psi_2(h_{old}, h) \geq 0 \geq \psi_1(h, h')$. Since the second term is discounted by $\rho_h \beta E[u_c(c', h')/u_c(c, h)]$, one would expect the total effect to be positive. Both effects from borrowing constraints and transactions costs explain why for young agents $h/c$ can be lower than the ratio derived in the frictionless version above.

*In understanding the effects that housing and related government policies have on households and the economy, recognizing the economic and demographic diversity among households is crucial.*

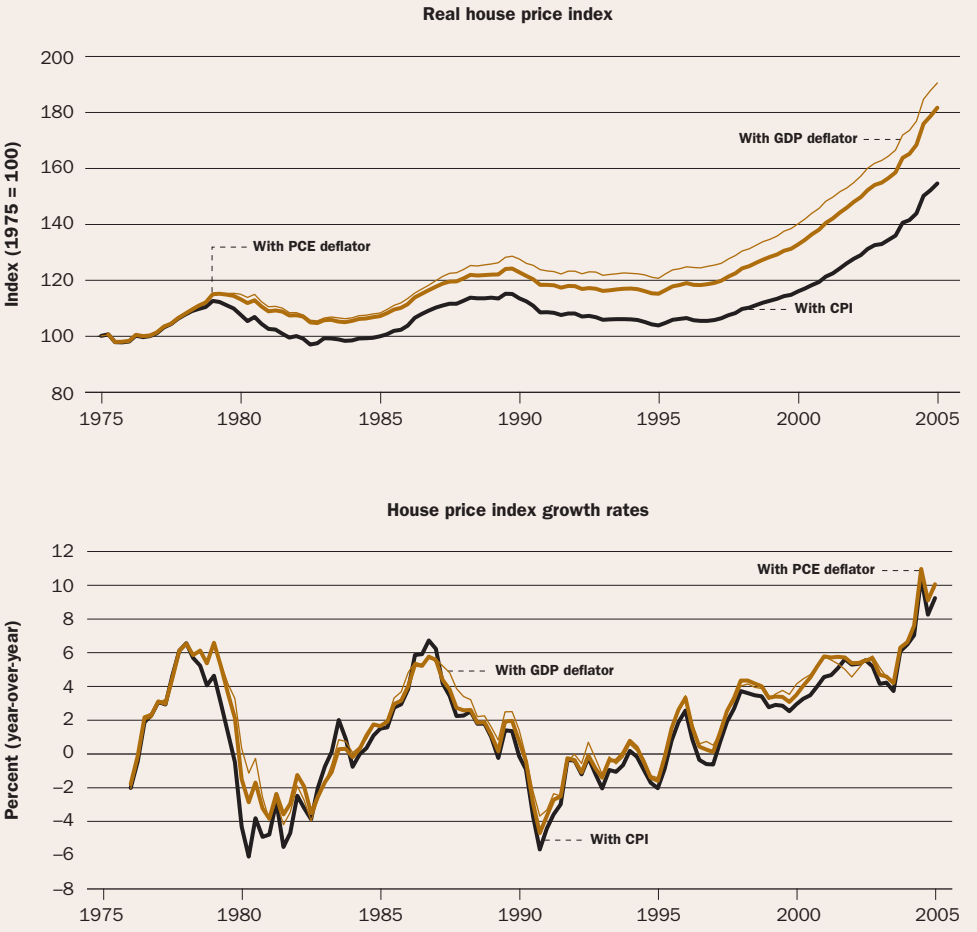
For older agents the borrowing constraint will likely no longer bind. Instead, all of the action comes from the transaction cost terms. The reverse of the argument for young agents applies. Older agents want to reduce consumption because of discounting due to lower survival probabilities. But picking $h_{old} \geq h \geq h'$ implies $\psi_2(h_{old}, h) \leq 0 \leq \psi_1(h, h')$. Again, the last term in equation (5) is subject to discounting, so the $\psi_2(h_{old}, h) \leq 0$ dominates, making the $h/c$ ratio higher than in a frictionless economy. The bottom line of the Yang (2005) paper is that by adding the two modifications the model matches the data qualitatively and quantitatively.

## The Effect of House Price Changes

Li and Yao (2005) study the effect of house price changes on the macroeconomy. The first panel of Figure 4 displays the U.S. house price index, adjusted for inflation by deflating the nominal house price index with three different commonly used measures of inflation: namely, the consumer price index (CPI) and the deflators for both gross domestic product (GDP) and personal consumption expenditures (PCE). Between 1975 and 1995 real house prices more or less stagnated if deflated by the CPI but moderately increased by about 20 percent over the twenty years if deflated by GDP or PCE.[11] Of course, the path of house prices was not monotonic. A decline in real house prices between 1979 and 1983 resulted from the high inflation rates in the late 1970s and the two recessions in the early 1980s. Likewise, the 1991 recession

8. Recall that in addition there is maintenance cost, but that enters the budget constraint separately.
9. The shape of transaction costs is slightly different from that used in Yang's paper, or most of the literature for that matter. Using the convex transaction cost, however, it is easier to explain the intuition using the Euler conditions because the necessary optimality conditions will also be sufficient.
10. Notice that in equilibrium it is always true that $\gamma > (r + \delta^h)/(1 + r)$. If this were not the case, one can show that housing investment is strictly more profitable than purchasing bonds, and nobody would ever hold positive amounts of bonds. This outcome cannot be an equilibrium outcome since the bond market has to clear—that is, bond savings have to equal mortgage lending.
11. This result makes perfect sense: GDP and PCE deflators tend to be slightly lower than the CPI because the CPI computes price changes of a fixed basket while in the GDP and PCE consumers obviously substitute into cheaper goods, thus lowering those two indexes.

Figure 4
**Consumption over the Life Cycle**




Source: Office of Federal Housing Enterprise Oversight, Bureau of Labor Statistics, and Bureau of Economic Analysis

took a bite out of house prices, and even during the first four years of the recovery, house prices stagnated.

After 1995, however, house prices staged an astonishing increase of almost 50 percent if deflated by CPI and 60 percent if adjusted by either of the deflators. Moreover, no decline occurred in house prices around the 2001 recession, and no stagnation occurred after the recession. On the contrary, as the second panel of Figure 4 shows, year-over-year growth rates in real house prices even accelerated during and after the recession to a record pace of 8 to 10 percent in 2004.

What is the effect of these house price fluctuations, especially the large run-up in prices since the mid 1990s? In a model without heterogeneity, a life-cycle earnings profile, and borrowing constraints, the effect on macroeconomic variables is exactly zero in the following sense: Imagine an economy has only one representative con-

sumer, earning exactly the average income and holding exactly the average house worth 143 percent of average income. An increase in house prices would have a positive wealth effect and a negative effect because housing in the present and future becomes more expensive. One can show, however, that the two effects precisely cancel each other out, so relative price changes between $c$ and $h$ are neutral. Utilizing the earlier model, one can formalize this intuition. Imagine one unit of housing now costs $P_h$ units of the consumption good. Abstracting from transaction costs, the optimality condition for housing versus consumption now becomes

$$(6)\quad \frac{u_h(c,h)}{u_c(c,h)} = \left(1 - \frac{\mu}{u_c(c,h)}\right)\left(1 - \frac{P_h'}{P_h}\frac{1-\delta^h}{1+r}\right) + \frac{\mu}{u_c(c,h)}\gamma.$$

This structure is familiar: The marginal rate of substitution must equal the relative price of the two goods. Relative price is the household's subjective cost, which in this case is a mixture of the user cost formula $1 - (P_h'/P_h)[(1 - \delta^h)/(1 + r)]$ and the cost of the borrowing constraint. Notice that for prices $P_h = P_h'$ this user cost formula reduces to equation (4) without the terms for the transaction cost. Notice also that permanently changing all future real estate prices to the same level will have no effect on the $h/c$ ratio because house prices show up in the optimality conditions only as ratios of two consecutive house prices. One can then show that in the one-person economy the consumption path without the price change is still affordable after the price change. Also, because this path satisfies the new optimality conditions, the optimal behavior of the household is to leave the consumption path unchanged, meaning that a permanent house price change is completely neutral.

*Heterogeneity among households is the crucial ingredient of any model attempting to determine the distributional effects of housing policy.*

The only way house price changes can have nontrivial effects is if households are heterogeneous. Li and Yao (2005) use a model similar to Yang's, though modified in several dimensions. Most importantly, Li and Yao introduce a rental market in which households can attain housing services via two routes: either purchasing a home, with the purchase subject to a transaction cost, or renting it instead at a cost of $\alpha P_h$. One can think of the parameter $\alpha$ as the rental return per unit rented.

Li and Yao then simulate the economy after a permanent price change to housing. Not surprisingly, renters suffer from a house price appreciation. As opposed to the representative household mentioned above, renters have no positive wealth effect at all because, by definition, they do not own real estate. Renters are, however, hit with an increase in their future housing costs regardless of their future tenure decision: If they decide to buy a house, they have to pay more, and if they continue to rent, their cost per unit of rental goes up because the rental price is a fixed portion of the house price $P_h$. Li and Yao can even quantify the loss renters suffer. The loss in utility due to an increase in $P_h$ by two standard deviations (11.5 percent) is equivalent to permanently reducing all current and future consumption by about 4.5 percent.

The effect on renters is obvious and unambiguous, which is not necessarily the case for homeowners. Li and Yao show that house price appreciation can in fact be disadvantageous to some homeowners. The intuition for their result has to do with the pattern of homeownership displayed in Table 2 (on page 17). The average value of the primary residence increases for every age group until around age fifty. As mentioned before, this increase is most likely due to the binding borrowing constraint

that keeps young households from attaining consumption levels (both housing and nonhousing) implied by the permanent income hypothesis. Suppose a young household owns $100,000 worth of real estate and experiences a 20 percent increase in house prices. Instead of a $20,000 gain, the household may view this as a $20,000 loss if it was planning to upgrade to a larger house worth $200,000, which now costs $240,000, because the upgrade would now cost $120,000 instead of $100,000.

In conclusion, Li and Yao argue that a substantial portion of the population does not benefit from a house price appreciation. The break-even age for homeowners is at around age fifty, which incidentally is roughly the age at which households finish their upgrading. Older households gain substantially from house price appreciation, partly because many of them start moving into smaller houses or even become renters again and are thus able to cash out some of their capital gains. Even without downgrading, older households experience a welfare gain because the value of the bequest they leave to their heirs increases.

## Different Mortgage Contracts and the Tenure Decision of Young Households

One can interpret the paper by Chambers, Garriga, and Schlagenhauf (2005b) as a follow-up to their earlier study (2005a), which uses a life-cycle model much like the ones presented earlier to explore potential reasons for the path of the homeownership rate over the past forty years. Homeownership stayed roughly constant at 64.5 percent between 1965 and 1995 but then increased significantly to now almost 68 percent.

Chambers, Garriga, and Schlagenhauf (2005a) argue that the main reason for this increase is the availability of mortgage contracts requiring lower down payments. They simulate the model economy, once with a tighter and once with a looser down payment constraint, and determine that relaxing the down payment constraint can indeed quantitatively account for the increase in homeownership. Specifically, they study the effect of going from a 20 percent down payment constraint to an 80-15-5 combo loan—that is, an 80 percent first mortgage, a 15 percent second mortgage, and a 5 percent down payment.

How can the arrival of a new mortgage contract increase homeownership, especially among young agents? Recall that younger households, say, between the ages of twenty and forty, are characterized by three features: They are far less likely to be homeowners and have both lower income and lower net worth than the average population. Equation (4) reveals the main culprit for low homeownership among young households: Because of the binding borrowing constraint, the subjective housing cost for young agents is higher than the rental cost, discouraging young agents from buying. Reducing the tightness of the borrowing constraint (lowering parameter $\gamma$) is then the most direct way to encourage homeownership among those agents with a positive Lagrange multiplier $\mu$ (a binding borrowing constraint). Chambers, Garriga, and Schlagenhauf (2005a) show that lowering $\gamma$ from 20 percent in the benchmark to 5 percent under the 80-15-5 combo loan will indeed increase homeownership among young households and can quantitatively account for the rise in homeownership.

The great innovation in the two papers by Chambers, Garriga, and Schlagenhauf is that they model mortgages with much more care than in the rest of the literature. The other three papers discussed here—Yang (2005), Li and Yao (2005), and Jeske and Krueger (2005)—take one crucial shortcut when modeling mortgages: Households roll over mortgage debt every period, that is, they choose whatever amortization schedule suits them, constrained only by the borrowing constraint (in Yang and Li and Yao) or the interest rate dependence on the leverage ratio (in Jeske and Krueger). Chambers, Garriga, and Schlagenhauf, however, assume that a mort-

gage follows a fixed amortization schedule as it would for, say, a thirty-year fixed rate mortgage. One can think of a mortgage contract then as not only an interest rate $r$ and a down payment ratio $\gamma$ but rather a whole sequence of ratios over the lifetime of the loan. Suppose a thirty-year mortgage with an 80 percent loan-to-value ratio stipulates a sequence of $\gamma$ starting at 20 percent and increasing very slowly over the first couple of years (because initially most of the mortgage payment goes toward interest rather than principal). Toward the end of the loan more and more of the payment goes toward reducing the principal, and eventually $\gamma$ reaches 100 percent when the loan is paid off.

*Relaxing the down payment constraint can greatly increase homeownership among younger households because it reduces the subjective user cost.*

Another example would be a balloon loan, in which case $\gamma$ stays at 20 percent for the duration of the loan and after which the balloon payment comes due, which requires the household to either pay off the loan ($\gamma = 1$) with one large payment or refinance into another mortgage. Naturally, different loan designs imply different payment schedules. Since the model is detailed enough to mimic a whole range of different mortgage contracts, Chambers, Garriga, and Schlagenhauf (2005b) can study how different contract types affect the tenure decision, especially of young agents.

The conclusion of their paper is that mortgage contracts that tend to have low payments early on, such as a balloon mortgage, are most successful in relaxing the borrowing constraint. This conclusion is intuitive: Instead of relaxing just one borrowing constraint, a balloon mortgage has the same effect as relaxing the constraints over the whole life of the mortgage relative to a fixed rate mortgage, with the exception of the last period when the entire principal is due. In general, the authors show that mortgage contracts with an increasing payment schedule and thus lower $\gamma$ over time tend to encourage more young households to become homeowners but decrease homeownership among older households. This finding makes perfect sense: Young households experience an increasing earnings profile (see the first two panels of Figure 2), which makes a mortgage with an increasing payment schedule a good match for them. The opposite is true for middle-aged and older households whose income, at least on average, is stagnating or even decreasing, making a mortgage with an increasing payment schedule a mismatch to their earnings profile.[12]

Chambers, Garriga, and Schlagenhauf have therefore pointed to the great importance of the borrowing constraint in determining tenure decisions over the life cycle. Relaxing this constraint can greatly increase homeownership among younger households because it reduces the subjective user cost in equation (4) if $\mu > 0$. This effect occurs both for the borrowing constraint in the initial period when a household purchases a house as well as over the duration of the mortgage.

## The Effect of Subsidies

The papers discussed so far clearly specify a dual role of housing—namely, a consumption role ($h$ enters the utility function) and an investment role because housing is also a form of saving as long as the depreciation rate is below 100 percent. One

12. Notice that in the model the authors assume that only one single mortgage contract exists because adding multiple contracts into the same model involves too big a computational burden. This result should not interpreted as an argument against, say, balloon mortgages because they reduce ownership among older households. In reality, several types of mortgages are available, so older households can choose mortgages that suit them better than a balloon mortgage.

ingredient missing in the literature so far has been the fact that housing investment is risky. Jeske and Krueger (2005) set up a model in which, in addition to idiosyncratic labor income risk, households face an idiosyncratic house depreciation shock. This setup generates some realistic features: Most importantly, households may end up with negative equity on their home if the depreciation shock is large enough. This feature is one way to generate mortgage foreclosure, which has not been done before in the general equilibrium literature. Jeske and Krueger also assume that there is a deadweight loss from foreclosure, taking the following shape: If a property goes into foreclosure, the bank receives proportionally less than the value of the property. In reality, this foreclosure loss is substantial. Pennington-Cross (2004) estimates the deadweight loss of foreclosure to be 22 percent of the value of the property. This loss becomes crucial in the policy experiment later because a subsidy on mortgage rates will likely increase mortgage default as well as the deadweight loss.

*A main finding of Jeske and Krueger is that interest subsidies, if they are indeed passed on to households, can cause overinvestment in housing and larger mortgages.*

In addition, Jeske and Krueger allow households to choose any leverage ratio they desire. This feature differs from the papers studied so far, in which households face a sharp borrowing constraint that requires a minimum or even fixed down payment ratio. Jeske and Krueger assume that the interest rate households pay is a function that increases in the leverage ratio they are taking on. Specifically, a bank will price the foreclosure risk into the mortgage, and, if a household chooses higher leverage, a smaller depreciation shock will be required to trigger a default—that is, default is more likely—implying a higher interest rate. Jeske and Krueger must also make some simplifying assumptions. As opposed to the previously discussed papers, they do not consider life-cycle effects—that is, they assume that households are infinitely lived, and they abstract from transaction costs. Both assumptions are necessary to make the computational burden tractable.

Suppose that a government were to provide a subsidy that lowers mortgage interest. Jeske and Krueger try to study the macroeconomic effects of such a subsidy. The subsidy the authors have in mind has a counterpart in reality—government-sponsored enterprises (GSEs) like Fannie Mae and Freddie Mac, which are in the business of borrowing in the bond market to then purchase large portfolios of home mortgages. The GSEs receive a benefit from the federal government in the form of an implicit bailout guarantee.[13] The Congressional Budget Office (CBO) estimates that the subsidy is worth about 42 basis points; that is, GSEs can borrow at rates 0.42 percentage points below what other entities with a similar credit rating would have to pay in the bond market. Passmore, Sherlund, and Burgess (2004) estimate that only a small fraction—7 basis points—actually makes it to the homeowner while the GSEs keep the lion's share of the subsidy to pass on to their shareholders. Blinder (2004), on the other hand, defends the GSEs by estimating that they indeed reduce mortgage rates by almost the entire 42 basis points.

Jeske and Krueger (2005) do not attempt to judge which of the two studies is correct but rather do the following thought experiment: Supposing that the GSEs indeed pass on the entire 42 basis points to homeowners as they claim, then what is the effect on aggregate macroeconomic variables as well as the distribution of wealth and income?

A main finding of Jeske and Krueger is that interest subsidies, if they are indeed passed on to households, can cause overinvestment in housing and larger mortgages.

In fact, mortgage debt increases proportionally more than the housing stock, implying more leverage on housing. This finding, of course, is less than surprising: The subsidy makes housing cheaper and therefore causes more housing investment, and lower mortgage rates encourage more leverage. Thanks to the endogenous borrowing constraint, Jeske and Krueger are able to study one new aspect of housing policy, namely, how mortgage subsidies affect the aggregate level of mortgage default. They find that the subsidy increases the proportion of mortgages in default by one half, which in turn makes the deadweight loss 50 percent larger.

The next provocative question is, What are the distributional consequences of the subsidy? The subsidy has little effect on homeownership rates. In other words, the assistance goes mainly to existing homeowners—those households already well off due to high incomes and high net worth—which are able to afford even larger houses while poor households are almost unaffected in their tenure choice.

This result is consistent with the findings of the generic housing model outlined earlier, especially equation (4): Lowering interest rate $r$ and thus the user cost for housing $(r + \delta^h)/(1 + r)$ may have only a small effect on households that are constrained by the down payment condition.[14] The main part of their subjective user cost, after all, comes from the borrowing constraint (the second term on the right-hand side of equation [4]) lowering current consumption. In the Jeske and Krueger model, another reason why poor households will not necessarily invest in real estate and instead keep renting is that the rental rate is determined endogenously; that is, the rental rate is a market-clearing price ensuring that rental demand equals rental supply. With a mortgage subsidy that causes more investment in housing, the rental price will necessarily drop, making renting more attractive relative to housing. These results may disappoint the proponents of government assistance for mortgage financing. Either the subsidy in the form of implicit guarantees and lower interest rates goes to shareholders of GSEs or the interest rate reduction is passed on to homeowners but ends up in the hands of already well-off households while doing nothing to promote homeownership for poorer households.

The findings, by the way, are consistent with empirical work that suggests that mortgage rate reductions have only a marginal impact on homeownership rates while the borrowing constraint seems to play a much larger role.[15] Jeske and Krueger confirm these findings but are able to do so in a logical and coherent framework using general equilibrium instead of microsimulation techniques.

## Conclusion

Households in different age cohorts and different positions in the income and wealth distribution have very different homeownership ratios, real estate wealth, and mortgage debt. Likewise, renters and homeowners look very different in their path of income and asset holdings over the life cycle. This heterogeneity poses a challenge to researchers trying to answer questions regarding housing policy because one cannot study the effect of a policy merely on the average household. Studying housing in the framework of macroeconomic models therefore requires incorporating heterogeneity of households. Likewise, the size of the housing stock is too large to ignore general equilibrium effects.

13. See Frame and Wall (2002) for a survey on this topic.
14. One must use caution here: The interest rate itself depends on the leverage ratio, so households will not view it as exogenous as they did in the other three papers. The intuition, though, would be similar in a model with an interest rate as a function of the leverage ratio.
15. See Painter and Redfearn (2002) or Feldman (2001) for a summary of a variety of empirical studies.

Thanks to advances in computational techniques and faster computer hardware, economists are now able to write down and solve models detailed enough to accommodate housing. The conference hosted by the Federal Reserve Bank of Atlanta featured (among others) four papers, all of which are variations of the generic housing model described here even though they touch on a large variety of issues. Yang (2005) points out the importance of transaction costs and borrowing constraints in life-cycle models with housing. Li and Yao (2005) study the effect of house price changes. Chambers, Garriga, and Schlagenhauf (2005b) incorporate financial innovations in the form of different mortgage contracts. And Jeske and Krueger (2005) study the effect of a subsidy on mortgage finance. In each paper, one can deduce the main intuition from the simple equilibrium conditions.

## REFERENCES


Aiyagari, S. Rao. 1994. Uninsured idiosyncratic risk and aggregate saving. *Quarterly Journal of Economics* 109, no. 3:659–84.

Attanasio, Orazio, and Martin Browning. 1995. Consumption over the life cycle and over the business cycle. *American Economic Review* 85, no. 5:1118–37.

Attanasio, Orazio, and Guglielmo Weber. 1995. Is consumption growth consistent with intertemporal optimization? Evidence from the Consumer Expenditure Survey. *Journal of Political Economy* 103, no. 6:1121–57.

Blinder, Alan S. 2004. Letter on Federal Reserve working paper to CEO Raines from Dr. Alan S. Blinder, February 23. <www.fanniemae.com/commentary/pdf/022304.pdf> (September 13, 2005).

Blundell, Richard, Martin Browning, and Costas Meghir. 1994. Consumer demand and the life-cycle allocation of household expenditures. *Review of Economic Studies* 61, no. 206:57–80.

Board of Governors of the Federal Reserve System. 2005. Flow of funds statistics. June 9 release. <www.federalreserve.gov/releases/z1/> (September 13, 2005).

Chambers, Matthew, Carlos Garriga, and Donald Schlagenhauf. 2005a. Accounting for changes in the homeownership rate. Florida State University working paper.

———. 2005b. Mortgage contracts and housing tenure-duration decision. Paper presented at the Federal Reserve Bank of Atlanta conference "Housing, Mortgage Finance, and the Macroeconomy," May 19–20.

Feldman, Ron J. 2001. Mortgage rates, homeownership rates, and government-sponsored enterprises. Federal Reserve Bank of Minneapolis annual report, 3–23.

Fernandez-Villaverde, Jesus, and Dirk Krueger. 2002. Consumption and saving over the life cycle: How important are consumer durables? Proceedings of the 2002 North American summer meetings of the Econometric Society.

———. 2004. Consumption over the life cycle: Some facts from Consumer Expenditure Survey data. University of Pennsylvania working paper.

Frame, W. Scott, and Larry D. Wall. 2002. Financing housing through government-sponsored enterprises. Federal Reserve Bank of Atlanta *Economic Review* 87, no. 1:29–43.

Gourinchas, Pierre-Olivier, and Jonathan Parker. 2002. Consumption over the life cycle. *Econometrica* 70, no. 1:47–89.

Jeske, Karsten, and Dirk Krueger. 2005. Housing and the macroeconomy: The role of implicit guarantees for government sponsored enterprises. Paper presented at the Federal Reserve Bank of Atlanta conference "Housing, Mortgage Finance, and the Macroeconomy," May 19–20.

Li, Wenli, and Rui Yao. 2005. The life-cycle effects of house price changes. Paper presented at the Federal Reserve Bank of Atlanta conference "Housing, Mortgage Finance, and the Macroeconomy," May 19–20.

Painter, Gary, and Christian L. Redfearn. 2002. The role of interest rates in influencing long-run homeownership rates. *Journal of Real Estate Finance and Economics* 25, nos. 2–3:243–67.

Passmore, Wayne, Shane Sherlund, and Gillian Burgess. 2004. The effect of housing government-sponsored enterprises on mortgage rates. Federal Reserve Board Finance and Discussion Series No. 2005-6.

Pennington-Cross, Anthony. 2004. The value of foreclosed property: House prices, foreclosure laws, and appraisals. Office of Federal Housing Enterprise Oversight Working Paper 04-1, January.

Yang, Fang. 2005. Consumption along the life-cycle: How different is housing? Paper presented at the Federal Reserve Bank of Atlanta conference "Housing, Mortgage Finance, and the Macroeconomy," May 19–20.